The problem with ‘5 whys’

Alan J Card

Correspondence to
Dr Alan J Card, Evidence-Based Health Solutions, LLC, PO Box 62, Notre Dame, IN 46556, USA; alan.j.card@gmail.com

Accepted 13 August 2016
Published Online First 2 September 2016

To cite: Card AJ. *BMJ Qual Saf* 2017;**26**:671–677.

‘The Problem with…’ series covers controversial topics related to efforts to improve healthcare quality, including widely recommended but deceptively difficult strategies for improvement and pervasive problems that seem to resist solution.

BACKGROUND

The ‘5 whys’ technique is one of the most widely taught approaches to root-cause analysis (RCA) in healthcare. Its use is promoted by the WHO,[1] the English National Health Service,[2] the Institute for Healthcare Improvement,[3] the Joint Commission[4] and many other organisations in the field of healthcare quality and safety. Like most such tools, though, its popularity is not the result of any evidence that it is effective.[5–8] Instead, it probably owes its place in the curriculum and practice of RCA to a combination of pedigree, simplicity and pedagogy.

In terms of pedigree, ‘5 whys’ traces its roots back to the Toyota Production System (TPS).[9] It also plays a key role in Lean[10] (a generic version of TPS) as well as Six Sigma,[11] another popular quality improvement (QI) methodology. Taiichi Ohno describes ‘5 whys’ as central to the TPS methodology:

> The basis of Toyota’s scientific approach is to ask why five times whenever we find a problem … By repeating why five times, the nature of the problem as well as its solution becomes clear. The solution, or the how-to, is designated as ‘1H.’ Thus, ‘Five whys equal one how’ (5W=1H). (ref. 9, p. 123)

This quote also makes the case for the technique’s simplicity. Asking ‘why’ five times allows users to arrive at a single root cause that might not have been obvious at the outset. It may also inspire a single solution to address that root cause (though it is not clear that the ‘1H’ side of the equation has been adopted as widely).

‘5 WHYS’ AS A TEACHING TOOL

The pedagogical argument for ‘5 whys’ is that it creates an ‘aha moment’ by revealing the hidden influence of a distant cause, which illustrates the importance of digging deeper into a causal pathway. This quick and easy learning experience can be a powerful lesson in systems safety and QI.

Possibly the most famous ‘5 whys’ case study to be used in this way focuses on efforts to preserve the Washington Monument.[12 13] Details vary slightly depending on the source, but it usually looks something like this:

Problem: The Washington Monument is deteriorating

> **Why?** Harsh chemicals are being used to clean the monument
>
> **Why?** The monument is covered in pigeon droppings
>
> **Why?** Pigeons are attracted by the large number of spiders at the monument
>
> **Why?** Spiders are attracted by the large number of midges at the monument
>
> **Why?** Midges are attracted by the fact that the monument is first to be lit at night.

Solution: Turn on the lights one hour later.

This is a great teaching example because the ‘root cause’ is so unintuitive. Who would think, before exploring the issue in depth, that lighting choices could endanger a marble monument? But, as is so often the case, reality is messier than this simple illustration.

Joel Gross[12] investigated the foundation of this example and discovered that many of the details are incorrect. And, crucially, the broader story it tells is incomplete.

In terms of the story’s details, the monument is question was actually the Lincoln Memorial, and it was not being damaged by the use of harsh chemicals. The real culprit was simply water. Pigeons were not an issue at all, and while there were ‘tiny spiders’ (ref. 14, p. 8) at the memorial, they were not a major problem. Instead, most of the cleaning was necessary because swarms of midges were dazzled by the lights and flew at high speed into the walls of the memorial, leaving it splattered with bits of the insects and their eggs.[12 14]

But that only speaks to the details that *were* described. The analysis is also incomplete in a number of more important ways. For instance, it only addresses one potential source of deterioration: cleaning water.

The first 'why' could just as easily have tackled other causes, such as rain or acid rain (a significant concern at the time), rising damp, erosion from windborne particles or damage from freeze-thaw cycles.[15] Or, if the goal had been to prevent harm to future monuments, the first 'why' could have focused on the use of marble as a building material, the choice of building site, etc.

However, the most important problem with this example is that, while the solution was 'effective' in one sense, it still failed:

> Messersmith [the consultant entomologist who worked on this project] thought that because the insects swarmed only at sunset, a one-hour delay in turning on the monument lights would go far in solving the problem. The technique worked, reducing the number of midges in the monuments by about 85 percent.
>
> 'But tourists who had driven hundreds of miles to have their photographs taken at the monuments were not happy,' he said. 'They complained every day, and the lights went back on.'[16]

The logic of the solution was sound, as far as it went. But it was predicated on an incomplete understanding of the broader system, its stakeholders and the purpose of the monument itself. If anything, this window on the complexity of real-world problem solving adds to the value of this teaching example. If the first 'aha moment' is followed by this second one, trainees will not only learn that distal causes can have unexpected outcomes, but also that systems thinking requires both depth *and breadth* of analysis.

