Failure mode and effects analysis: too little for too much?

Bryony Dean Franklin,[1,2] Nada Atef Shebl,[2] Nick Barber[1,2]

[1]Centre for Medication Safety and Service Quality, UCL School of Pharmacy and Imperial College Healthcare NHS Trust, London, UK
[2]Department of Practice and Policy, UCL School of Pharmacy, London, UK

Correspondence to
Professor Bryony Dean Franklin, Director, Centre for Medication Safety and Service Quality, UCL School of Pharmacy and Imperial College Healthcare NHS Trust, Pharmacy Department, Charing Cross Hospital, London W6 8RF, UK; bryony.deanfranklin@imperial.nhs.uk

ABSTRACT

Failure mode and effects analysis (FMEA) is a structured prospective risk assessment method that is widely used within healthcare. FMEA involves a multidisciplinary team mapping out a high-risk process of care, identifying the failures that can occur, and then characterising each of these in terms of probability of occurrence, severity of effects and detectability, to give a risk priority number used to identify failures most in need of attention. One might assume that such a widely used tool would have an established evidence base. This paper considers whether or not this is the case, examining the evidence for the reliability and validity of its outputs, the mathematical principles behind the calculation of a risk priority number, and variation in how it is used in practice. We also consider the likely advantages of this approach, together with the disadvantages in terms of the healthcare professionals' time involved. We conclude that although FMEA is popular and many published studies have reported its use within healthcare, there is little evidence to support its use for the quantitative prioritisation of process failures. It lacks both reliability and validity, and is very time consuming. We would not recommend its use as a quantitative technique to prioritise, promote or study patient safety interventions. However, the stage of FMEA involving multidisciplinary mapping process seems valuable and work is now needed to identify the best way of converting this into plans for action.

INTRODUCTION

Healthcare has borrowed two forms of risk assessment from other high risk industries. The first, often referred to as root cause analysis, involves a structured retrospective review of a 'critical incident', an event that caused serious patient harm or a near miss that carried substantial potential for harm.[1] However, there also exists a need in high risk industries for prospective risk assessment ie, identifying the likely ways in which a complex process or technology might fail, together with the likely impacts of such failures. Particularly for new processes or technologies (or major modifications to existing processes), it seems appropriate to have a systematic approach to risk mitigation ahead of time rather than waiting for risks to materialise. There are many structured methods of prospective risk assessment, of which failure mode and effects analysis (FMEA) is the most widely used in healthcare.[2]

FMEA involves a multidisciplinary team mapping out a high-risk process of care in order to identify the failures that can occur (see figure 1). Briefly, the team then characterise each 'failure mode' in terms of three characteristics: the probability of occurrence, severity of effects and detectability (the degree to which something can be discovered and rectified before harm results). A single risk priority number (RPN) is then calculated for each failure by multiplying scores for each severity, probability and detectability (usually each using a 10-point scale, accompanied by written descriptions for the numerical scores). Thus, RPN is intended to guide the team's efforts by highlighting the failures with the highest RPNs, and thus most in need of attention. Conducting these steps should also encourage the team to thoroughly understand the processes involved. Finally, the team makes recommendations to prevent or mitigate the failure modes, in which those with the highest RPNs are usually addressed first. To achieve this entire process, the multidisciplinary group have to meet multiple times, typically for an hour or two each time.

