People make mistakes. Systems fail. There is no getting around it. Risk Management and Patient Safety professionals have made great strides in understanding the human and system factors that effect performance and cause mistakes, errors, and failures. This increase in understanding has moved risk managers away from asking “what happened” and “who did it” to understanding “why”. It is through this understanding of “why” a failure occurred that effective safeguards preventing reoccurrence can be found. For more than a decade, Healthcare Risk Managers have been using a structured process called a Root Cause Analysis (RCA) to investigate the failures which cause (or could have caused) serious patient injury.

Similarly, the Failure Mode Effects and Criticality Analysis (FMECA or FMEA) is an analysis methodology that has been effectively used in many industries for years. The purpose of the FMEA is to proactively look at the components in a process and determine how each might fail. This is done before an actual failure so that the system can be designed from the beginning not to fail. The FMEA has only relatively recently been adapted to healthcare, but the impact has been tremendous. New processes, procedures, and care delivery methods are being “tested” to see if there are weaknesses that could reach a patient and cause harm. Corrections in design are being made before an injury occurs.

The RCA and FMEA are not only effective in situations where there is actual or potential physical harm, they are also effective were there is any real or potential loss. Because they are resource intensive tools, they are most effectively used where the loss is significant to the organization. As we all know, failures in our care delivery systems, documentation systems, coding, billing, and reimbursement systems can cause a significant loss to an organization. The RCA and FMEA are therefore perfectly suited to examine real and/or potential failures in our compliance programs.

To understand how to effectively use the Root Cause Analysis and the Failure Mode and Effect Analysis one first needs to understand a bit about human factors and the systems approach to failures.
Understanding Human Factors Analysis

James Reason, a British psychologist and professor studied aviation safety and the human factors that caused plane crashes. His “theory of error” has been used by the airlines, military, nuclear power plants, automotive manufacturers, financial institutions, transportation, food handling, and many other industries. In 1990, Reason differentiated between active failures and latent failures. According to Reason, active failures are those actions or inactions of individuals that appear to “cause” the failure. In a patient safety context, this would be the nurse who gave the wrong medication or the surgeon who cut off the wrong leg. In the compliance context this may be the coder who routinely codes complex pneumonia instead of simple pneumonia or the registration clerk who never provides an Advance Beneficiary Notices when required.

Latent failures on the other hand, are the errors committed by individuals somewhere else in the chain of command or elsewhere in the process. This could be the supervisor who refuses to correct known problems or the person doing employee training who is not familiar with the process being trained. It could be poorly written policies, unclear instructions, incorrect or incomplete data, poor hiring practices, ineffectve discipline for violations, poor communication, unclear roles, overly complex processes, an organizational culture that does not expect members to follow the rules, or any number of other factors. We can begin to see that a mistake (the active failure) may be the end result of a long chain of other, often organizational, problems. Unfortunately, these latent failures are often difficult to detect and can exist in an organization for a very long time before there is an active failure.

In 1974, Frank Bird promoted the “Domino Theory” where mishaps are the end result of a series of errors. Bird likened the process that leads to injury or damage to a row of dominoes standing on edge. When any domino falls, it can disturb the other dominoes and a sequence is started which eventually leads to the fall of the final piece (the active failure). The theory implies that if any one of the dominoes is removed from the sequence, or is strong enough to withstand the impact, then the chain of events will be broken and the ultimate event will never happen.

James Reason presented an updated version of the domino theory often referred to as the “Swiss Cheese” model. Reason suggests that within complex process there are a number of barriers or safeguards that are established to protect the system from failure. These safeguards may include education and training, hiring, policies and procedures, technology controls, supervision, oversight, auditing and monitoring, culture of compliance, etc. Unfortunately it is rare that any safeguard is perfect, like a piece of Swiss cheese, it will have some holes. For example, you may have a very comprehensive training program however it is so dull no one can stay awake through the entire program. In complex systems when one safeguard fails, we will generally have other safeguards that will catch a problem before a mistake is made. Sometimes these are formal safeguards; an excellent hiring and selection program providing you with well trained individuals. Often however, we develop informal back-up safeguards; the co-workers on the job know how boring the training is and “watch out” for the new guy by filling in
training gaps on-the-job. Reason proposed that there will be times when the holes in the safeguards line up just right and a problem (latent failure) will make it all the way through your safeguards and result in an accident, mistake, or other active failure.