THE PROBLEM WITH '5 WHYS' IN RCA

'5 whys' has been the subject of a number of caveats and critiques. For instance, Minoura, one of Ohno's successors at Toyota, highlights the potential for users to rely on off-the-cuff deduction, rather than situated observation when developing answers, as well as difficulty in prioritising causes, if multiple '5 whys' are used.[17] Mark Graban, a thought leader in the Lean community, points out that '5 whys' is just one component of what should be a far more comprehensive problem-solving process.[18] And Serrat clarifies that users should not feel constrained by the arbitrary number in the tool's title: more, or fewer, than five 'whys' may be required.[19]

But the real problem with '5 whys' is not *how* it is used in RCA, but rather that it so grossly oversimplifies the process of problem exploration that it should not be used at all. It forces users down a single analytical pathway for any given problem,[13] insists on a single root cause as the target for solutions[9 13 20] and assumes that the most distal link on the causal pathway (the fifth 'why') is inherently the most effective and efficient place to intervene.

A single causal pathway

A credible '5 whys' approach to a wrong patient medication error might look like this (adapted from Battles *et al*):[21]

Incident: Wrong patient medication error
  Why? Wristband not checked
  Why? Wristband missing
  Why? Wristband printer on the unit was broken
  Why? Label jam
  Why? Poor product design

But another team could easily come up with five wholly different and equally valid 'whys'. And any single string of '5 whys' can provide only a blinkered view of the complex causal pathway that led to the incident. This is illustrated by figure 1, a causal tree diagram (or, more accurately, a 'causal and contributing factors tree diagram') depicting the underlying issues that gave rise to the adverse event.

It is clear from the tree diagram that the causal pathway related to the wristband printer is neither the *only* relevant cause of the incident nor indisputably the most important. A serious effort to solve the myriad problems that gave rise to this incident would have to tackle a number of other causal pathways as well.

These might include pathways related to a maladaptive workplace culture,[22–25] clinical and information technology (IT) staffing, orientation of agency staff and the absence of a forcing function[26] to ensure that patients are properly identified before medication is administered. It could also include a focus on improved infection control and better preparedness for infectious disease outbreaks. Solutions based on the '5 whys' in the example above would leave all of these issues unaddressed.

There is also no objective or reliable means of mapping out the causal pathway, which is a critical failing when only one pathway will be examined. Consider the variant below, which follows essentially the same causal reasoning as the first example:

Incident: Wrong patient medication error
  Why? Wristband missing
  Why? Wristband printer on the unit was broken
  Why? Healthcare system purchased an unreliable printer
  Why? Poor process for evaluating and purchasing 'non-clinical' equipment
  Why? Equipment deemed 'non-clinical' is not seen as safety-critical

This version skips the step of asking why the wristband was not checked and moves directly to asking