Within healthcare, conducting an annual FMEA or similar prospective risk assessment is a requirement for Joint Commission accreditation in the USA, and for Accreditation Canada. FMEA was also used as a tool in the UK's Safer Patients Initiatives.[6] Studies have been published describing the use of FMEA to compare new and old processes of care, to investigate the likely consequences of introducing new technologies or medical equipment, laboratory-related processes and
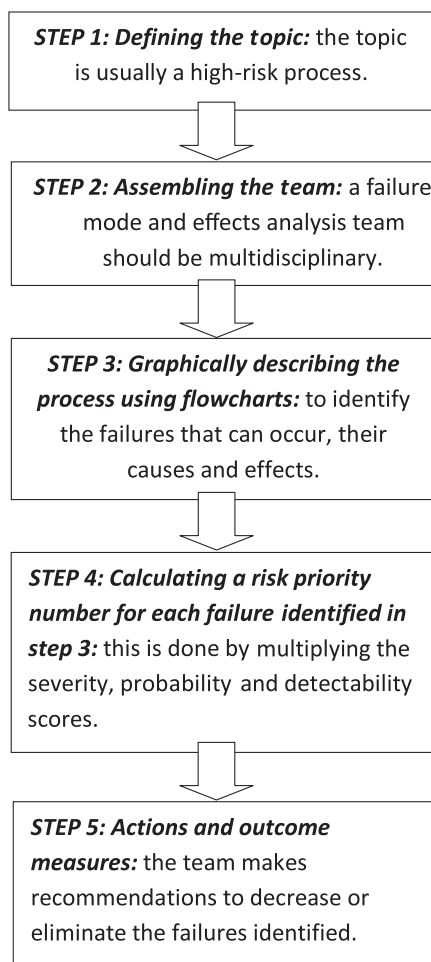
STEP 1: *Defining the topic:* the topic is usually a high-risk process.

STEP 2: *Assembling the team:* a failure mode and effects analysis team should be multidisciplinary.

STEP 3: *Graphically describing the process using flowcharts:* to identify the failures that can occur, their causes and effects.

STEP 4: *Calculating a risk priority number for each failure identified in step 3:* this is done by multiplying the severity, probability and detectability scores.

STEP 5: *Actions and outcome measures:* the team makes recommendations to decrease or eliminate the failures identified.

Figure 1 Failure mode and effects analysis steps.[3–5]

even hospital layout designs.[7] FMEA has been used both as a quality improvement tool and as a research tool to quantitatively compare alternative processes or to identify and prioritise failures.

One might assume that a tool associated with such widespread use would have an established evidence base. Here we consider whether this is the case, and question if, and how, FMEA should be used within healthcare. We recognise that there may be differences in how FMEA is used in routine practice in healthcare, by healthcare professionals with little or no experience or training in its use, and how it might be used by experienced risk engineers. Here, we focus on how it is typically used within healthcare.

RELIABILITY AND VALIDITY

Any tool used to guide decision-making might be expected to have reasonable reliability and validity, particularly if its outputs are used quantitatively. However, little work has been done to examine the reliability and validity of the FMEA process or its outputs. The few studies that have explored these issues in healthcare, disappointingly, do not show favourable results.

A reliable technique would produce the same results regardless of who actually performed the technique. In 2009, we explored the reliability of FMEA by recruiting two similar multidisciplinary groups from within the same organisation, to conduct separate FMEAs in parallel on the same topic. The two groups created similar process maps with similar steps in the process of care, but identified different failures and very different RPNs.[8]

Validity of a technique refers to the extent to which it measures what it is purported to measure. We explored the validity of FMEA's outputs using four different methods and concluded that there are significant methodological challenges in validating FMEA.[7] Face validity was found to be positive as the FMEA participants documented the same processes of care as mapped by the researcher, following detailed observation of the process. However, both FMEA teams missed potential failures identified by other healthcare professionals, including the category of failure that was most commonly spontaneously reported within the study organisation.

In addition to researchers who have shed doubts on FMEA's reliability and validity, users have also questioned these characteristics; many of the Safer Patients Initiative participants who conducted FMEAs in UK hospitals had concerns about its reliability and/or validity.[9]

PROBLEMATIC ASPECTS OF THE RPN

Capturing the risk of failure in terms of probability, severity and detectability makes sense conceptually. However, calculating the RPN by multiplying severity, probability and detectability scores gives rise to several mathematical problems. First, such calculation of the RPN breaches the mathematical properties of the ordinal scales used, as ordinal numbers cannot meaningfully be multiplied or divided. Second, Bowles[10] highlights further mathematical limitations of the RPN as used in FMEA. Assuming 10-point scales are used, he showed that 1000 is the largest number, 900 the second largest, followed by 810, 800, 729 and 720. The differences between consecutive possible numbers cover a very variable range. Is the difference between 720 ($8\times9\times10$) and 729 ($9\times9\times9$) the same or less than the difference between 900 ($9\times10\times10$) and 1000 ($10\times10\times10$)? Third, the majority of RPN values can also be formed in several ways. For example, RPN values of 60, 72 and 120 can each be formed from 24 different combinations of severity, probability and detectability, so although the RPN values may be identical, their risk implications may be different. Finally, small variations in one of the three parameters can lead to very different effects on the RPN, depending on the values of other parameters (table 1).

