Heroes and Martyrs series
....

The Heroes and Martyrs series: job descriptions for health care quality improvement professionals?

D Neuhauser
....

A review of the Heroes and Martyrs series so far

This issue brings the 14th in the ''Heroes and Martyrs of Quality and Safety'' series. More are to come. With a new editor of this journal on board, it is a good moment to define the content of the series. It is about people—present and past—who have found an opportunity to improve health, their efforts to carry out the needed improvements and to measure the results. These people are patient centred, careful with their evaluations, and committed to making changes. Not all of them succeed and some suffer the consequences, thereby becoming martyrs and our heroes. Each story highlights some aspect of the tryptic of quality improvement which is also referred to as customer mindedness, statistical mindedness, and organizational transformation.

These heroes are not just compassionate caregivers, or researchers whose story ends with publication, or health administrators and change agents. Our quality heroes undertake all three aspirations and need a combination of skills to actually improve quality.

PATIENT CENTREDNESS

Among our heroes, Ernest A Codman,[1] Florence Nightingale,[2] and Ignaz Semmelweis[3] clearly cared about the human suffering they encountered. They saw opportunities for improvement, measured outcomes, and advocated change with varying degrees of success. All the people portrayed in this series were impassioned human beings, such as Avedis Donabedian writing love poetry in his old age[4] and W Edwards Deming, an observant patient and composer.[5] In searching these historical records there is often a lack of description of the human side of our heroes.

THE EVIDENCE

These are stories of evidence, research design, and statistics including the controlled trial of Biblical Daniel,[6] Cotton Mather and the use of numbers,[7] the blinded evaluation of Mesmerism,[8] James Y Simpson's severity adjustments,[9] and the use of randomization.[10] Louis Pasteur's rabies vaccination was used to make a point about the importance of statistical process control.[11] The reader interested in the history of medical evidence per se is strongly urged to visit Sir Iain Chalmers' website (www.jameslindlibrary.org).

SYSTEM CHANGE

Many of our heroes came to grief when the results of their evidence collided with the interests of powerful organizations. Bruce Psaty,[12] John Williamson,[13] and Dwain Harper[14] inadvertently suffered the consequences. Semmelweis and Codman were spectacularly incompetent agents for change. Nightingale and Deming were geniuses at change—focus and constancy of purpose being their most powerful levers.

We had planned that the 14th article in this series would be about Dr Betty Dong whose negative evaluation of a drug was criticised by its manufacturer. This piece was to be a reprint from the new book by Drs Deyo and Patrick, ''Hope and Hype'', but the journal's legal advisors recommended that this report should not be published because it might be construed as libellous.

Interested readers can read about Dr Dong in this newly published book.[15] I recommend it.

This series is therefore a collection of stories about passionate and compassionate people improving health, changing care, and measuring their results. Taken as a whole, it is a painless textbook on research methods. The series defines a curriculum and job description for present and future healthcare quality improvement professionals.

Correspondence to: Professor D Neuhauser, Department of Epidemiology and Biostatistics, Case School of Medicine, Case Western Reserve University, Cleveland, OH 44106-4945, USA; dvn@case.edu