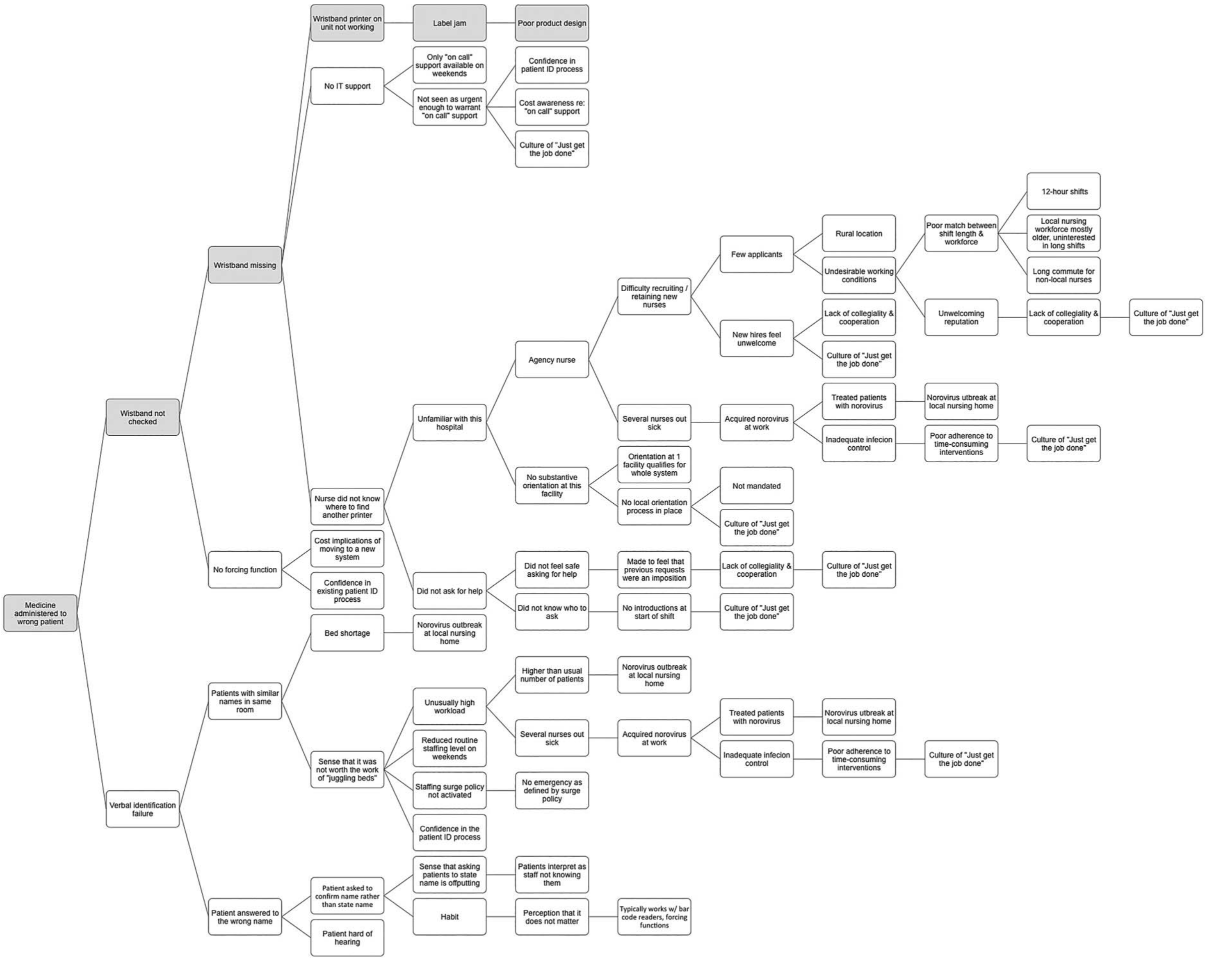
Medicine administered to wrong patient
Wistband not checked
Verbal identification failure
Wristband missing
No forcing function
Patients with similar names in same room
Patient answered to the wrong name
Wristband printer on unit not working
No IT support
Nurse did not know where to find another printer
Cost implications of moving to a new system
Confidence in existing patient ID process
Bed shortage
Sense that it was not worth the work of "juggling beds"
Patient asked to confirm name rather than state name
Patient hard of hearing
Label jam
Only "on call" support available on weekends
Not seen as urgent enough to warrant "on call" support
Unfamiliar with this hospital
Did not ask for help
Norovirus outbreak at local nursing home
Unusually high workload
Reduced routine staffing level on weekends
Staffing surge policy not activated
Confidence in the patient ID process
Sense that asking patients to state name is offputting
Habit
Poor product design
Confidence in patient ID process
Cost awareness re: "on call" support
Culture of "Just get the job done"
Agency nurse
No substantive orientation at this facility
Did not feel safe asking for help
Did not know who to ask
Higher than usual number of patients
Several nurses out sick
No emergency as defined by surge policy
Patients interpret as staff not knowing them
Perception that it does not matter
Difficulty recruiting / retaining new nurses
Several nurses out sick
Orientation at 1 facility qualifies for whole system
No local orientation process in place
Made to feel that previous requests were an imposition
No introductions at start of shift
Norovirus outbreak at local nursing home
Acquired norovirus at work
Typically works w/ bar code readers, forcing functions
Few applicants
New hires feel unwelcome
Acquired norovirus at work
Not mandated
Culture of "Just get the job done"
Lack of collegiality & cooperation
Culture of "Just get the job done"
Treated patients with norovirus
Inadequate infecion control
Rural location
Undesirable working conditions
Lack of collegiality & cooperation
Culture of "Just get the job done"
Treated patients with norovirus
Inadequate infecion control
Culture of "Just get the job done"
Norovirus utbreak at local nursing home
Poor adherence to time-consuming interventions
Poor match between shift length & workforce
Unwelcoming reputation
Norovirus utbreak at local nursing home
Poor adherence to time-consuming interventions
Culture of "Just get the job done"
12-hour shifts
Local nursing workforce mostly older, uninterested in long shifts
Long commute for non-local nurses
Lack of collegiality & cooperation
Culture of "Just get the job done"
Culture of "Just get the job done"

Figure 1 A causal event tree (adapted from ref. 21). ID, identification; IT, information technology.

why it was not there. It also sticks to the high-level issue of the printer being broken, without delving into the details of the label jam. 'Skipping' these questions allows the analysis to go deeper because it leaves more 'whys' available. This example also maintains a focus on issues *within* the organisation, rather than the design of the printer. This would lead to very different solutions.