As described by Reason and others, we need to look beyond the “what” and “who” to truly understand “why” something goes wrong and a system fails. Yale organizational theorist Charles Perrow suggests that a structured investigation requires a thorough review of Design, Equipment, Procedures, Operations, Supplies and Materials, and Environment (he uses the acronym DEPOSE). English psychologist Charles Vincent breaks it down into the elements of Institution, Organization and Management, Work Environment, Team, Individual, Task, and Patient.

In addition to reviewing the latent failures we need to look at the human factors which exist that affect systems performance. These human factors may appear as a component of the active failure or a latent failure. Generally human errors can be categorized as skill-based errors, decision based errors, and perception based errors. These are the “honest mistakes” we all make without intending to. These are generally products of the way the system is designed and can be managed through system redesign, improved communication, training, education, changes in procedures, technology improvements, and improved work environment.

Skill-based behaviors are the basic skills that are performed without significant conscious thought. For most of us, this is like driving a car. You don’t really think about what you are doing, you just do it. Failures in this type of behavior are often failures of attention or reliance on memory. For example, you get in the car to go to the grocery store and end up going to work. You didn’t choose to make the mistake, you just weren’t paying attention to what you were doing.

Decision errors are intentional behavior that proceeds as intended, yet the plan is inadequate or inappropriate for the situation. These are generally the result of inadequate training, incomplete or misinterpreted information, or simply poor choices given the circumstances.

Perception errors occur when we don’t have enough information so we use our “best guess”. Errors occur here when we incorrectly “fill in the gaps” because we make assumptions about how things usually work.

Additionally, there is unintended at-risk behavior. This occurs when an individual is unaware of the risks or are given some incentive for the at-risk behavior (i.e. better performance reviews for higher productivity even when at the expense of quality). Errors may also occur when individuals routinely violate rules. These routine violations are like driving 5-10 miles an hour over the speed limit. You rarely get caught, the increased risk is minimal, there is some incentive, and “everyone does it”. These are the policies that no one ever follows but don’t seem to ever get rewritten or removed from the manual. Routine violations are often ignored by supervisors and may even be encouraged until a significant error occurs.
Finally, there is the exceptional violation or the reckless behavior. These errors and behaviors are not the product of system design, but are willful and intentional and can only be managed through appropriate employee disciplinary action.

Once we understand that most errors are unintentional, products of the way our current systems are designed, and that latent failures are as important to understanding why an error occurred then we can begin to understand how to effectively use the tools for investigating errors (Root Cause Analysis) and proactively addressing potential system failures before they occur (Failure Mode and Effect Analysis).

**Root Cause Analysis**

Simply stated, a root cause analysis is a process designed for investigating what, how and *why* something happened and to figure out how to prevent the same thing from happening again. Understanding why an event occurred is the key to developing effective recommendations. Imagine you find a registration clerk who does not provide Medicare Beneficiaries with the Advance Beneficiary Notice (ABN) when it is required. The typical compliance investigation would probably conclude that the registration clerk was the cause of the error. However, stopping here in the investigation only tells us what happened and how it happened. We have not probed deeply enough to understand the reasons for the mistake and therefore, don’t know how to prevent its reoccurrence. In this case we are likely to see corrective action that involves some retraining, employee discipline, and/or reminders to pay more attention to what the individual is doing.

