International Journal of Quality & Reliability Management

Failure modes effects analysis (FMEA) for review of a diagnostic genetic laboratory process

Karen Claxton, Nicola Marie Campbell-Allen,

Failure modes effects analysis (FMEA) for review of a diagnostic genetic laboratory process

Karen Claxton and Nicola Marie Campbell-Allen
School of Engineering and Advanced Technology, Massey University, Palmerston North, New Zealand

Received 6 May 2015
Revised 22 October 2015
Accepted 16 November 2015

Abstract

Purpose – For any improvement tool to be successfully integrated into an organizations' quality improvement or risk management programme, it needs to be relatively easy-to-use and proven to provide benefits to the customer and organization. Many healthcare organizations are facing fiscal constraints and increasing complexity of tests, putting strains on resources, particularly for those on "the shop floor" who are "hands on" in the design, delivery and improvement of products or services. Within a laboratory setting, there is often limited time for formal extensive process reviews; with the pressure to meet "turn-around times" for often "clinically urgent" results. Preventative and corrective actions are often identified through audits or root-cause analysis in some cases after an event has occurred. The paper aims to discuss these issues.

Design/methodology/approach – Failure modes effect analysis (FMEA) is a risk management tool, used to identify prospective failures within processes or products, before they occur. Within laboratory healthcare, risk management for prevention of failure (particularly an inaccurate result) is imperative, and underpins the design of all steps of sample handling. FMEA was used to review a laboratory process for a "gene mutation test" initially considered to have few opportunities for improvement. Despite this perception, a previous review of the process, and the time restrictions for review, new improvements were identified with implications to patient management.

Findings – This study shows that FMEA can yield benefits, for prospective risk management and general process improvement, within a laboratory setting where time and team input is restricted, and within a process that was considered to have few "problems".

Originality/value – The study was undertaken in a large metropolitan public health system laboratory – one of the largest in the country. This laboratory is a significant contributor to the health outcomes of patients in the local region, and through its contribution to national laboratory testing and reporting. This was the first use of FMEA in this laboratory setting.

Keywords Process improvement, Laboratory, Risk management, Failure modes effects analysis (FMEA), Quality management

Paper type Case study

1. Introduction

Within healthcare settings accuracy and speed of reporting of laboratory results (turn-around time (TAT) – from receipt to reporting), is vital to clinical management. Delays or mistakes in reporting can negatively influence treatment initiation and procedure choice. These treatments and choices may affect "critical to health" outcomes. Laboratory tests are evolving as technological and treatment advances are made. Continuous reviews, through quality improvement initiatives are important for meeting and adjusting to these changing needs. Within many laboratory settings however, there is often limited time for more proactive formal, team-based and extensive review processes of new and existing tests, with increasing pressure to meet TAT for often clinically urgent results, increasing referral numbers, cuts in staffing and the increasing complexity of tests. Preventative and corrective actions may only occur following audits or root-cause analysis after an event has occurred, or a customer complaint has been received.

This paper discusses the results of a year-long project taken as part of studies towards a Post Graduate Diploma in Quality Systems at the Massey University, New Zealand. The study was undertaken in a large metropolitan public health system laboratory – one of

International Journal of Quality & Reliability Management
Vol. 34 No. 2, 2017
pp. 265-277

0265-671X
DOI 10.1108/IJQRM-05-2015-0073

the largest in the country. This laboratory is a significant contributor to the health outcomes of patients in the local region, and through its contribution to national laboratory testing and reporting.

Failure modes effect analysis (FMEA) is a risk management tool, used to identify prospective failures within a process, product or service, before they occur, so that proactive steps can be taken to design and implement robust processes. FMEA has been implemented in a number of varying scenarios and its name can reflect this usage – for example it can be implemented:

- in the early design stages of a new product or service – and is referred in that instance as design FMEA;
- prior to implementation of manufacturing and/or service delivery processes in production or service settings – sometimes referred to as potential FMEA, FMEA and criticality analysis (FMECA), process FMEA; and
- as part of continuous improvement strategies, and pre- and post-process alterations.

Within laboratory healthcare, risk management for prevention of failure (particularly an inaccurate result) is imperative, and underpins the design of all steps of sample handling from receipt to reporting. FMEA in this setting can be used to examine and document in detail, all steps of a process for failures (actual and/or potential), significance (or criticality) to the customer, and possible causes and existing controls to detect, prevent or reduce occurrence.

FMEA is used as an accreditation requirement for proactive risk management within healthcare organizations overseas (Shebl *et al.*, 2012; Van Tilberg *et al.*, 2006). Within New Zealand, ISO15189, 4.11 (NZS/ISO), stipulates that laboratories shall have documented procedures for determining where potential non-conformities exist, through review of data and information – but it does not specify a method. Audits have the capacity for preventative action limited to the scope of each review. FMEA is a proactive "self-auditing" improvement tool that provides for a detailed review and analysis of processes, and a structured systematic format for proactive identification of corrective and preventative actions to reduce risks to the customer.