But because this approach skips past the question of why the wristband was not checked, it closes the door to questions about other reasons why it was not checked. In figure 1, this would include the lack of a forcing function. But in another scenario, it might include a desire to avoid waking the patient;[27] an unreadable wristband (eg, smudged, crinkled or occluded);[28] the lack of a label on the medication;[29] confusion caused by multiple wristbands;[30] lack of trust in the wristband data due to frequent errors[31] or any of a number of other causes.[28–31]

Users could also go down an entirely different causal pathway. An equally reasonable '5 whys' for this incident could look like this:

Incident: Wrong patient medication error

Why? Patients with similar names in the same room

Why? Not feasible to try 'juggling beds'

Why? Not enough nurses to deal with the influx of patients

Why? Nurses affected by an outbreak of norovirus

Why? Poor adherence to time-consuming infection control interventions

Why? A culture of 'just get the job done'

There are many 'correct' ways a team might use '5 whys' to assess even this one incident. And it is unlikely that any two teams would independently arrive at exactly the same results. This subjectivity is critically important because '5 whys' focuses on only one root cause at the end of one causal pathway.

More sophisticated practice in the use of '5 whys' might produce two causal pathways, focusing on the main service failures uncovered, rather than the event itself (ie, a set of '5 whys' for 'wristband not checked' and another for 'verbal identification failure'). But this is not how use of the tool has generally been taught in the healthcare industry.[1 3 32] And even this unusually thorough approach would identify only 2 of the 30 causal pathways shown in the tree diagram.

A single root cause

Forcing users down a single causal pathway should be disqualifying by itself. But '5 whys' narrows the scope for improvement even further by insisting that risk control efforts must focus on a single root cause for each causal pathway. In the first healthcare example above, for instance, the root cause would be 'poor product design', and this would serve as the sole target for improvement efforts.

But accidents are seldom the result of a single root cause.[33] So focusing exclusively on one (or even a few) arbitrarily determined 'root causes' is not a reliable method for driving improvement—especially in a system as complex as healthcare. As Wieman and Wieman wrote: "Unfortunately … restricting the number of variables [considered] in a complex system only results in an increased potential for errors of omission" (ref. 34, p. 117).

How much might be omitted when using '5 whys'? The tree diagram for our example uncovers more than *75 whys* (causes and contributing factors), each of which is a potential target for action to reduce the risk of a recurrence. The '5 whys' approach would identify only one (or possibly two) root cause as target for action. At best, this represents <3% of the opportunities for improvement identified using the tree diagram.

Targeting only the most distal cause

Not only are users of '5 whys' limited to one root cause per causal pathway, but they are also limited to selecting only the most distal cause (conventionally, the fifth 'why'). There is, however, no logical reason to assume that this is always the most effective or most efficient target for intervention.

Actually, if it were possible to magically place a 100% effective risk control at any one point on the tree diagram, it would be best used on a *proximate* cause. For instance, making it impossible to administer medication without checking the wristband would render all the more distal causes moot for the purpose of preventing a recurrence.

And, while 100% effective risk controls are seldom available, an action plan that includes a proven[26] (if certainly imperfect)[29] intervention like a well-designed bar-code reader with a forcing function for patient identification (ID) is more likely to prevent another serious 'wrong patient' medication error than switching to a well-designed printer.

This is not to suggest that more distal causes are not appropriate targets for improvement efforts. In the example presented in figure 1, for instance, there is clearly a profound need to change the culture from one that is task-oriented and sometimes hostile to one that is outcomes-oriented and psychologically safe. The pervasive impact of such a culture change would be far more important than merely reducing the risk that this particular incident might recur; it would influence almost every quality and safety issue in the organisation.

And, in contrast to that powerful-but-difficult lever for shifting outcomes, sometimes more distal causes represent 'low-hanging fruit' that can be addressed while a more proximate solution is in the works. In figure 1, educating patients about why clinicians will be constantly asking them to identify themselves would be far from foolproof. But it *would* be fast, cheap and easy. And it might reduce an important barrier to best practice in verbal identification.

Appropriate targets for intervention may occur anywhere along the causal continuum and on any causal pathway. And efforts to improve safety and quality will often require more than one intervention targeting more than one underlying hazard. It is useful to identify all the key hazards that gave rise to an incident and ensure that each of these is either addressed or intentionally accepted.[35] (See, for instance, the Options Evaluation Matrix.)[36] But the use of '5 whys' makes this impossible.

CONSIDERING THE VIRTUES OF '5 WHYS'

What, then, of the virtues of '5 whys?' Are the issues above redeemed by the tool's simplicity and pedigree?

Simplicity

Simplicity is a complicated virtue when it comes to the frameworks, tools and techniques of QI. For instance, the conceptual simplicity of the Plan-Do-Study-Act (PDSA) framework is one of its main selling points, but it may also lead organisations to underestimate the messy work involved in applying PDSA to real-world problems.[37]

But, as this paper has shown, the '5 whys' approach has clearly overshot the mark: it is not simple, but *simplistic*. It is, as Leveson describes, "… perhaps the most simplistic [accident analysis technique] and … leads to the least amount of learning from events".[13]

Charles Vincent famously called for RCA to serve as "a window on the system".[38] If that is the goal, then '5 whys' is doomed to fail. It purposely discards the vast majority of what might be learned about the system being interrogated.