Table 1 Example of the RPN and its sensitivity to small changes[10]

Scenario	Severity	Probability	Detectability	RPN
A	3	8	8	192
	4	8	8	256
B	8	3	8	192
	9	3	8	216

Small variations in one parameter can lead to very different effects on the RPN, depending on the values of the other factors. A 1-point change in the severity in scenario A causes a 64-point change in the RPN, whereas for scenario B, a 1-point change in severity causes only a 24-point change. RPN, risk priority number.

FMEA is increasingly used as a research tool or as part of a quality improvement cycle, in which reduced RPNs are seen as proof of an improved process. The above mathematical challenges to the interpretation of the RPN call into question the value of these activities.

FURTHER VARIATION IN PRACTICE

In addition to the lack of evidence for its underlying validity and reliability, there is also substantial variation in how FMEA is used in practice. Although FMEA theoretically involves following a set of standard steps, there is considerable variation in how these steps are applied within healthcare. Such variation includes team composition, frequency of meetings, duration of the FMEA, the FMEA steps followed, how the severity, probability and detectability (if included) are quantified, and how failures are prioritised for action.[7] Complicating things further is the development of other derivatives of FMEA such as healthcare FMEA (HFMEA), which was developed and introduced by the U.S. Veteran's Administration National Centre for Patient Safety in 2001.[5] Although both FMEA and HFMEA have similarities at their core, HFMEA uses 4-point instead of 10-point scales and HFMEA detectability scores are only determined if the failure identified warrants further action, as determined by a decision tree.

These variations may be important. For example, Ashley and Armitage[11] compared two different scoring procedures for the same failures by the same FMEA team. First, a mathematical procedure in which scores are assigned independently by each team member and averaged. Second, a consensus procedure, in which the scores are agreed via discussion. The two scoring procedures yielded notably different scores which in turn resulted in a clear difference in the failures' prioritisation.

IS IT WORTH IT?

Potential advantages of FMEA, as suggested in the literature, are that it is a useful tool to aid multidisciplinary

groups in mapping, understanding and prioritising improvements to a process of care, it allows teams to consider vulnerabilities within a process of care before they actually occur, and that it can be used as an educational tool.[12–17]

However, a problem in conducting FMEA is that it is very time-consuming. We identified 10 published studies of FMEA in healthcare which stated the number of meetings that had been required. There was an average of eight meetings (range 2–19), which had a mean duration of 1.5 h each. Of 26 studies which cited the number of participants, the average was eight (range 2–22).[7] This corresponds to 96 h of healthcare professionals' time per FMEA. This number and length of meetings may result in inconsistent attendance due to work schedules and time commitments, resulting in loss of expertise and continuity.

SHOULD HEALTHCARE CONTINUE TO USE FMEA?

In short, FMEA in healthcare is associated with a lack of standardisation in how the scoring scales are used and how failures are prioritised. Different team members and different scoring methods yield dissimilar results, and the concept of multiplying ordinal scales to prioritise failures is mathematically flawed. The FMEA process is subjective, but the use of numerical scores gives an unwarranted impression of objectivity and precision. FMEA is therefore a tool for which there is a lack of evidence. It is surprising that such a commonly used and widely promoted technique within healthcare appears to have no evidence that its outcomes are valid and reliable; particularly as it is used to prioritise patient safety practices and requires so much staff time. Similar concerns have also been raised about root cause analysis, which has not been evaluated for effectiveness.[18]