FMEA can be viewed as an extension of Shewhart's PDCA cycle (Deming, 1986), examining not only the current, but the potential "situation" for each process step, "planning and doing" through recommended actions, "acting and checking" from following through and re-evaluation for continuous improvement (McCain, 2006). FMEA can be used at the development stage, or for review of an existing process or product (Teng *et al.*, 2004).

Implementation of FMEA involves the creation of a risk priority number (RPN) which results by rating the severity of each potential failure to the customer (S), the likelihood of occurrence of the failure (O) and the likelihood of detection (D) before the effect of the failure reaches the customer (McCain, 2006). The rating numbers are multiplied to determine a RPN, which is then used for prioritizing action (higher numbers generally taking priority), and measuring the effect of change after improvements are implemented (Tague, 2004; Rodriguez-Perez and Pena-Rodriguez, 2012). FMEA (and its variants) have been seen as a time-consuming tool and has not been used in this laboratory setting previously.

FMEA was used to review the "analytical" and "post analytical" phases of a new laboratory process – a gene mutation test (GMT), (refer Table I), one year after implementation, and nine months after a process review (using brainstorming), resulting in improved workflow. The aims of the FMEA were to use the same team to determine if further improvements could be identified, and to examine the utility of the FMEA's structured analysis and ranking system in prioritizing and minimizing risk in this setting.

The GMT uses a "black box" technology for the analytical phase, reliant on steps set by the manufacturer, with restricted variation and limited access to process details.

KEY; PID = patient identification

Check = steps are where a 2nd scientist checks patient identification (PID) & watches each step

Pre-analytical: (Specimen receipt & registration processes are not included in the scope)

A) *Organize a 'run'* (batch of patients to test- check TAT due dates).

Analytical:

B) *Retrieve kits and reagents*
C) *Tumour removal* off slides (match target to be removed with a stained slide) - (Check 1).
D) *DNA extraction* (Check 2) & *quality check*.
E) *Dilute DNA* (Check 3), *'plating'* (transfer each patient's DNA & controls to 3 wells each, on a shared plate - Check 4). '*Analysis'* (enter each patient identification (PID) into analyzer in matching order & start programme).

Post analytical:

F) *Reporting* (transcribe results into hospital report system - Check 5). Organize follow up tests.

Constraints:

Patients must be batched (4-10 /run).
The shared plate (E) cannot be labelled with the PID.
DNA quality (D) can only have one form of PID.
If a result fails, the test can be repeated one more time (found from the earlier review).

Table I.
GMT FMEA scope: process map framework of the six main functions

In spite of this, the previous review, and time restrictions on team member participation, new improvements were identified (including the highest RPN), with potential for clinical implications to patient management.

2. Method

The team consisted of three scientists, each involved in the weekly handling and analysis of specimens for the GMT – their experience ranged from 4 to 12 months. It was not possible in this study to obtain input from the clinical (doctor's) perspective. The project owner (PO), responsible for FMEA development and documentation, was the section leader of the team, who was the most familiar with FMEA, and had overseen the previous initial review of the GMT protocol. A brief training session on the use of FMEA including the aims, methodology for each assessment and use of RPNs for prioritizing action was provided to the team by the PO.

As part of the project, a process map from the protocol (developed by the team in the previous review) was developed by the PO as an alternative reference to define and summarize the scope – this was reviewed by the team. A process map framework was also developed to identify the key steps involved in the process – the total process contained 59 process steps and this methodology identified more manageable sized steps to study. A summary of the scope (Steps A-F), and process constraints for the GMT project, are shown in Table I.

The methodology (approach and scaling system) applied was tailored from a process tool template from the American Society for Quality (Tague, 2004) which recommends FMEA users apply a rating scale based on 1-10.

The FMEA ratings and scales used by the team are provided in Table II. A criticality (C) score was also assigned to each function. Although intermediary numbers were not defined for S, O and D, a flaw identified in the process during analysis, the team discussed the possible scenarios in between 1 and 10 for each. The scenarios for S, perceived to be the most important due to potential clinical consequences, were defined and used as a guide for rating each failure mode (Table III). S was defined in this manner as potential health impacts

and consequences to the customer are foremost in the scientist's mind throughout all steps of a process.

As there was no direct clinical (doctor's) input, the rating scale for severity was based on the laboratories' perspective of customer requirements (accuracy of result, within required TAT), and the perceived clinical impact of not meeting these requirements. GMT is considered a priority test – a positive or negative result determines the course of targeted therapy. Thus a (clinical) severity score of 10 (inaccurate result) would be the worst-case scenario and may seriously impact the course of therapy given by the clinician. A severity score of 1 (accurate result, within TAT) is the best scenario, and considered the target for all GMT tests. A score between 1 and 10 would indicate a "delayed result" to "no result" that might or might not, impact on clinical management (the scale was left open – a fault in this FMEA). The likelihood ratings were based on the experience of the scientists, examination of potential process flaws and a "walk through" analysis.

