Root Cause Analysis in Clinical Environments

Jeffrey C. McIlwain
Consultant, Clinical Risk Management
St Helens & Knowsley Teaching Hospitals NHS Trust
Warrington Road, Prescot, Merseyside L35 5DR
jeff.mcilwain@sthk.nhs.uk

Root Cause Analysis (RCA) is exactly what it says - a means to ascertain the facts that have lead to an event and so help evaluate what "went wrong".

Contextual Setting

The origins of RCA can be found in industry and engineering, when major catastrophes such as the Piper Alpha oil rig fire in 1976, the Challenger space shuttle disaster in 1986 or the more recent Grayrigg derailment in 2007 near Preston resulted in major investigations to ascertain "why?"

Over the decades risk-averse industries in the UK have developed specialised investigation units such as the Civil Aviation Authority and the Rail Accident Investigation Branch. In healthcare by contrast, there have been few or belated attempts to understand why clinical errors have occurred. The traditional response has been silence amidst a court case of alleged medical negligence. Over the past two decades there has been growing concern about events that have caused harm to patients and to find out why and to ascertain if ways can be found to prevent a recurrence. Formally this (aside from academic interests) has become encapsulated in the formation of the National Patient Safety Agency (NPSA) which has developed reporting mechanisms. This has been linked to changes in medical indemnity. Previously the NHS as a quasi-governmental organisation had Crown Immunity from prosecution. This has gone, to be replaced with crown indemnity whereby NHS bodies or NHS employees can be prosecuted. To insure itself against prosecution and so liability for medical negligence and third party liability NHS organisations insure themselves with the NHS litigation Authority (NHSLA) which also has some links with the NPSA. So various systems such as incident reporting and risk assessment have become established tools within the NHS, with RCA a relative latecomer.

RCA Principles

In effect the RCA process differs little from standard clinical consultation which in itself aims to elucidate causality and so guide clinical management. Clinical history taking, examination, forming a differential diagnosis and further investigation mirrors the RCA process and so is relatively familiar territory to clinicians. There are various well documented examples of RCA methods, but there are three that predominate in healthcare settings: the "five whys", the care-service timeline and the "fishbone" more properly called the Ishikawa diagram after its inventor.

The underlying principles of each are the same: to ascertain in a formatted manner the who, what, when, where and then the why of how an event/accident/mistake occurred to "learn lessons from" and so develop "action plans" to prevent a recurrence. The detail differs between each method and the quicker easier methods can underpin subsequent more detailed analysis.

It is also a vital part that causal analysis is not about blame but about establishing facts – it must be a neutral act. This is in contrast to a criminal prosecution that begins with a suspicion and a collection of evidence to prove an individual is guilty beyond all reasonable doubt. Causal analysis is about assessing systems, equipment and humans; not individuals from the outset. Anyone who fails to understand neutrality in factual evidence collection and has a pre-set perception of the outcome will automatically prejudice the outcome report and lay themselves open to challenge for process flaw and prejudice.

The five "whys"

This is a quick and simple technique. If 'A' happened one asks "why" which leads to 'B'. Then ask why 'B' happened which leads to 'C' and so on. It is usual to end at the fifth "why" but may stop at three or be more than five.

Let's look at an example – the wrong drug given to a patient.

Q1. "Why did you do give the wrong drug to the patient?"

A1 "I was distracted".

Q2. "why were you distracted?"

A2. "I had to answer the phone"

Q3 "why did you have to answer the phone?"

A3. "because we have a policy that all inbound phone calls must be answered within 30 seconds".

Q4. "why must all phone calls be answered within 30 seconds?"

A4. "to ensure that relatives are not kept waiting and are fully informed."

Q5. "why is there such a policy?"

A5 "because we were getting complaints about phone calls not being answered or a delay in answering.