As discussed above, mistakes do not just happen, but can be traced to some latent causes. In this case we might ask, “Why didn’t they follow the procedure?” “Is the procedure clear, understandable, easy to follow?” “Are we asking the employee to interpret clinical information beyond the scope of their training?” “Are there incentives for not obtaining the ABN such as productivity bonuses?” “Do the employees have someone to ask if they have questions?” “Are any other employees having the same issues?” Each question you ask will lead you to more questions. This constant questioning, asking “why” “why” “why”, will help determine the root causes so that corrective actions are meaningful in preventing reoccurrence.

The RCA is a four-step process. The first step is to collect the data. Generally a team of individuals is brought together to review what happened, when, how, and begin to explore why. This may include participants or witnesses, individuals who are expert in potential aspects of the event (Human Resources, Information Technology, Clinicians, Coders, etc…), and organizational leaders who can authorize resources required for corrective actions.

The second step is to chart the causal factors. This provides a structure for investigators to organize and analyze the information gathered during the investigation and identify gaps and deficiencies in knowledge as the investigation progresses. The causal factor chart is a sequence diagram. When the entire occurrence has been charted out, the
investigators are in a good position to identify the major contributors to the incident, the
causal factors. The diagram will help to show the cause and effect relationship between
factors, even if significantly removed from each other in the system.

Step three is to identify the root causes. After all the causal factors have been identified,
the investigators begin the root cause identification. This step generally involves the use
of a decision diagram or “fishbone” diagram. This diagram structures the reasoning
process of the investigators by helping them answer question about why a particular
causal factor exists or occurred. For every event there will likely be a number of causal
factors. For each causal factor there will likely be a number of root causes.

The final step of the process is to generate recommendations for corrective action. A
summary table can help organize the information obtained during the data analysis. The
Joint Commission on Accreditation for Healthcare Organizations (JCAHO) offers an
excellent tool for this purpose.

**Failure Mode Effects and Criticality Analysis**

The Failure Modes, Effects and Criticality Analysis (FMECA or simply FMEA) is a
bottom-up analysis of component-level failures and their effects on higher-level systems.
FMEAs differ from Root Cause Analysis (RCA) in that they are intended to result in
preventative actions, they are not “after-the-fact” exercises done to determine the root
cause of a failure that has already occurred. The Automotive Industry Action Group
describes a FMEA as a systematic group of activities intended to:
1. Recognize and evaluate the potential failure of a product or process and its effects.
2. Identify actions that could eliminate or reduce the chance of the potential failure
   occurring.
3. Document the process.5

The FMEA was developed in the United States Military in 1949 as a reliability evaluation
technique and has been used effectively in many industries for years. Effective July 1,
2001, the Joint Commission on Accreditation of Healthcare Organizations (JCAHO)
implemented new standards in support of error reduction programs in hospitals.
Articulated in the intent of this standard the JCAHO identifies the need for a proactive
risk assessment that identifies failure modes, analyses of the effects and seriousness of
the failure modes, determines why the variation/failure could occur and redesigns the
process to minimize the potential for critical failures. The JCAHO has indicated that
using the FMEA technique is a good way to approach this standard.

The FMEA process is both rigorous and detailed. Done properly, it requires a significant
amount of time and resource.

The first step in conducting a FMEA is to define the process that you want to analyze.
Since you can expect to invest significant time and resources into the FMEA process, you
should select a topic that is high-risk or a high-vulnerability area. For example an FMEA
may be an excellent tool to use if you are considering adding a wound care program since
the OIG has cited concerns related to medical necessity, billing, and appropriate controls in this setting. It is expected that the FMEA process be evidence based. That is, you have data available to support your analysis. You will need to look at external data such as OIG audit reports, the OIG work plan, peer-reviewed literature, DHS and/or other external data. You will also want to collect data using your internal audit and monitoring systems as well as interview individuals who are knowledgeable about the process.

Once you have selected your topic for review you will want to put together a multidisciplinary team. This team should include subject matter experts and represent various viewpoints. The team should include not only management, but also staff who actually carry out the work. You may want to include someone who knows very little about the process to identify holes and assumptions made by the experts.