The team chose to go to the *Gemba* to understand the GMT process, and undertook a one-and-half-hour walkthrough "simulation". *Gemba* (or *Genba*) is a Japanese term referring to "the real place", "crime scene" or "place where value is created". In manufacturing the *Genba* is the factory floor. It can be any "site" such as a construction site, sales floor or where the service provider interacts directly with the customer (Imai, 1997).

Each process step (function) was examined with the process map and documented protocol at hand – potential failures, causes, effects and current controls were identified. Wherever possible, data from evidence of existing issues (collated by the PO) were discussed and incorporated. Recommendations were identified. After tabulating the findings, the PO reviewed the findings with the most experienced team member before distribution to the team. Due to time restrictions, this second draft was then reviewed separately by each team member.

Where failures had more than one effect, the rating applicable to the worst-case scenario (most severe to the customer) was used (Tague, 2004). In this study, S, O and D scores were averaged this had the potential for skewing data – outliers were therefore examined first (this point is discussed later in this paper). Where there was a duplication of process steps,

Table II.
GMT FMEA rating and scale

Rating	Scale
Severity to the customer if failure occurred (Severity – S)	1 = Least severe 10 = Most severe
Likelihood of the failure occurring (Occurrence – O)	1 = Least likely 10 = Most likely
Likelihood of being able to detect the failure (Detection – D)	1 = Most likely 10 = Least likely
The Criticality of the step to the process (Blank, or C, S or I)	C = Critical to the process S = Significant to the process I = Important to the process

Table III.
GMT FMEA "severity" rating and scale

Severity rating possible error scenario (decreasing severity)	Scale
Accurate result, no error – for appropriate therapy. Within TAT – for clinical management	1
Accurate but delayed result (repeat test) – does not delay therapy	Not quantified
Accurate but delayed result (repeat test) – delays therapy	Not quantified
No result – cannot repeat	Not quantified
Inaccurate result – undetected	10

data were compared and combined, using the highest score (only if the step and the effect was the same). The S, O and D scores were multiplied to produce a RPN.

Analysis of the final FMEA results took two forms. First, errors were separated into the six main process functions (A-F, Table I), which linked issues clearly to their location within the process flow. Second, the RPN, S, O and D scores were compared, in descending order of importance, for each process function. Traditionally process functions with high RPN scores – (those associated with the highest risk), are actioned first to minimize risk – S × O can also be used (Tague, 2004). Individually high S, O and D scores may also indicate the need for prioritizing action – S is the highest concern. However, improving controls may improve D and reduce O.

The data therefore were split into two groups, Group A containing the highest 23 RPN scores and all of the S and D scores of 10 (there were no O scores of 10), and Group B was constructed from the remaining 65 RPN scores of less than 10. The ranking system was then examined qualitatively to evaluate the scores for consistency with prioritizing action aligned with meeting the needs of the customer. Key effects of "delayed run, delayed result, failed result, invalid result/run and inaccurate result" (increasing in severity to the customer – refining the open scale from Table III) were used to search the effect column for consistency with increasing RPN and S scores (Group A results are tabulated within Tables IV and V).

Analysis of the FMEA results enabled priority "improvement" actions to be determined. Changes deemed "easy" or "urgent and critical to the process" were made immediately, or were allocated to a team member with a defined target and review date.

3. Results

In total, 88 potential failure modes were identified within the 59 GMT process steps. The FMEA resulted in eight improvements (within the highest 23 RPNs), seven of which had not been identified by the previous review (Table V). The highest RPN (200) resulted in a preventative action, critical to the integrity of test results – a second scientist rechecking the order of patient's samples into the analyser. The development of the FMEA was worthwhile, for improving controls of this one risk alone. Evaluation of RPNs and the S, O and D scores, although undertaken, was not required for prioritization, as all actions were clear, uncomplicated "quick fixes" (Table V).

3.1 Score comparisons – RPN, S, O and D

The majority of S, O and D scores had "low" values – a range of 1-4.5 (Table VI). The highest O score was 5.7, in an area that did not require action. However there were 15 S scores and three D scores of 10. The S and D scores of 10 were used to form Group A, containing the highest 23 RPNs (range 10-200), with 26 per cent of 88 identified possible failures. Group B had lower S and D scores, and contained the lower 65 RPNs (range 1-9) – 74 per cent of all identified possible failures. Occurrence was roughly the same for both groups (69.6 and 66.1 per cent $\leqslant 2$) – see Groups A and B data in Table VI.