Many of the problems with FMEA relate to the number of meetings required and to the weaknesses in calculating the RPN—interestingly, this is the most time consuming element of FMEA. We suggest that the most effective attribute of FMEA is that it involves gathering a multidisciplinary team to map out a process of care and identify the failures that may occur. This allows participants to gain an insight into their colleagues' daily practice and challenges faced, especially since most healthcare processes require teamwork rather than an individual approach, and also serves an educational purpose. Our work suggests that the process mapping stage is the part of FMEA that is most valid and reliable.[7] [8] Since it can be used prospectively, there may well be situations where mapping out the process and anticipating the likely failures may be useful. Its use may bring qualitative benefits in terms of mapping and sharing understanding. However, we argue that FMEA's

quantitative outputs lack sufficient validity and reliability to be used as a sole method of prioritising patient safety interventions, or as a quality improvement or research tool, while requiring a substantial time commitment.

The benefits of gathering a multidisciplinary team to discuss a process of care are clear; however, the RPN scores have, inappropriately in our view, become the focus, and the aim of FMEA then becomes the reduction of the RPN values rather than finding and evaluating solutions to avoid harming our patients.

RECOMMENDATIONS

We would therefore suggest that the focus should be on the qualitative part of FMEA: mapping and understanding the process in a structured way, rather than focusing on the RPN.

If calculation of RPNs is not useful for prioritising failures, other approaches are needed. Developments in psychology in the last few years have given many illustrations in which the use of intuition or simple rules of thumb have been equal to or superior to sophisticated analytical methods in decision making. This literature is well described in Gerd Gigerenzer's excellent book 'Gut Feelings'.[19] There may also be other general principles of system improvement that are helpful in identifying and prioritising recommendations, such as getting things right first time rather than focusing on later corrective feedback loops, starting where the patient will experience most difference, and starting with the most common failures (if good quality data are available about their frequency). We would recommend that existing data on different types of failure, where available, be used as part of this process. These suggestions could be brought together into an integrated approach which recognises the necessary subjectivity of the process, allows a consensus on actions and decisions, and is far less resource intensive than FMEA. This might be termed 'failure mapping and corrective action (FMCA)'; work is now needed to explore these options in more detail.

Finally, we recommend that some of the many other methods of human reliability analysis are explored for use in healthcare. A possible reason for the limitations of FMEA may be that it was originally developed for use in engineering, where systems are largely deterministic and failure rates more easily quantifiable. However, in healthcare, human-based systems introduce variation, which is much harder to quantify. It may therefore be that other methods are more appropriate. Lyons et al[20] identified popular human reliability analysis techniques used in other industries and considered their feasibility for use in healthcare, concluding that there is considerable scope to use other techniques. Ward et al[2] also point out that prospective risk assessment is not a single method but an approach, with a whole range of tools. They considered a wide range of methods and produced a toolkit to support their use within a healthcare context. Further work should therefore explore these techniques in terms of their practicality and value within healthcare, and their relative advantages and disadvantages, and reliability and validity.

CONCLUSION

Although FMEA is popular and many published studies have reported its use within healthcare, there is little evidence to support its use for the quantitative prioritisation of process failures. It lacks both reliability and validity, and is very time consuming. We would not recommend its use as a quantitative technique to prioritise, promote or study patient safety interventions. However, the initial multidisciplinary mapping process seems valuable and work is now needed to identify the best way of converting this into plans for action.