REFERENCES

1 **Neuhauser D**. Ernest Amory Codman, MD. *Qual Saf Health Care* 2002;**11**:104–5.
2 **Neuhauser D**. Florence Nightingale gets no respect: as a statistician that is. *Qual Saf Health Care* 2003;**12**:317.
3 **Best M**, Neuhauser D. Ignaz Semmelweis and the birth of infection control, Vienna, 1847. *Qual Saf Health Care* 2004;**13**:233–4.
4 **Best M**, Neuhauser D. Avedis Donabedian, father of quality assurance and poet. *Qual Saf Health Care* 2004;**13**:472–3.
5 **Best M**, Neuhauser D. W Edwards Deming: father of quality management, patient and composer. *Qual Saf Health Care* 2005;**14**:000–0.
6 **Neuhauser D**, Diaz M. Daniel: using the Bible to teach quality improvement methods. *Qual Saf Health Care* 2004;**13**:153–5.
7 **Best M**, Neuhauser D, Slavin L. ''Cotton Mather, you dog, dam you! I'll inoculate you with this; with a pox to you'': smallpox inoculation, Boston, 1721, *Qual Saf Health Care* 2004;**13**:82–3.
8 **Best M**, Neuhauser D, Slavin L. Evaluating Mesmerism: Paris 1784. *Qual Saf Health Care* 2003;**12**:232–3.
9 **Neuhauser D**. Surgical experience, hospital size and severity adjusted mortality. James Y Simpson, 1869. *Qual Saf Health Care* 2005;**14**:67–8.
10 **Diaz M**, Neuhauser D. Shuffle the deck, flip that coin: randomization comes to medicine. *Qual Saf Health Care* 2004;**13**:315–6.
11 **Diaz M**, Neuhauser D. Pasteur and parachutes. When statistical process control is better than a randomized clinical trial. *Qual Saf Health Care* 2005;**14**:140–3.
12 **Deyo R**. Bruce Psaty and the risks of calcium channel blockers. *Qual Saf Health Care* 2002;**11**:294–6.
13 **Neuhauser D**. John Williamson and the terrifying results of the Medical Practice Demonstration Project. *Qual Saf Health Care* 2002;**11**:387–9.
14 **Neuhauser D**, Harper D. Too good to last: did Greater Cleveland Health Quality Choice leave a legacy and lessons to be learned? *Qual Saf Health Care.* 2002;11: 202–3; 391).
15 **Deyo R**, Patrick D. *Hope or Hype. The obsession with medical advances and the high cost of false promises.* New York: AMACOM, 2005.

Assessing safety culture

....

Assessing safety culture: guidelines and recommendations

P Pronovost, B Sexton

....

A step nearer to the reliable measurement of safety culture

''The journey of a thousand miles begins with one step.'' Latsu

Safety culture is increasingly recognized as an important strategy—and perhaps a necessary precursor—to improving the widespread deficits in patient safety. The Joint Commission for Accreditation of Healthcare Organizations (JCAHO) included an annual assessment of safety culture in its 2007 patient safety goals. The Institute of Medicine (IOM) report ''To Err is Human'' spurred healthcare organizations to implement initiatives that improve patient safety.[1]

Despite this, culture is defined and measured in various ways. Definitions of culture commonly refer to values, attitudes, norms, beliefs, practices, policies, and behaviors of personnel. In essence, culture is ''the way we do things around here'', whereby the word ''here'' refers not to the hospital, but rather to a particular work unit. In a safe culture employees are guided by an organization-wide commitment to safety in which each member upholds their own safety norms and those of their co-workers. A number of tools are available to measure safety culture, but each instrument has unique domains of culture, limited validity and reliability data, and average response rates that vary from poor (29%) to excellent (83%). The science of measuring safety culture is evolving, even as the demand for rigorous cultural assessment intensifies.

MEASUREMENT OF SAFETY CULTURE

Many organizations have embarked on efforts to measure safety culture. It is not uncommon for senior leaders in these hospitals to use culture survey scores as a system level measure of patient safety to hold managers accountable, often with the use of bonuses. Although these efforts are laudable, the enthusiasm for measuring culture may be outpacing the science. Due perhaps to the nascent nature of cultural assessment in health care, culture researchers lack consensus and clarity about

domains important in a culture of safety; how to score and present improvements in culture over time; the relationship between culture and clinical and operational outcomes; and ultimately, how to package a tool kit to measure, score, and improve culture.

An important and perhaps glaring gap in our knowledge of cultural assessment is understanding the sources of variation in culture—that is, we do not understand whether staff characteristics, the patient care area, the department (where applicable), or the hospital explain variation in culture. We must understand these sources of variation in order to target who to measure, how to score, where to focus efforts to improve culture, and to hold accountable for improving culture.[2] Failure to understand these important issues can cause managers to make incorrect inferences regarding scores on cultural assessments, and potentially cause additional harm by either rewarding the wrong behaviors or diverting scarce resources away from important efforts.