3.2 Consistency of ratings with "prioritizing" for customer needs

High O and D scores can produce high RPN ratings, however in this study the majority of rating ranges were low (Table VI). The high RPNs appear to be consistent with increasing severity of effects, high S scores, and the need for action. Comparison of Group A data show 10/23 risks were associated with possible "inaccurate result" (most severe failure to the customer) with $S = 10$ (shaded, Table IV). For 5/23 (the remaining five with $S = 10$, Table IV, not shaded), "inaccurate result" was not possible due to technical/internal kit controls. However these steps were classified as critical to the process, and were critical to result integrity (involving tumour removal C, and plating; transfer of all patient's and control DNA to a shared plate, E,) and were therefore rated as $S = 10$.

Table IV.
Group a FMEA results (top 23 RPNs)

	Type of failure (summary) Listed in order of GROUP A RPNs (highest to lowest)	Location (function)	Delayed run	Delayed result	Failed result	Invalid result/ run	Inaccurate result	Severity	Occurrence	Detection	Risk priority number	Criticality	Action $n = 8$ (see Table V)
1	Patients entered into analyser in incorrect order	E					Y	10	2	10	200	C	1
2	Incorrect orientation of tumour targets	C			Y		Y	10	3.7	3	111	C	2
3	Removal of tumour target (no target tissue removed)	C			Y		Y	10	2.7	4	108	C	2
4	Removal of tumour target (some removed)	C			Y		Y	10	2.7	4	108	C	2
5	Reagent (not mixing)	B			Y			2	3.5	10	70	C	none* Table V
6	Removal of tumour target (no tissue at all removed)	C	Y	Y	Y	na	na	10	2	2	40	C	2
7	Plating (patient DNA transfer to another patient well)	E			Y		Y	10	1.7	2	34	C	3
8	Plating (patient DNA transfer to empty well)	E			Y		na	10	1.7	2	34	C	3
9	Plating (patient DNA transfer to control well)	E			Y	Y	na	10	1.7	2	34	C	3
10	Reagent wrong concentration	B	(Y)					1.5	2	10	30	S	8
11	Plating (not changing tips)	E					Y	10	1.5	2	30	C	none*
12	Plating (error with seal)	E				Y	Y	10	1.5	2	30	C	none*
13	Plating (reagent to wrong well)	E				Y	na	10	1.7	1.5	25.5	C	3
14	Plating (positive control to wrong well)	E				Y	na	10	1.7	1.5	25.5	C	3 and 6
15	Result entry (incorrect reporting)	F					Y	10	1.7	1.5	22.5	C	1
16	Plating (no seal)	E			Y			2.5	1.7	1.5	19	C	1
17	Check (incorrect report recipients)	F		(Y)				1.5	5	2	15	I	none*
18	Miscalculation of TAT	A		(Y)				1	4.7	3	14		4

(*continued*)

Table IV.

	Type of failure (summary) Listed in order of GROUP A RPNs (highest to lowest)	Location (function)	Delayed run	Delayed result	Failed result	Invalid result/ run	Inaccurate result	Severity	Occurrence	Detection	Risk priority number	Criticality	Action $n = 8$ (see Table V)
19	Result entry (DNA quality into wrong patient –for DNA dilution)	D			Y			2	2	3	12	S	4
20	Result entry – Failed result	F			Yes!			4	3	1	12	I	6
21	Inadequate slide drying	C			Y			1.5	2	3.5	10.5	C	none*
22	Removal of tumour target (into wrong patient tube)	C					Y	10	1	1	10	C	5
23	Checker does not pick up mistake	D					Y	10	1	1	10	C	7
	Total ($n = 23$)		2	3	12	4	10					18 C	13

Notes: *None (no action required – (reminders to follow standard operating procedure). Modification of the FMEA table showing a summary of the top 23 RPNs, associated failures, location in the process and the effect to the customer (increasing severity; delayed run to inaccurate result); Shaded = most critical to integrity of results (could result in inaccurate result). 2, 3 and 4 could result in either a false negative result (false normal) or an accurate result (if sufficient tumour is outside the target)

Recommendations within Group A	No. of potential failures	Location	RPN	Identified at previous review?	Immediate quick fix?
1. New check step – 2nd scientist checks order of patients into analyser	3	E	200, 22.5, 19	N	Y (highly critical)
2. Check slides for tumour target sooner (at receipt – Lean troubleshooting)	4	C	111, 108 (x 2), 40	Y	Y
3. Include "Well" number on worksheet (Lean)	4	E	22.5, 34(x 3)	N	Y
4. Electronic data transfer (Lean)	2	A, D	14, 12	N	Y
5. All slides in one tray, numbered in order for analyser (Lean)	1	C	10	N	Y
6. Correction/addition to SOP (standard operating procedure)	2	E, F	25.5, 12	N	Y
7. Move check to earlier Step C (Lean)	1	D	10	N	Y
8. Second scientist performs repeat test (failure of run /result due to reagent mistake) – under consideration	1	B	30	N	(case by case basis)
	5	B, C, E, F	70,10.5,30 (x2),15	N	n/a