The reality of most healthcare processes and systems is that we face classic design problems: problems that are highly contextualised and often "ill-defined, ill-structured, or 'wicked'"[39] (ref. 40, p. 224). A '5 whys' analysis ignores this. Most of the causal pathways that led to an event are amputated from the start, and consideration of those that remain is limited to a single root cause.

This creates a *toy problem*[i] in which it is assumed that simple optimisation of one, or at most a handful of variables will lead to improvement, without any need to consider the rest of the system. This flies in the face of everything we know about solving problems in complex adaptive systems like healthcare.[41]

---

[i] A toy problem is: "A deliberately oversimplified case of a challenging problem used to investigate, prototype, or test algorithms for a real problem. Sometimes used pejoratively".[62]

Pedigree

The positive reputation enjoyed by TPS/Lean provides an aura of credibility for '5 whys'. But how applicable is this to the question at hand? The reputation of TPS/Lean was built in a very different context. And the use of '5 whys' as an RCA tool is by no means the same thing as the use of the full TPS/Lean methodology.

Healthcare organisations are not automobile factories. And while there is much to be learned from the automotive industry and other high-reliability organisations (HROs), healthcare delivery will never be truly comparable to automobile manufacturing. Despite efforts in the healthcare industry to adopt the tenets of HROs,[42 43] current practice provides recommended care only about 70% of the time.[44] And the percentage of hospital patients who experience an adverse event may be as high as 25–33%.[45–50]

HROs commonly aim for a reliability rate of 'six sigma' (three errors per million opportunities). By these measures, healthcare is struggling to move beyond *two sigma* (308 500 errors per million opportunities, or a 30.85% error rate).[51]

Reliability in the healthcare industry can improve, and indeed it has (cf. ref. 52). But healthcare is far more complex[34 53] than automobile manufacturing, and takes place amid processes and systems that are woefully underdesigned in comparison to a modern factory. Further, the safety and quality workforce in healthcare is only beginning to move towards professionalisation[54 55] and often lacks formal training in engineering, human factors, ergonomics or similar domains.

As a result, approaches developed for solving problems in the automotive manufacturing context may not be as effective in the healthcare arena. And, indeed, the evidence base for the use of Lean/TPS in healthcare is weak[56–58] and increasingly negative.[59 60]

It is also important to differentiate between the use of '5 whys' as a QI *method* and the use of Lean/TPS as a QI *methodology*. Though the two are sometimes conflated in both the literature[59] and practice,[61] they are by no means equivalent. And the use of '5 whys' in healthcare RCA is not typically part of a full-scale Lean management approach.

If the use of '5 whys' does not imply the adoption of Lean, and if the evidence to date does not support the effectiveness of Lean in healthcare in any case, there is little reason to be swayed by the pedigree argument.

Using '5 whys' undermines an already weak RCA process

A recent article by Peerally *et al*[33] describes a number of important weaknesses in healthcare RCA practice. Some of these have been explored above (eg, focusing on a single root cause or a small handful of them; and poor quality investigations), but it also raises a number of other issues, such as misuse of the RCA process to pursue (or avoid) other agendas; failure to support feedback loops and double-loop learning; a focus on individual and isolated incidents; a confused approach to blame; the 'problem of many hands' (see also ref. 63) and the shortfalls of a retrospective approach.

The authors also note that RCA often results in poorly designed and/or poorly implemented risk controls.[33] While the goal of learning from incidents is to reduce risk and prevent future harm, the actual tools and techniques of current practice focus exclusively on diagnosing problems; they provide no direct support for prescribing and managing *treatments* for the organisational pathologies they uncover.[64 65]

Some organisations have adopted the PDSA approach to continuous improvement of their risk control action plans.[37 66] But this is a high-level framework, akin to the scientific method. And like the scientific method, it must be implemented through an appropriate set of tools and techniques to produce reliable results.[37] Although a handful of such tools have been introduced in recent years,[35 36 67–72] they have not yet been adopted as the current standard of practice.

Perhaps as a result, there is little evidence to suggest that current practice in RCA improves outcomes.[65]

These challenges do not mean that RCA is never worthwhile; it certainly can be a source of important learning[73] and improvement.[74] But it *does* mean that we cannot afford to compound these problems through the use of an RCA tool that is so deeply and fundamentally flawed. Other more systems-focused techniques, such as fishbone[75] or lovebug diagrams,[72] causal tree diagrams,[21] Causal Analysis based on Systems Theory (CAST)[76] or even prospective risk assessment approaches,[77–81] should be considered instead.

CONCLUSION

When used carefully, '5 whys' may play a powerful role in the classroom. It can illustrate both the need for depth (as a positive example) and the need for breadth (as a negative example) when analysing complex problems.

As a tool for conducting RCAs, however, especially in the area of patient safety, the use of '5 whys' should be abandoned. As the (apocryphal) quote goes: "For every complex problem, there is an answer that is clear, simple and wrong".[82] When it comes to accident investigation, '5 whys' is that answer.