In five simple questions and answers it is apparent that a local policy of insisting that ward telephone calls must be answered within 30 seconds impacts upon safe clinical care. The cause has been elucidated. This is a quick and simple technique but doesn't get to an actual root cause. It is best reserved for simple errors or frequent errors were no actual harm has resulted.

The care – service timeline

For many more crucial investigations this is the workhorse methodology. The principle is to divide clinical management into two categories – clinical care (the actual "hands on" clinical work) and clinical service (the systems and processes that support and facilitate the clinical care). These are set against the timeline of an event -for example analysing a complete clinical pathway from admission date and time through daily times. It usually is a documentary review but can be supplemented by interviews of people involved to add more information. Such interviews can be appended as appendices to a final report.

To analyse and record the pathway of an event it is usual to construct three columns. The first column is the date and then time, the second the care and the third the service. A brief hypothetical example of an event timeline is as follows:

DATE	TIME	CARE	SERVICE
Saturday 02/02/02	00:45	Admitted via ED and triaged	
	00:50	Diagnosed hypothermia on chemotherapy ? neutropenia	
	01:15	No IV line inserted - no pump available	Bed booked on oncology
	03:56		Transferred to oncology to avoid target breach
	04:15		House officer called to give intravenous antibiotics and clerk
	05:15		House officer called again – busy
	06:00	IV access team sets up IV line	House officer called again – still busy
	06:30	Patient collapses	
	06:45	Resuscitation begins	Cardiac Arrest team called
	07:00	Patient dies	

This very simple illustration of an emergency admission of neutropenic shock resulting in rapid death resulted from an analysis of the timeline of micro-events both in clinical care and in clinical service. Whilst there are some positive aspects to the events it is apparent that deterioration has rapidly occurred and both care and service failed the patient in an obvious chain of events that culminated in death. The death may not have been preventable as a natural outcome but there are obvious questions that extend beyond the record as annotated - such as time of first antibiotic, intravenous line insertion, moving from a safe environment to a more unknown one, staff availability, roles and duties. It is easy to dismiss the above as unlikely – however this illustration is drawn from true events albeit made brief for illustration. It is also probable that the "that would never happen on my watch " viewpoint is thought to be true, however only by analysing the timeline sequence of micro-events can various and many aspects come to light.

Within our own Trust we have been using this tool in an adapted form for over a decade. This includes specific questions of quality to answer after the timeline facts and recommendations. It is a trust-wide system of peer clinical professionals and assessors who receive any cause of clinical concern for analysis as well as random sampling of over 200 cases. It is usual that there are often up to 26 micro-events in the chain, some of which are significant and some not. It is also true that some micro-events repeat over referred cases suggesting an embedded set of problems. Some of these are internal and some external – such as the effect of Government targets.

This method supplies much greater detail and logical sequencing of clinical pathways. It also, as shown, demonstrates cross-over between care matters and service matters. But it takes time to perform: staff require some training in assessing detail from all documentary sources and may become complex. It is important that analysing staff remain neutral and objective to fact elucidation. It is probably best used for moderate-harm events or high-frequency events.

General Matters of Causality

Before discussing the final large tool it is probably pertinent to provide some background understanding of what and why things go wrong.

Human factors

The human factors in causality (courtesy of Professor James Reason) can be divided into human error and human violation (figure 1).