Table V.
Group a recommendations (23 highest RPNs)

Notes: Four improve controls associated with integrity of results (1, 2, 3 and 5 – shaded). Five processes to be made "more Lean" (2, 3, 4, 5, 7), includes the one RPN located in pre-processing (A)

Range of ratings	RPN	S	O	D
All data ($n = 88$)	1-200	1-4.5 (83%) 10 (17%)	1-4.5 (94.3%) 4.6-5.7 (5.7%)	1-4.5 (96.6%) 10 (3.4%)
Group A highest 23 RPNs, and S and D = 10	10-200 (26%)	1-4 (34.8%) 10 (65.2%)	1-3.7 (91.3%) 4.7-5 (8.7%)	1-4.5 (96.6%) 10 (3.4%)
Group B lowest 65 RPNs and S, D < 10	1-9 (74%)	1-4.5 ($\leqslant$2, 93.8%)	1-5.7 ($\leqslant$2, 66.1%)	1-3 ($\leqslant$2, 98.5%)

Table VI.
Ranges and per cent of ratings for Groups A and B RPNs

This shows that the rating system is not discrete. An interplay of factors influences the true meaning of S, and the evaluation of prioritizing actions. Of the ten risks associated with "inaccurate result", two had no action except reminders to "take care" in carrying out steps, and eight had action to "improve controls" associated with maintaining result integrity through 4 recommendations (shaded in Table V). The other four recommendations were minor.

Although eight recommendations stemmed from 18/23 possible failures within the high RPN category (Group A, Tables IV and V), examination of a failure mode with a low RPN may prompt identification of a related failure mode of higher severity. Two of the eight recommendations within Group A (1 and 3, Table V) were prompted by interrogation of failure modes within Group B (data not shown). For example, "incorrect ID into analyser" (RPN 6.6) prompted possible "incorrect order" of patients, resulting in identification of the highest RPN (200 – Table IV) with higher severity to the customer, inaccurate result (controls for RPN 6.6 were sufficient, resulting in rechecking patient identification and retest if required).

The remaining Group B failure modes resulted in seven other minor recommendations involving 9/65 RPNs related to improving workflow and one was a correction to the standard operating procedure (SOP). None of these were urgent, and this was reflected in the low RPN and S scores (Table VI).

Two of the $D = 10$ scores in Group A were related to potential errors in reagent making. Although critical and significant to the process, these were generic and less severe to the customer. Controls were in place but limited to checks of the small volumes handled. Recommendation 8 (Table V) has not been implemented.

3.3 Utility of FMEA in identifying improvements

One aim of FMEA is to eliminate or minimize risk by improving controls. The main benefit of the analysis was the preventative action from the highest RPN (200, $S = 10$) "patients entered into analyser in incorrect order" (at E). The FMEA highlighted this as a high risk not addressed by the team within the previous review. The FMEA brought to light that controls were insufficient, with the potential for "patient mix up" and the most severe effect for the customer "inaccurate result" (Table IV), with possible clinical consequences such as not being offered targeted therapy or being given inappropriate targeted therapy. A new check step using a second scientist to recheck patient order with paperwork was introduced. For this one improvement, the FMEA was worthwhile. This control was also introduced for "result entry at reporting" at F (15th highest RPN 22.5, S = 10). FMEA analysis shows its value here – with a systematic detailed review of process steps. The previous review however, using *Gemba* and brainstorming, may not have been as thorough, concentrating on pre-processing (A – which required addition to the SOP, whereas functions B-F had concentrated on review of the manufacturer's SOP (with no prescriptive check steps throughout, and no *Gemba*).

Table IV shows that all of the other top 23 RPNs were located within Steps B-F. Two areas known to have high risk were confirmed, and the check steps (controls) improved; recommendation 2, Table V (for "tumour removal issues" at C, second, third, fourth and sixth highest RPNs, Table IV), and recommendation 3 (for five possible "plating" errors, at E, seventh-ninth, 13th and 14th highest RPNs). For the five areas not requiring action, controls were considered sufficient.

3.4 Revealing waste

FMEA revealed waste within the GMT process, resulting in five "Lean" improvements (Table V). Unnecessary repetition in three separate tube labelling and checking steps (2 identified within group B) were streamlined by moving to one "check" step, without introduction of new risk. Manual data entry, identified as a risk within two steps (RPNs 12 and 14, group A) was eliminated (recommendation 4) increasing efficiency and reducing risk.