The team will first break the process down into its component parts, or sub-processes. For example, if you are looking at your ABN process, you might break that down into obtaining a complete physician order, screening for medical necessity, delivery of the ABN to the patient, coding, billing, etc. Once the team has identified the sub-processes, they will decide which ones they want to look at. With complex processes, your team may be overwhelmed if they try to look at everything. Start with those sub-processes that you know to be most problematic. The team will flow-chart the sub-process to be examined.

For each step in the process, the team will determine how that step could fail (failure modes). Using our ABN example, if your sub-process step is to check current Local Coverage Decisions (LCD) the failure steps may include (1) automated screening system down, (2) LCD information not updated, (3) unable to read physician order, (4) physician order does not provide diagnosis, etc. The team will do some brainstorming to determine the failure modes but should also rely on source material such as OIG audit reports, internal audit reports, OIG’s workplan, etc.

For each failure mode, the team will define the failure effects. Ask the question, “What would likely happen if this failure occurred?” Determining the effect of failure is as important as defining failure itself. In our ABN example above, the effects of the physician not providing diagnostic information on the order may be delay in registration, patient dissatisfaction, failure to obtain ABN, obtaining an ABN unnecessarily, loss of revenue, improper registration, physician dissatisfaction, etc…

Once you know the failure modes and effects the team will ask:

1. “What are the chances this failure will occur?” This is called the Probability of Occurrence measure. Typically, this is expressed using a scale of 1 to 10, 10 being very likely that the failure will occur.

2. “How serious would it be if this failure occurred?” This is called the Severity of Failure. Again, this is expressed with a score of 1 to 10, 10
meaning that failure would be catastrophic.

3. “What are the chances that this failure would be detected before it is too late to fix?” This is the Difficulty of Detection score. On a scale of 1 to 10, 10 would mean that it would be very difficult to detect this failure before it could be stopped.

You will get a “Risk Priority Number” by multiplying the probability, severity, and detectability scores together. For example if through your data collection you determined the following:

Sub-process: Outpatient registration screening for medical necessity
Failure mode: Physician did not provide diagnostic information
Probability of Occurrence score – 7 (pretty likely)
Severity of Failure score – 3 (not particularly severe)
Detectability score – 2 (highly likely you will notice this and correct)
Risk Priority Score = 7x3x2 = 42

Once you have calculated your Risk Priority Score for each failure mode, you can rank the failure modes that are more likely to create your most significant loss. The higher the Risk Priority Score the more likely to adversely effect your compliance. You will concentrate the remainder of your analysis on only the “significant few” failure modes.

The next step in the process is to identify the causes for the failures. For each significant failure mode you will conduct a “mini” root cause analysis. Even though the failure is theoretical, you will want to look at the human factors and latent factors that could contribute or cause the failure. Recommendations for improvements and corrective actions will be developed. Corrective actions should be piloted to ensure that they are effective and do not in fact create other failures in the system. If effective, the corrective actions can be fully implemented and should be monitored to ensure on-going effectiveness.

Using the Root Cause Analysis and Failure Mode and Effect Analysis in your Compliance Program

Risk managers and loss prevention specialists have effectively used the root cause analysis and failure mode and effects analysis tools in many industries for years. Though newer in healthcare, these tools have proven to be extremely effective in improving patient safety. They are valuable tools used to flush out the root causes of failures and to identify potential failures before they occur. They evaluate not only what happened or could happen, but look at the entire system to identify weaknesses in safeguards (the holes in the Swiss cheese) and processes. They look beyond the active failure to identify the latent system failures as well. They propose solutions that can and do prevent the reoccurrence of failures. Similarly, these tools can be used in our compliance efforts to better investigate and analyze failures and to proactively evaluate processes before they become compliance failures.
Compliance professionals clearly understand the dire consequences, to our organizations and individuals within our organizations, of significant compliance failures. Putting these proven tools to work within our compliance program will go a long way to mitigating the potential for these failures. As you begin to evaluate the use of these tools in your program there are some practical realities. The RCA and FMEA take time, energy and a commitment of resources to effectively conduct the process and to correct issues found. Here is my advice to you:

**Pick your processes for review carefully**

Not every investigation will require a formalized RCA. Not every process requires evaluation through a FMEA. A formal, well-documented RCA should be conducted when you are investigating a significant compliance failure. It should be used in those situations where reoccurrence would have a significant impact. You would not typically use a RCA for simple mistakes, inadvertent slips, or one-time events. You would consider using an RCA when those simple mistakes turn into a trend.