A first step in this effort to measure safety culture is to ensure that the survey instruments are valid and reliable—that is, that they measure what they intend to measure and produce similar results upon repeat measurement. The paper by Kho and colleagues[3] published in this issue of *QSHC* does much to advance our science of measuring culture. They adapted and administered the Safety Climate Survey to measure safety culture in four Canadian University affiliated intensive care units (ICUs). Using a novel technique, they achieved a 75% response rate and evaluated the validity and reliability of three different scoring methodologies. They measured internal consistency to estimate validity, and to determine reliability they measured test-retest reliability. They found that the Safety Climate Survey overall (22 items) and the 13-item scale had construct validity and sufficient reliability while the 7-item scale lacked construct validity and, as such, was not considered further. The high score on internal

consistency suggests that the questions on the full and 13-item subset measure a single construct—safety culture. The authors are to be applauded for their rigorous evaluation of construct validity and reliability of these scoring methodologies. They add important new knowledge on how to measure culture.

ORIGIN OF SAFETY CLIMATE SURVEY

It is important to recognize the origin of the Safety Climate Survey used by Kho *et al*. In 2002 we extracted a subset of items related to safety climate from the larger Safety Attitudes Questionnaire.[4] This subset of safety climate items did not elicit attitudes along any dimensions such as teamwork climate or perceptions of management. We provided a copy of this survey with instructions and comparison data to the Institute for Healthcare Improvement for posting on their new website (www.qualityhealthcare.org).

Our own analyses of the Safety Climate Survey scoring methodologies showed test-retest reliability of 0.85–0.90 and Cronbach α values of 0.75–0.88.[5] These psychometrics are sound, yet they tell us nothing about *which* domains of culture are most appropriate. In our experience with the multidimensional Safety Attitudes Questionnaire, we have 30 items that measure six domains: safety climate, teamwork climate, perceptions of management, stress recognition, job satisfaction, and working conditions. We are only beginning to appreciate what we gain from multidimensional cultural assessments. Clearly, the additional dimensions are informative and allow feedback to hospitals and work units to consider a number of their cultural strengths and weaknesses.

When we feed back the results of these surveys, which we have now done at over 500 hospitals, caregivers and managers generally find the multidimensional cultural feedback and comparison data highly informative. We have shown that focused interventions can improve the safety and teamwork climate in a given work unit, yet these often improve at the expense of another cultural dimension—namely, stress recognition. For example, as work units evolve to be highly efficient with greater trust, collaboration, and openness, caregivers sometimes develop a sense of personal invulnerability as a by-product of working on a stellar unit. Unfortunately, this reduction in acknowledging stressors is a dark path that leads otherwise outstanding work units to host seemingly unexplainable sentinel events. In other words, we need the more diagnostic multidimensional
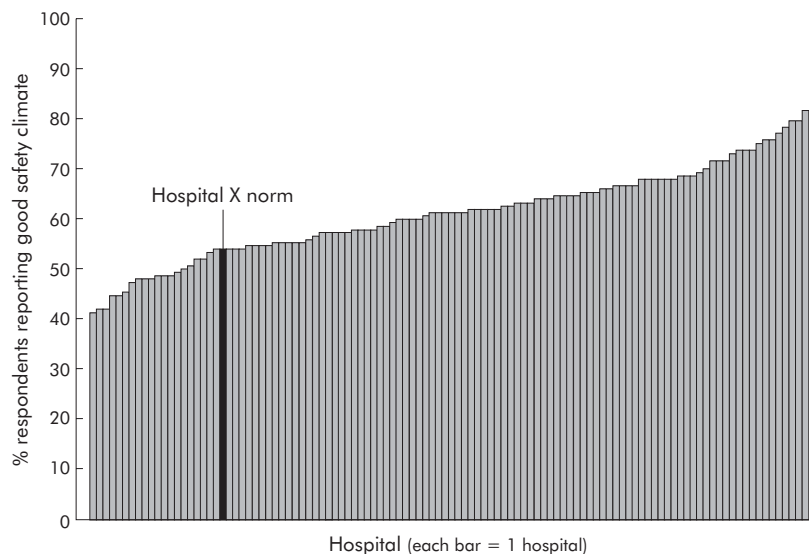
Figure 1 Safety climate across 100 hospitals.