4. Discussion

4.1 Research limitations/study design flaws

This review was the first FMEA analysis conducted within the study organization. Training given to the project team on FMEA and the use of this tool as a continuous improvement methodology was undertaken by the PO. Time for dedication to the project was limited for all participants due to daily work constraints. There are flaws to be noted in this study.

4.1.1 The rating scale. The rating system was not as clearly defined as it could have been and would be made more specific if the study were to be repeated. For example, criteria were not specified for values between 2 and 9 of the rating scales. This meant that there was an increased level of subjectivity to the assessments of S, O and D scores. The gap observed in the S and D ranges in Group A (Table VI) may be due to this flaw – however this did not appear to hinder identification and resolution of issues. The modified rating, used in analysis of the final FMEA, which incorporated customer expectations (Table III) would have been more useful to the team at an earlier stage. If this analysis were to be repeated, or applied to another process, the S, O and D scales would be defined with intermediary numbers (expansion of Table III for instance), and tested before implementation.

4.1.2 Team involvement. Involvement of team members is all steps of the FMEA's development and analysis was limited due to time and work constraints. The team consisted of a specialised group of genetic scientists. Their primary role was to process patient samples in order to meet TATs on, or before due dates, in order to assist with the timely delivery of clinical care which is critical to the health outcomes of the patient.

Therefore, although meetings were scheduled for times agreed to by team members, not all participants were able to make face-to-face meetings. Elements of the project had to be undertaken via e-mail, with subsequent individual reviews of the draft FMEA then being compiled by the PO. This may have resulted in a gap in consensus decision making and a reduction in the reliability of the final scores. More robust data may have been obtained through greater team involvement. Were FMEA to be repeated, a more robust planning and team selection process (also including a broader cross-section of staff, as outlined in Section 4.1.4) would be applied to provide as many viewpoints as possible.

4.1.3 Rating bias. The data set in this study was small, consensus meetings were not able to be held, and scores were averaged, thus leaving room for skewed results. Team consensus could reduce bias, and use of a nominal group technique would add robustness to a further study.

4.1.4 Cross-functional team. The team was not "cross-functional". There was no input from medical (i.e. doctors) or management staff, and the input from laboratory staff was limited (Section 4.1.2). This may result in a rating scale bias, as the genetic laboratory scientists involved in the process may not have the same extensive knowledge of the consequence of a potential failure to the patient. Thus, the overall "effects of failure" may not be as reflective of clinical practice as they could be.

There was no direct input from the designer and manufacturer of the GMT and reasons for a failed or invalid test within a "black box" procedure are not always transparent and visible to laboratory staff using the test.

Future improvement initiatives would involve introduction of the aims and scope of the project to a broader cross-section of staff (rather than just the genetic scientists). This would assist with gaining "buy in" and involvement from a more multidisciplinary and cross-functional team, thus strengthening the review process.

4.1.5 Ambiguity within failure mode statements. There is some vagueness within the failure mode statements. For example "delayed result" is actually an effect. Such ambiguity may unintentionally bias ratings e.g. an S score can be influenced by the length of a delay and how urgently a result is needed. In Table IV, the 20th and 23rd highest RPNs are examples of poorly worded modes.

4.2 Benefits achieved and alternative approaches

Although this FMEA study has some failings and there are opportunities for improvement, it yielded benefits, within review of a "black box" GMT process. New improvements were identified in spite of the previous review, including one significant preventative action.

4.2.1 A reduction in the scope of the study. The PO required extensive time, outside of daily work, to compile findings from the 88 failure modes that resulted in the eight recommendations from the 26 per cent highest RPNs. A smaller scope (sub-process) may have been more manageable for using the FMEA tool, reducing fatigue from data overload, and making compilation and analysis less taxing. Some authors report a smaller scope can result in more specific recommendations (Van Tilberg *et al.*, 2006), as FMEA can result in many potential failure modes that are irrelevant (Lipol and Haq, 2011).

4.2.2 Alternative approaches to FMEA ratings. Several alternative approaches have been identified in the literature rather than the linear approach used here, and have been recommended to improve the efficiency of analysis. These include:

(1) prioritization matrices: to determine which areas to investigate further, a prioritization matrix could be used before tabulation, to narrow efforts to process steps and inputs that have the biggest effect on customer (Rodriguez-Perez and Pena-Rodriguez, 2012);

(2) hazard scoring matrices: these could be used after listing failure modes (Van Tilberg *et al.*, 2006) to prioritize areas for action;

(3) Pareto analysis: a Pareto chart could be used to rank and display the top RPNs (McCain, 2006; Lipol and Haq, 2011); and

(4) multiple analyses: analysis could be split between two teams for S and O and D rankings after cause-and-effect analysis (Ramu, 2009) thus adding another level of assessment and robustness to the analysis process.