Competing interests None declared.

Provenance and peer review Commissioned; internally peer reviewed.

REFERENCES

1 World Health Organisation. *The WHO patient safety curriculum guide: multi-professional edition*. Geneva, 2011.
2 NHS Improving Quality. Improvement tools. http://www.nhsiq.nhs.uk/improvement-programmes/patient-safety/improvement-resources/improvement-tools.aspx (accessed 28 Jan 2016).
3 Institute for Healthcare Improvement. Ask 'why' five times to get to the root cause. Improv. stories. http://www.ihi.org/resources/Pages/ImprovementStories/AskWhyFiveTimestoGettotheRootCause.aspx (accessed 28 Jan 2016).
4 The Joint Commission. Sentinel event policy and procedures. 2016. http://www.jointcommission.org/sentinel_event_policy_and_procedures/ (accessed 28 Jan 2016).
5 Card AJ, Ward JR, Clarkson PJ. Getting to Zero: evidence-based healthcare risk management is key. *J Healthc Risk Manag* 2012;32:20–7.
6 Shojania KG, Grimshaw JM. Evidence-based quality improvement: the state of the science. *Health Aff* 2005;24:138–50.
7 Marshall M, Pronovost P, Dixon-Woods M. Promotion of improvement as a science. *Lancet* 2013;381:419–21.
8 Shekelle PG, Pronovost PJ, Wachter RM, *et al*. Advancing the science of patient safety. *Ann Intern Med* 2011;154:693–6.
9 Ohno T. *Toyota production system: beyond large-scale production*. Portland, OR: Productivity Press, 1988.
10 Zidel TG. A lean toolbox: using lean principles and techniques in healthcare. *J Heal Qual* 2006;28:7–15.
11 Yuniarto HA. The shortcomings of existing root cause analysis tools. *Lect Notes Eng Comput Sci* 2012;III:4–7.
12 Gross JA. 5 Whys Folklore: The Truth Behind a Monumental Mystery. The KaiZone, 2014. http://www.webcitation.org/6jHDZM1Gc (accessed 25 Jul 2016).
13 Leveson NG. Applying systems thinking to analyze and learn from events. *Saf Sci* 2011;49:55–64.
14 Einhorn Yaffee Prescott Architecture and Engineering PC, Messersmith DH. *Lincoln memorial lighting and midge study (complete analysis of existing system)*. Washington DC, 1993.
15 GSA. Marble: Characteristics, Uses and Problems. Outdoor Sculpt. Man.—Cent. Public Build. 2016. http://www.gsa.gov/portal/content/111858 (accessed 2 Aug 2016).
16 Sheppard N Jr. *Monumental mess*. Chicago Trib. 1995. http://articles.chicagotribune.com/1995-09-01/features/9509010003_1_monuments-park-service-lincoln-memorial
17 Toyota Motor Corporation. *The 'thinking' production system: TPS as a winning strategy for developing people*. 2003.
18 Graban M. *All you have to do is 'ask why five times?' Nope.* Try Eight Steps. 2016. https://www.linkedin.com/pulse/all-you-have-do-ask-why-five-times-nope-mark-graban (accessed 11 Feb 2016).
19 Serrat O. *The five whys technique*. Washington DC, 2010. http://digitalcommons.ilr.cornell.edu/intl/198/
20 Latino RJ, Flood A. Optimizing FMEA and RCA efforts in health care. *J Healthc Risk Manag* 2004;24:21–8.
21 Battles JB, Dixon NM, Borotkanics RJ, *et al*. Sensemaking of patient safety risks and hazards. *Heal Serv Res* 2006;41:1555–75.
22 Laschinger HKS. Impact of workplace mistreatment on patient safety risk and nurse-assessed patient outcomes. *J Nurs Adm* 2014;44:284–90.
23 Lewis MA. Nurse bullying: organizational considerations in the maintenance and perpetration of health care bullying cultures. *J Nurs Manag* 2006;14:52–8.
24 Laschinger HKS, Grau AL, Finegan J, *et al*. New graduate nurses' experiences of bullying and burnout in hospital settings. *J Adv Nurs* 2010;66:2732–42.
25 Rydon-Grange M. 'What's Psychology got to do with it?' Applying psychological theory to understanding failures in modern healthcare settings. *J Med Ethics* 2015;41:880–4.
26 Poon EG, Keohane CA, Yoon CS, *et al*. Effect of bar-code technology on the safety of medication administration. *N Engl J Med* 2010;362:1698–707.
27 McDonald CJ. Computerization can create safety hazards: a bar-coding near miss. *Ann Intern Med* 2006;144:510–16.
28 Koppel R, Wetterneck T, Telles JL, *et al*. Workarounds to barcode medication administration systems: their occurrences, causes, and threats to patient safety. *J Am Med Inform Assoc* 2008;15:408–23.
29 Schnock KO, Dykes PC, Albert J, *et al*. The frequency of intravenous medication administration errors related to smart infusion pumps: a multihospital observational study. *BMJ Qual Saf* 2016. Published Online First: 23 Feb 2016.
30 Dale JC, Renner SW. Wristband errors in small hospitals. A College of American Pathologists' Q-Probes study of quality issues in patient identification. *Lab Med* 1997;28:203–7.
31 Renner S, Howanitz P, Bachner P. Wristband identification error reporting in 712 hospitals. A College of American Pathologists' Q-Probes study of quality issues in transfusion practice. *Arch Pathol Lab Med* 1993;117:573–7.
32 NHS Institute for Innovation and Improvement. Root cause analysis using five whys. 2008. http://www.institute.nhs.uk/quality_and_service_improvement_tools/quality_and_service_improvement_tools/identifying_problems_-_root_cause_analysis_using5_whys.html (accessed 20 Mar 2016).
33 Peerally MF, Carr S, Waring J, *et al*. The problem with root cause analysis. *BMJ Qual Saf* 2016. Published Online First: 23 Jun 2016.
34 Wieman TJ, Wieman EA. A systems approach to error prevention in medicine. *J Surg Oncol* 2004;88:115–21.
35 Card AJ, Ward JR, Clarkson PJ. Rebalancing risk management—Part 1: the Process for Active Risk Control (PARC). *J Healthc Risk Manag* 2014;34:21–30.
36 Card AJ, Ward JR, Clarkson PJ. Rebalancing risk management—Part 2: the Active Risk Control (ARC) Toolkit. *J Healthc Risk Manag* 2015;34:4–17.
37 Reed JE, Card AJ. The problem with Plan-Do-Study-Act cycles. *BMJ Qual Saf* 2015. Published Online First: 23 Dec 2015.
38 Vincent CA. Analysis of clinical incidents: a window on the system not a search for root causes. *Qual Saf Health Care* 2004;13:242–3.
39 Rittel H, Webber M. Dilemmas in a general theory of planning. *Policy Sci* 1973;4:155–69.
40 Cross N. Designerly ways of knowing. *Des Stud* 1982;3:221–7.
41 Glouberman S, Zimmerman B. *Complicated and complex systems: what would successful reform of Medicare Look Like?* Commission on the Future of Health Care in Canada, 2002.
42 Chassin MR, Loeb JM. High-reliability health care: getting there from here. *Milbank Q* 2013;91:459–90.
43 Sutcliffe KM, Paine L, Pronovost PJ. Re-examining high reliability: actively organising for safety. *BMJ Qual Saf* 2016. Published Online First: 21 Mar 2016.
44 Agency for Healthcare Research and Quality. 2014 National Healthcare Quality and Disparities Report. Rockville, MD: 2015. http://www.ahrq.gov/sites/default/files/wysiwyg/research/