REFERENCES

1. Bagian JP, Gosbee J, Lee CZ, et al. The Veterans Affairs root cause analysis system in action. *Jt Comm J Qual Improv* 2002;28:531–45.
2. Ward J, Clarkson J, Buckle P, et al. *Prospective Hazard Analysis: Tailoring Prospective Methods To A Healthcare Context. Research Project PS/035 Report to the Patient Safety Research Programme of the Department of Health.* http://www.haps2.bham.ac.uk/publichealth/psrp/documents/PS035_-_Revised_PHA_Final_Report_v1.1_with_Toolkit__July_2010_.pdf (accessed 25 Jan 2012).
3. NPSA-National Patient Safety Agency NHS. *Seven Steps to Patient Safety.* 2004. http://www.npsa.nhs.uk/site/media/documents/499_sevensteps_overview(2).pdf (accessed 19 Jun 2006).
4. Wetterneck TB, Skibinski KA, Roberts TL, et al. Using failure mode and effects analysis to plan implementation of smart i.v. pump technology. *Am J Health Syst Pharm* 2006;63:1528–38.
5. VA NCPS- Veteran's Administration National Center for Patient Safety. *The Basics of Healthcare Failure Mode and Effect Analysis.* 2005. http://www.patientsafety.gov/SafetyTopics/HFMEA/HFMEAIntro.doc (accessed 2 Mar 2012).
6. Safer Patients Initiative Program. http://www.ihi.org/IHI/Programs/StrategicInitiatives/SaferPatientsInitiative.htm (accessed 10 Nov 2011).
7. Shebl NA. *Promoting Patient Safety Using Failure Mode and Effect Analysis. PhD Thesis, the School of Pharmacy.* London: University of London, 2010.
8. Shebl NA, Franklin BD, Barber N. Is failure mode and effect analysis (FMEA) reliable? *J Patient Saf* 2009;5:86–94.
9. Shebl N, Franklin B, Barber N, et al. Failure Mode and Effect Analysis (FMEA): the views of hospital staff in the UK. *J Health Serv Res Policy* 2012;17:37–43.

10. Bowles JB. An assessment of RPN prioritization in a failure modes effects and criticality analysis. *Journal of the Institute of Environmental Sciences and Technology* 2003;47:51–5.
11. Ashley L, Armitage G. Failure mode and effects analysis: an empirical comparison of failure mode scoring procedures. *J Patient Saf* 2010;6:210–15.
12. Redfern E, Brown R, Vincent CA. Improving communication in the emergency department. *Emerg Med J* 2009;26:658–61.
13. Ford EC, Gaudette R, Myers L, *et al*. Evaluation of safety in a radiation oncology setting using failure mode and effects analysis. *Int J Radiat Oncol Biol Phys* 2009;74:852–8.
14. Riehle MA, Bergeron D, Hyrkäs K. Improving process while changing practice: FMEA and medication administration. *Nurs Manage* 2008;39:28–33; quiz 34.
15. Nickerson T, Jenkins M, Greenall J. Using ISMP Canada's framework for failure mode and effects analysis: a tale of two FMEAs. *Healthc Q* 2008;11:40–6.
16. Cheung DS, Maygers J, Khouri-Stevens Z, *et al*. Failure modes and effects analysis: minimizing harm to our bariatric patients. *Bariatric Nursing and Surgical Patient Care* 2006;1:107–14.
17. Wetterneck TB, Hundt AS, Carayon P. FMEA team performance in health care: a qualitative analysis of team member perceptions. *J Patient Saf* 2009;5:102–8.
18. Wu AW, Lipshutz AK, Pronovost PJ. Effectiveness and efficiency of root cause analysis in medicine. *JAMA* 2008;299:685–7.
19. Gigerenzer G. *Gut Feelings*. London: Penguin Books, 2007.
20. Lyons M, Adams S, Woloshynowych M, *et al*. Human reliability analysis in healthcare: a review of techniques. *Int J Risk Saf Med* 2004;16:223–37.
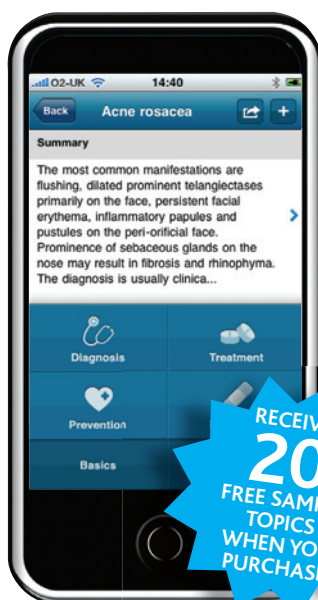
BMJ Quality & Safety

Failure mode and effects analysis: too little for too much?

Bryony Dean Franklin, Nada Atef Shebl and Nick Barber

Updated information and services can be found at:
http://qualitysafety.bmj.com/content/21/7/607

These include:

References	This article cites 14 articles, 2 of which you can access for free at: **http://qualitysafety.bmj.com/content/21/7/607**#BIBL
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