4.2.3 Utility of FMEA as a model. Regardless of the approach, there are questions regarding the utility of FMEA. Studies have questioned the validity of the ranking system, with calculations for the RPN being mathematically flawed (Shebl *et al.*, 2012; Bowles, 2004). The same RPN number is possible from different combinations of the S, O and D scores, however the meaning behind each rank is different, and there is a limited number of possible RPN scores (Shebl *et al.*, 2012; Bowles, 2004). Where the same RPN is found, prioritization by S, then O then D may be used (Lipol and Haq, 2011). An alternative method uses S, O and D or S and D scores without multiplication (Reid, 2009).

4.2.3.1 Definitions of scales. Doubt over validity of RPNs for prioritization is further strengthened by possible subjectivity when choosing scores for each failure mode. Definitions of scales used should be customized to the area of review (McDermott *et al.*, 2009), and designed by the team for clear understanding to reduce variation of interpretation. Data from the current situation, collected beforehand, may reduce variation. In situ simulations can strengthen scores, rather than relying on brainstorming "from memory alone" (Davis *et al.*, 2008). Voting can bring consensus, or using two teams to share and peer review results (Ramu, 2009). Interpretation of the importance of a factor – S, O and D, regardless of approach, are however, still influenced by personal values and belief. One study illustrates the interaction of factors with the outcome of FMEA conducted simultaneously and independently by two teams on the same process, with different (but both useful) results, influenced by the definition of the scales chosen, the approach (use of a SOP or walk through) and subjectivity of individuals (Oldenhof *et al.*, 2011).

4.2.3.2 Interpretation of modes and effects. Analysis is not always straightforward for failure modes. Modes can be confused with effects (Teng *et al.*, 2004), or may have multiple possible effects, and ideally should be listed and ranked separately (Ramu, 2009). Rating each possible effect separately, and for all customers, would be cumbersome, particularly for a large scope.

Severity scores may not be discrete due to a variety of environmental factors outside of the analysis scope. Within healthcare S to a patient can vary depending on their symptoms and clinical needs. Additionally the effect of the same failure and S, could vary from patient to patient due to the inherent variability in health. Thus S is not discrete – and should the majority or the average situation be used? It is difficult to rate S without bias when occurrence for the majority (and thus S), is low. If a serious adverse event is possible however then one could argue that this cannot be ignored.

This FMEA, study used the most severe possible effect to the patient for S scores thus even with an O of 1, data are skewed. An example of this is for the second, third and fourth highest RPNs (Table IV), with low O – all could result in "some tumour" or "no tumour" being removed, depending on the content of the whole specimen (outside of the intended target), varying from patient to patient. In this study the "worst-case scenario" was used.

4.2.3.3 Use for prioritizing. The interacting factors mentioned, suggest that RPNs should not be used on their own, and should be more of a guide, particularly within healthcare (Shebl *et al.*, 2012). A low RPN does not exclude the possibility of O. Improvements may also have unintended upstream or downstream effects; RPN can be calculated to examine possible unintended effects (Williams, 2010). Where prioritization is unclear however,

ratings may be useful. High scores of S, O and D are worth examining, individually and together. Reducing S (severity to the customer) is a priority, however high O and D with low S may also be harmful to the internal/external customer, particularly in healthcare. For example, if results are frequently delayed, work could be lost, or clinical management delayed, thus having possible compounding effects outside of the scope of the review.

Some approaches support examining issues over a certain number with an aim to reduce RPNs by a chosen amount. However this approach may detract from the true objectives, of improving procedures to consistently meet customer requirements. FMEA is a living document, with the contributing factors to risk, changing with the varying influences of the environment. An FMEA should be reviewed regularly to identify new risks, and after implementation of a corrective or preventative action (McCain, 2006; Tague, 2004). In analysis, it should be remembered that FMEA is a snapshot in time of many interacting factors.

5. Conclusions

For any improvement tool to be successfully integrated into an organizations' quality improvement or risk management programme, it needs to be relatively easy-to-use and proven to provide benefits to the customer and organization. Many healthcare organizations are facing fiscal constraints and increasing complexity of tests, putting strains on resources, particularly for involvement of those on the "shop floor", and who are "hands on" in the production of a product or service.

FMEA was useful within this review. Two relatively brief reviews by the team, increased awareness of risk and improved controls, documentation and workflow within the GMT. The ranking system illustrated concordance with needed actions (eight recommendations from 18/23 top RPNs, and seven minor changes in the lower 63 RPNs). However prioritization using the scores were not required as needs to fulfil customer requirements were obvious and solutions simple to implement.

FMEA broadens the scope for improvement, improves the depth of understanding of a process and identifies the possible failings within process steps. Within this study, the FMEA provided further improvements to an already reviewed process that was perceived to have few problems.

FMEA does not require complicated statistics however, it does require time, patience and planning for efficient and effective application. A team with current detailed knowledge of the area under review, including all inputs and all customers (a cross-functional team) will increase the power of its application.