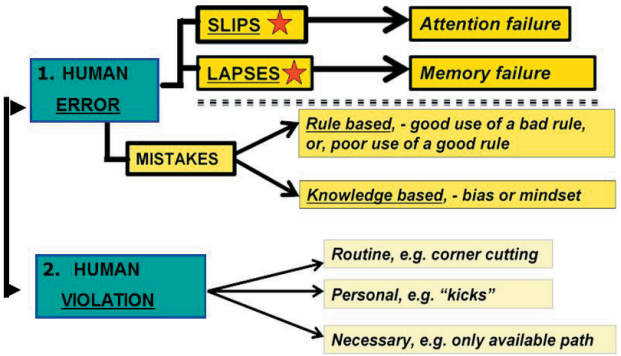

Figure 1: *Human factors in causality*

Human error can be subdivided into slips, lapses and mistakes. Slips are a failure in attention to the task such as being distracted and lapses are a failure in memory – simply forgetting. Slips and lapses are by far the commonest causes of human error. Mistakes are in a sense systems or knowledge failures i.e. rule-based failures or knowledge-based failures (in the latter a failure to keep up to date or a bias in the mindset). The rule-based mistakes are either a good rule used badly or a bad rule used diligently. The former has to do with training and education and the latter in policy/operating procedure design and implementation. So in summary we have in human error, slips, lapses, rule-based and knowledge-based mistakes.

Human violation is however a psychological mindset when something significant goes wrong e.g. an unexpected death. The unexpectedness and consequent grief may trigger investigation and the investigator may find it hard not to have prejudicial thinking - "this should never have happened" - and so veer towards blame. Yet human violation in an event is a rarity. Violation comes in three forms; corner cutting (the bored routine), personal (someone doing it deliberately for "kicks") and sometimes out of necessity i.e. deviation as this, at the time, was the only available path to follow. The former two have to do with unprofessional behaviour in an absolute form whereas the latter may be a proper professional decision, or not - as the case may be.

So in summary, that humans deliberately or with malice do harm is actually rare and it is more likely that distraction or tiredness or excess task needs leading to a slip or lapse is more common, which may be underpinned by a bad policy or a poorly implemented good policy that places the person in that situation. Altruism trumps media speculation in harm every time.

Other factors

There are in essence only three things that can go wrong – humans, equipment and systems. Equipment failure is rare as engineers design equipment and test it at the design, component, sub-assembly and assembly stages to form a final product. Human factors have been discussed above although work by James Reason can expand upon this in detail (see further reading).

System failures are more difficult to design, test and assess. The main systems within healthcare are policies, standard operating procedures (SOPs) training, education, competency assessing etc. These though are designed by humans for humans and have the potential to have inherent and dormant hidden errors that an event may expose. Once again turning to James Reason the classic visual model of threats and defences is the so called Swiss Cheese Model (figure 2).

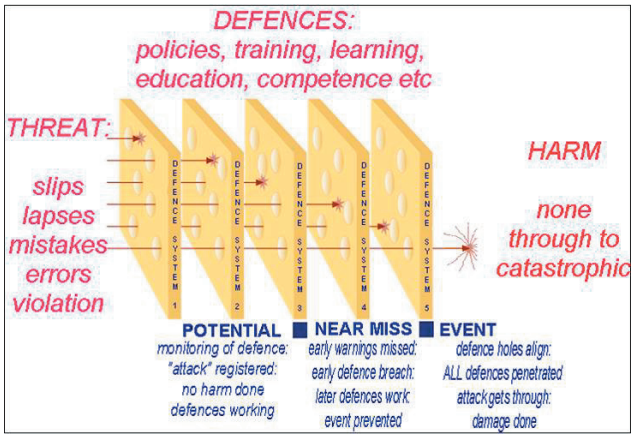

Figure 2: *The Swiss Cheese Model. As successive slices of cheese (representing defence systems) are traversed a potential event can become a "near miss" which can become a harmful event.*

In essence this describes several slices of Swiss Cheese that are characterised by having holes. Each slice is a line in your defences against harm of any description. If one stacks each slice side by side vertically it is usual that the holes do not line up so one could not pass a pencil through each slice through aligned holes. It is taken that each slice is a defence and the holes are inherent hidden errors. This is the natural state of affairs in systems - that the holes of each slice do not generally line up. So the defence stands due to this lack of hole [hidden error] alignment. But should the holes line up then a threat (human or equipment or system of process etc) may penetrate the defences and cause harm. In a sense a detailed version of this underpins the next and final RCA method – the Ishikawa diagram.