findings/nhqrdr/2014chartbooks/hispanichealth/2014nhqdr-hispanichealth.pdf

45 Landrigan CP, Parry GJ, Bones CB, *et al*. Temporal trends in rates of patient harm resulting from medical care. *N Engl J Med* 2010;363:2124–34.

46 HHS OIG. *Adverse events in hospitals: incidence among Medicare beneficiaries*. Washington DC, 2010.

47 Unbeck M, Schildmeijer K, Henriksson P, *et al*. Is detection of adverse events affected by record review methodology? An evaluation of the 'Harvard Medical Practice Study' method and the 'Global Trigger Tool'. *Patient Saf Surg* 2013;7:10.

48 Suarez C, Menendez MD, Alonso J, *et al*. Detection of adverse events in an acute geriatric hospital over a 6-year period using the Global Trigger Tool. *J Am Geriatr Soc* 2014;62:896–900.

49 Classen DC, Resar R, Griffin F, *et al*. 'Global trigger tool' shows that adverse events in hospitals may be ten times greater than previously measured. *Health Aff* 2011;30:581–9.

50 HHS OIG. *Adverse events in rehabilitation hospitals: national incidence among Medicare beneficiaries*. Washington DC, 2016.

51 Pande P, Holpp L. *What is Six Sigma?* McGraw-Hill, 2002.

52 McGlynn EA, Asch SM, Adams J, *et al*. Quality of health care delivered to adults in the United States. *N Engl J Med* 2003;348:2635–45.

53 Snowden DJ, Boone ME. A leader's framework for decision making. *Harv Bus Rev* 2007;85:68–76, 149.

54 Card AJ. Patient safety: this is public health. *J Healthc Risk Manag* 2014;34:6–12.

55 Emanuel L, Berwick D, Conway J, *et al*. What exactly is patient safety. In: Henriksen K, Battles JB, Keyes MA, *et al*., eds. *Advances in patient safety: new directions and alternative approaches*. Vol. 1: Assessment. Rockville, MD: AHRQ, 2008:1–18. http://ahrq.hhs.gov/downloads/pub/advances2/vol1/Advances-Emanuel-Berwick_110.pdf (accessed 28 Feb 2014).