A clear scope with defined scales for ranking are also required. The rating system can be a useful guide however; analysis of the causes, controls and effects of the possible failures may be enough in itself to point the way for action to minimize risks to the customer. When definitions of S are clear and focussed on the customer, actions may be obvious or urgent. Where risks can be minimized with simple solutions to improve controls without introducing new risks, it may be possible to implement improvements immediately.

References

Bowles, J.B. (2004), "An assessment of RPN prioritization in a failure modes effects and criticality analysis", *Journal of the IEST*, Vol. 47, pp. 51-56.

Davis, S., Riley, W., Gurses, A.P., Miller, K. and Hansen, H. (2008), "Failure modes and effects analysis based on in situ simulations: a methodology to improve understanding of risks and failures", in Henriksen, K., Battles, J.B., Keyes, M.A. and Grady, M.L. (Eds), *Advances in Patient Safety: New Directions and Alternative Approaches (Vol. 3: Performance and Tools)*.

Deming, W.E. (1986), *Out of the Crisis*, MIT Centre for Advanced Engineering Study, Cambridge, MA.

Imai, M. (1997), *Genba Kaizen: A Commonsense Low-cost Approach to Management*, McGraw-Hill Professional, New York, NY, p. 13.

Lipol, L.S. and Haq, J. (2011), "Risk analysis method: FMEA/FMECA in the organizations", *International Journal of Basis and Applied Sciences*, Vol. 11 No. 5.

McCain, C. (2006), "Using FMEA in a service setting", *Quality Progress*, available at: http://asq.org/data/subscriptions/qp/2006/0906/qp0906mccain.html (accessed 22 July 2013).

McDermott, R.E., Mikulak, R.J. and Beauregard, M.R. (2009), *The Basics of FMEA*, 2nd ed., Productivity Press, Taylor & Francis Group, LLC.

Oldenhof, M.T., van Leeuwen, J.F., Nauta, M.J., de Kaste, D., YMCF Odekeren Rombouts., Vredenbregt, M.J., Weda, M. and Barends, D.M. (2011), "Consistency of FMEA used in validation of analytical procedures", *Journal of Pharmaceutical and Biomedical Analysis*.

Ramu, G. (2009), "FMEA minus the pain", *Quality Progress*, Vol. 43 No. 3, pp. 36-42.

Reid, D. (2009), "Major upgrade. New FMEA manual offers more flexibility", *Quality Progress*, Vol. 43 No. 5.

Rodriguez-Perez, J. and Pena-Rodriguez, M.E. (2012), "Fail-safe FMEA: combination of quality tools keeps risk in check", *Quality Progress*, Vol. 45 No. 1, pp. 30-35.

Shebl, N.A., Franklin, B.D. and Barber, N. (2012), "Failure mode and effects analysis outputs: are they valid?", *BMC Health Services Research*, Vol. 12 No. 150.

Tague, N.R. (2004), "Failure modes effects analysis", available at: http://asq.org/learn-about-quality/process-analysis-tools/overview /fmea.html (accessed 22 July 2013).

Teng, G.S., Ho, M.S., Shumar, D. and Liu, P.C. (2004), "Implementing FMEA in a collaborative supply chain environment", *International Journal of Quality Management and Reliability Management*, Vol. 23 No. 2, pp. 179-196.

Van Tilberg, C.M., Leistikow, I.P., Rademaker, C.M.A., Bierings, M.B. and van Dijk, A.T.H. (2006), "Health care failure mode and effect analysis: a useful proactive risk analysis in a pediatric oncology ward", *Quality and Safety in Health Care*, Vol. 15, pp. 58-64.

Williams, T. (2010), "Minimizing risks: how to apply FMEA in services", available at: www.isixsigma/tools-templates/fmea/minimizing-risks-how-apply-fmea-sevices/ (accessed 13 August 2013).

Further reading

NZS/ISO 15189 (2012), *International Standard, Medical Laboratories – Requirements for Quality and Competence*, 3rd ed.

About the authors

Karen Claxton has completed studies towards a Post Graduate Diploma in Quality Systems and works full-time in the laboratory of New Zealand's largest metropolitan hospital. This study was undertaken as part of her studies that required a significant industry improvement project to be undertaken using learning from previously studied papers in the Quality Systems Suite of Programmes at the Massey University.

Nicola Marie Campbell-Allen is a Member of the School of Engineering and Advanced Technology at Massey University in New Zealand. Her specialty areas include leadership and management development, quality management, process improvement and business excellence. Nicola Marie Campbell-Allen is a National Evaluator for the New Zealand Business Excellence Awards, and joint Co-ordinator for the suite of qualifications available within the Quality Systems framework taught at the Massey University. Nicola Marie Campbell-Allen is the corresponding author and can be contacted at: n.m.campbell-allen@massey.ac.nz