The Ishikawa diagram

This is more usually known as the "fishbone". It is a technique that many wish to employ, but is reserved for catastrophic events or effects because it is a large tool with significant resource

implications in time and staff. It was developed by Dr. Kaoru Ishikawa of the University of Tokyo in 1943. It has become an industry standard for events such as major engineering or industrial catastrophes such as the Piper Alpha oil rig, rail and air crash investigations. The visual description is of a fish skeleton with the head attached. The spine represents how factors come together. Each rib attached to the spine is a causal topic and side ribs to these ribs are the sub-topics. It has been adopted by NPSA in a toolkit and covers specific topics in clinical care viz.

- Patient Factors *(clinical condition & social factors & mental & psychological factors & interpersonal relationships)*
- The Individual's Factors *(physical issues & psychological issues & personality)*
- Task Factors *(guidelines and policies & decision making aids & task design)*
- Communication Factors *(verbal & written & non-verbal)*
- Team & Social Factors *(role congruence & leadership & support & cultural factors)*
- Education & Training Factors *(appropriateness & supervision & availability)*
- Equipment & Resource Factors *(equipment & supplies & visual display & integrity & positioning & usability)*
- Working Conditions *(administrative & design of physical equipment & staffing & time)*
- Organisational & Strategy *(organisational structure & policy, standards, goals & externally imported risks & safety culture & priorities)*

It is vital to proper RCA that every primary factor title and each consequent sub-factor is assessed both for any negative as well as any positive or correct procedures. Unlike a criminal investigation, fact-finding neutrality is a fundamental essential foundation that focuses upon all aspects pertaining to the event - and some more distant -embracing systems, humans and equipment. It is not and never has been a method to examine an alleged human factor alone. Neutral, reasonable and fair are its critical watchwords. No further discussion around the Ishikawa diagram will be given as anyone interested in this technique must be appropriately trained in its use; it is not an "off-the-shelf" technique to be used as anyone desires.

The Report

In brief the distillation of the facts will illustrate event chains and also priorities within the chain links. A report should highlight good and bad practice, avoid apportioning blame unless appropriate and reasonable and should analyse the event to formulate remedies and other avoidance techniques. Report writing is a unique and separately learned skill and art form and as usual must not be prejudicial. Interim reports may be hurried, provided it is clear that such a report is noted to be in draft format and not final. Recommendations should be a part of any report based upon either the report writer's knowledge or other expert assistance.

Summary

RCA is a collection of various techniques that are neutral in opinion and by the nature of their implementation are retrospective in approach. They form part of the risk management armoury that embraces the prospective risk assessment tool that seeks to prevent and analyse potential events or effects. This means that anyone interested in risk management must familiarise themselves with risk avoidance tools (risk assessment) as well as risk event analysis tools (RCA). They will then find preventative measures and learning points and so dynamic management of risk factors that in themselves may alter or reduce future risk events. The methods are proportionate in resource allocation

to the gravitas of the risk event. Whilst risk assessment can be part of system planning, root cause analysis can influence this by the subsequent report.

Further Reading

It is difficult to get a good all round viewpoint of risk management. One document that has excellent background comes from the aviation industry specifically ICAO (International Civil Aviation Organisation) who have a huge document that can be downloaded as a pdf and in particular Chapters 2-5 & 7

http://www.icao.int/anb/safetymanagement/DOC_9859_FULL_EN.pdf is the link.

This gives a good generic background to various techniques from a non-healthcare but risk-averse industry perspective.

The NPSA have also published tool kits for e-learning or group learning at

http://www.nrls.npsa.nhs.uk/resources/?EntryId45=59847

Finally for a textbook: Clinical Risk Management published by BMJ publishing group in 2001. 2nd edition Editor Charles Vincent ISBN 0 7279 1392-1. In particular chapter one by James Reason covers human factors to an excellent high level as well as Chapter 2 by Eric J Thomas and Troyen A Brennan.

This paper is based on a talk Dr McIlwain gave at Darlington Memorial Hospital in October 2011.