cultural assessment tools to have the ability to track units on a variety of strengths and weaknesses. The Safety Attitudes Questionnaire has good construct validity and internal consistency, yet there is a tremendous amount of work still to be done on criterion validity for each domain which would show how dimensions of culture link to clinical and operational outcomes. This is an active area of research.

RECOMMENDATIONS FOR THE FUTURE

What would we recommend for healthcare organizations interested in measuring safety culture? Our recommendations are informed by our prior mistakes and continued research. We would use the full Safety Attitudes Questionnaire (rather than just safety

climate) and measure the entire hospital annually; this has already been done at the Johns Hopkins Hospital. We found that when you measure and feedback data in one work unit, other units quickly desire their own cultural assessment as well. The Safety Attitudes Questionnaire is the most widely used cultural assessment tool in health care to date. In the past 12 months we have assessed culture in over 100 hospitals with an average hospital-wide response rate of over 80%. Representativeness is critical as it makes the data easy to interpret and difficult to ignore. This is particularly true in pre-post or longitudinal cultural assessments where high response rates are essential to interpreting data over time. When response rates fall below 60%, the data represent opinions rather than culture

and the results should be used with caution.

In addition, the measurement of culture should include a presentation of results to staff as well as senior management, followed by a focused intervention to improve culture. There is limited evidence regarding interventions that improve culture. To our knowledge, the Comprehensive Unit-based Safety Program (CUSP) is the only published intervention that has been shown to improve culture.[6] Even with a valid measure of culture, if culture is not responsive to interventions there is no point in measuring it. Although far from perfect, the CUSP provides a practical framework for improving patient safety (culturally, clinically, and operationally) throughout an entire organization by focusing on individual work units and respecting the wisdom of their frontline providers.

Why focus on the unit level? Culture is local. Intervening in culture requires focusing at the local work unit level. Figure 1 shows the safety climate for 100 hospitals where the hospital level safety climate ranges from 40% to about 80% positive. We pulled out hospital X from fig 1 to show the typical variability within a hospital across the work units. Figure 2 shows that, within hospital X, positive safety climate scores range from 0% to 100% among work units. With few exceptions, we find more variability between work units within a hospital than we do between hospitals. For this empirical reason, it is critical to assess culture across all work units in an institution.

The research by Kho et al[3] provides us with new confidence in the ability to move towards measuring safety culture with methodological rigor. Future research is needed to evaluate whether measuring additional domains such as teamwork climate, perceptions of management, or stress recognition is useful. Let us hope that this well written and rigorously conducted paper is an early step in a long journey towards understanding safety culture and ultimately improving patient safety.

. . . .
Authors' affiliations
P Pronovost, B Sexton, Department of Anaesthesiology and Critical Care Medicine, The Johns Hopkins University, Baltimore, MD, USA

Correspondence to: Associate Professor P J Pronovost, Department of Anaesthesiology and Critical Care Medicine, The Johns Hopkins University, 600 North Wolfe Street, Meyer 295, Baltimore, MD 21287, USA; ppronovo@jhmi.edu

Figure 2 Safety climate across 49 work units in one hospital.