56 Vest JR, Gamm LD. A critical review of the research literature on Six Sigma, Lean and StuderGroup's Hardwiring Excellence in the United States: the need to demonstrate and communicate the effectiveness of transformation strategies in healthcare. *Implement Sci* 2009;4:35.

57 Andersen H, Røvik KA, Ingebrigtsen T. Lean thinking in hospitals: is there a cure for the absence of evidence? A systematic review of reviews. *BMJ Open* 2014;4:e003873.

58 DelliFraine JL, Langabeer JR II, Nembhard IM. Assessing the evidence of Six Sigma and Lean in the health care industry. *Qual Manag Health Care* 2010;19:211–25.

59 Moraros J, Lemstra M, Nwankwo C. Lean interventions in healthcare: do they actually work? A systematic literature review. *Int J Qual Heal Care* 2016;28:150–65.

60 Poksinska BB, Fialkowska-Filipek M, Engström J. Does Lean healthcare improve patient satisfaction? A mixed-method investigation into primary care. *BMJ Qual Saf* 2016. Published Online First: 10 Feb 2016.

61 Radnor ZJ, Holweg M, Waring J. Lean in healthcare: the unfilled promise? *Soc Sci Med* 2012;74:364–71.

62 Toy problem. Dict. Comput. http://www.webster-dictionary.org/definition/toy problem (accessed 2 Aug 2016).

63 Dixon-Woods M, Pronovost PJ. Patient safety and the problem of many hands. *BMJ Qual Saf* 2016; 25:485–8.

64 Card AJ, Ward JR, Clarkson PJ. Trust-level risk evaluation and risk control guidance in the NHS East of England. *Risk Anal* 2014;34:1471–81.

65 Card AJ, Ward J, Clarkson PJ. Successful risk assessment may not always lead to successful risk control: a systematic literature review of risk control after root cause analysis. *J Healthc Risk Manag* 2012;31:6–12.

66 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.

67 Godfrey AB, Clapp TG, Nakajo T, *et al*. Application of healthcare-focused error proofing: principles and solution directions for reducing human errors. In: *ASQ World Conference on Quality & Improvement Proceedings*. Seattle, WA: American Society for Quality, 2005:335–40.

68 Vacher A, d'Hollander A, El Mhamdi S, *et al*. Effectiveness of a tool for structuring action plan after analysis of adverse event. *Proc Hum Factors Ergon Soc Annu Meet* 2011;55:1631–4.

69 Card AJ. The Active Risk Control (ARC) toolkit: a new approach to designing risk control interventions. *J Healthc Risk Manag* 2014;33:5–14.

70 Card AJ, Simsekler MCE, Clark M, *et al*. Use of the Generating Options for Active Risk Control (GO-ARC) Technique can lead to more robust risk control options. *Int J Risk Saf Med* 2014;26:199–211.

71 Card AJ, Ward JR, Clarkson PJ. Generating Options for Active Risk Control (GO-ARC): introducing a novel technique. *J Healthc Qual* 2014;36:32–41.

72 Card AJ. A new tool for hazard analysis and force field analysis: the Lovebug Diagram. *Clin Risk* 2013;19:87–92.

73 Leistikow I, Mulder S, Vesseur J, *et al*. Learning from incidents in healthcare: the journey, not the arrival, matters. *BMJ Qual Saf* 2016. Published Online First: 1 Apr 2016.

74 Chen KH, Chen LR, Su S. Applying root cause analysis to improve patient safety: decreasing falls in postpartum women. *Qual Saf Heal Care* 2010;19:138–43.

75 Taylor-Adams S, Vincent C. Systems analysis of clinical incidents: the London protocol. *Clin Risk* 2004;10:211–20.

76 Leveson N, Samost A, Dekker S, *et al*. A systems approach to analyzing and preventing hospital adverse events. *J Patient Saf* 2016. Published Online First: 11 Jan 2016.

77 Horberry T, Teng YC, Ward J, *et al*. Guidewire retention following central venous catheterisation: A human factors and safe design investigation. *Int J Risk Saf Med* 2014;26:23–37.

78 Card AJ, Harrison H, Ward J, *et al*. Using prospective hazard analysis to assess an active shooter emergency operations plan. *J Healthc Risk Manag* 2012;31:34–40.

79 Card AJ, Ward JR, Clarkson PJ. Beyond FMEA: the structured what-if technique (SWIFT). *J Healthc Risk Manag* 2012;31:23–9.

80 Ward J, Clarkson J, Buckle P, *et al*. Prospective hazard analysis: tailoring prospective methods to a healthcare context. 2010. http://www.webcitation.org/6KZ0Y4R8E

81 Hyman WA, Johnson E. Fault tree analysis of clinical alarms. *J Clin Eng* 2008;33:85–94.

82 Sturmberg J, Topolski S. For every complex problem, there is an answer that is clear, simple and wrong: and other aphorisms about medical statistical fallacies. *J Eval Clin Pract* 2014;20:1017–25.