An FMEA is a particularly good tool to use when considering a new service, particularly if you are adding a service that has been identified by the OIG through their workplan or audit reports to be of concern (pain management programs, wound care programs, skilled nursing, partial hospitalization programs, research programs, etc.). The FMEA is also a good tool to use when considering deleting a service, program, or department. Sometimes when services and programs are eliminated, important safeguards are also inadvertently eliminated. A FMEA would be an excellent tool to identify these potential holes.

**Involve others in the work**

You can’t do this alone. You will not have the required expertise in all areas to evaluate. You can’t possibly have the time or energy it takes to thoroughly evaluate the root causes of a failure nor to evaluate processes that may result in compliance failures. You will need to involve others in the organization in the work.

**Participate in your organizations risk management program**

If you are working for an accredited organization, you will find individuals in your organization that are already expert in the use of the RCA and FMEA. Your organization will have adopted methods, forms, tools and resources for effectively conducting evaluations. Work with your risk managers to learn more about these tools as they are actually deployed in your organization. Find out what the barriers may exist in your organization to effectively using these tools. Ask if you can participate in the next RCA team or FMEA team, if only as an observer, to learn more about the process.

**Tag on to evaluations currently being considered in your organization**
Since your organization has likely chosen a process or two to evaluate using the FMEA, you may want to consider taking one of the sub-processes associated with that project and looking it from a compliance angle. For example, if your organization is looking at the medication delivery processes in the emergency department, you may want to look at your process for ensuring you are not billing Medicare for self-administered medications. As we know, the care delivery processes, quality of care and documentation of care provided are hot compliance topics. You will likely find compliance aspects that you could evaluate for virtually any project that your organization has chosen for a FMEA.

Consider it a way of thinking about investigations and processes

In many ways we are lucky, we don’t have accrediting bodies requiring that we use the RCA or FMEA in our compliance work. We don’t need to use the prescribed forms, we don’t have to convince anyone that our evaluations are “thorough” and “credible”. If we engage in this work, we do it because we want to find a structured way to find why systems fail and to prevent those failures. Even without the forms and teams you can use the concepts that drive the RCA and FMEA analysis. Understanding the need to constantly ask “why” to get to the real root of the issue, thinking about the human factors that may have contributed to an error, recognizing that latent failures in the system may have set the individual up for the failure, are all important considerations when you are investigating any failure. Similarly, thinking about the various components of a process and how failures may occur within those sub-processes should be part of your standard procedure for looking at your systems and the design of those systems to ensure compliance.

Learn more about human factors, error theory, RCA’s and FMEA’s

In addition to the risk management team in your organization there are a number of excellent resources on human factors, human error theory, root cause analysis and failure mode and effects analysis. Here are some of my favorites:

- VA National Center for Patient Safety, [www.patientsafety.gov](http://www.patientsafety.gov)
- Joint Commission on Accreditation of Healthcare Organizations, [www.jcaho.org](http://www.jcaho.org)
- National Patient Safety Foundation, [http://npsf.org](http://npsf.org)
- Agency for Healthcare Research and Quality (AHRQ), [www.ahrq.gov/qual/errorsix.htm](http://www.ahrq.gov/qual/errorsix.htm)
- FDA Patient Safety News: [www.fda.gov/cdrh/psn](http://www.fda.gov/cdrh/psn)

If you would like to learn more, I will be doing a session on the integration of FMEA’s and RCA’s into your compliance programs at the HCCA Compliance Institute in New Orleans, April 2005.

5 Automotive Industry Action Group, www.aiag.org