REFERENCES

1 **Kohn L**, Corrigan J, Donaldson M, eds. *To err is human: building a safer health system*, Institute of Medicine Report. Washington, DC: National Academy Press, 1999.
2 **Sexton JB**, Thomas EJ, Helmreich RL. Error, stress, and teamwork in medicine and aviation: cross sectional surveys. *BMJ* 2000;**320**:745–9.
3 **Kho ME**, Carbone JM, Lucas J, *et al*. Safety Climate Survey: reliability of results from a multicenter ICU survey. *Qual Saf Health Care* 2005;**14**:273–8.
4 **Sexton JB**, Thomas EJ, Helmreich RL, *et al*. *Frontline assessments of healthcare culture: Safety Attitudes Questionnaire norms and psychometric properties*, Technical Report 04-01. The University of Texas Center of Excellence for Patient Safety Research and Practice, 2004. AHRQ Grant #1PO1HS1154401. Available at http://www.utpatientsafety.org.
5 **Shtyenberg G**, Sexton JB, Thomas EJ. *Test retest reliability of the safety climate scale*, Technical Report 01-05. The University of Texas Center of Excellence for Patient Safety Research and Practice, 2005. AHRQ grants #1PO1HS1154401 and U18HS1116401. Available at http://www.utpatientsafety.org.
6 **Pronovost PJ**, Weast B, Rosenstein B, *et al*. Implementing and validating a comprehensive unit-based safety program. *J Patient Safety* 2005;**1**:33–40.

Clinical Evidence—Call for contributors

Clinical Evidence is a regularly updated evidence-based journal available worldwide both as a paper version and on the internet. *Clinical Evidence* needs to recruit a number of new contributors. Contributors are healthcare professionals or epidemiologists with experience in evidence-based medicine and the ability to write in a concise and structured way.

Areas for which we are currently seeking authors:
- Child health: nocturnal enuresis
- Eye disorders: bacterial conjunctivitis
- Male health: prostate cancer (metastatic)
- Women's health: pre-menstrual syndrome; pyelonephritis in non-pregnant women

However, we are always looking for others, so do not let this list discourage you.

Being a contributor involves:
- Selecting from a validated, screened search (performed by in-house Information Specialists) epidemiologically sound studies for inclusion.
- Documenting your decisions about which studies to include on an inclusion and exclusion form, which we keep on file.
- Writing the text to a highly structured template (about 1500–3000 words), using evidence from the final studies chosen, within 8–10 weeks of receiving the literature search.
- Working with *Clinical Evidence* editors to ensure that the final text meets epidemiological and style standards.
- Updating the text every six months using any new, sound evidence that becomes available. The *Clinical Evidence* in-house team will conduct the searches for contributors; your task is simply to filter out high quality studies and incorporate them in the existing text.
- To expand the topic to include a new question about once every 12–18 months.

If you would like to become a contributor for *Clinical Evidence* or require more information about what this involves please send your contact details and a copy of your CV, clearly stating the clinical area you are interested in, to Klara Brunnhuber (kbrunnhuber@bmjgroup.com).

Call for peer reviewers

Clinical Evidence also needs to recruit a number of new peer reviewers specifically with an interest in the clinical areas stated above, and also others related to general practice. Peer reviewers are healthcare professionals or epidemiologists with experience in evidence-based medicine. As a peer reviewer you would be asked for your views on the clinical relevance, validity, and accessibility of specific topics within the journal, and their usefulness to the intended audience (international generalists and healthcare professionals, possibly with limited statistical knowledge). Topics are usually 1500–3000 words in length and we would ask you to review between 2–5 topics per year. The peer review process takes place throughout the year, and our turnaround time for each review is ideally 10–14 days.

If you are interested in becoming a peer reviewer for *Clinical Evidence*, please complete the peer review questionnaire at www.clinicalevidence.com or contact Klara Brunnhuber (kbrunnhuber@bmjgroup.com).

BMJ Quality & Safety

Assessing safety culture: guidelines and recommendations

P Pronovost and B Sexton

Updated information and services can be found at:

http://qualitysafety.bmj.com/content/14/4/231.full.html

These include:

References

This article cites 2 articles, 2 of which can be accessed free at:
http://qualitysafety.bmj.com/content/14/4/231.full.html#ref-list-1

Article cited in:
http://qualitysafety.bmj.com/content/14/4/231.full.html#related-urls

Email alerting service

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/