Qualcon 2005

Confidence and Safety – A key dilemma in planning continuing Health Services

Tony Stephenson. BE, ME, FIQA, Snr Mbr ASQ
Operations Manager: International Standards Certifications Pty Ltd. A MoH Designated Auditing Agency
PO Box 89 039 Auckland NZ
Telephone 64 9 473 4346 Fax 64 9 473 4329
Email: certify@certificationz.com

In association with

Doug Brown RN Lead Auditor Dunedin NZ
Rene Darier RN MOHS(Auck) CEO Corporate Wellness Ltd NZ
Jill Gray Lead Auditor Consumer Representative Whangarei NZ
Ron Guthrie MI Mech E Lead Auditor Food Safety IMAT Ltd Dunedin
Fiona Worth. CQA, Dip Bus Qual
Formerly Consumer Liaison Team Leader Auckland DHB

ABSTRACT

Risk in a clinical/medical/surgical and caring environment requires a multiple approach to analysis. Risk is more generally a combination of events, few of which in their own right will cause death or even minor injury. There is a need to separate those combined risks whose individual probability of occurrence is low from those risks that are due to absolute incompetence and for which personnel selection and peer review processes have to take full responsibility for adverse events.

Against a background of Healthcare experiences in the UK, Malaysia, Japan, South Africa, Australia and New Zealand ranging from full surgical and medical through to Age and Mental Care as well as support functions like Air Ambulance/Flying Doctor, the authors present the case for Confidence in the results of an audit in a diverse range of Healthcare settings within the current NZ Health and Disability Standards framework and other diverse Health settings including Aged Care, Drug and Alcohol Abuse, Intellectually handicapped and Dementia both in home and RestHome/Hospital Environments. The push towards more Home Care brings a new set of risks that have been swept under the carpet under the guise of better care; these will be explored further.

The authors conclude that the Audit can only have some valid significance in Healthcare settings if it is possible to set a level of confidence on the audit findings. The cost of the audit will be linked to the level of confidence based on the time taken to sample the various processes in the service(s) being audited. The level of confidence will be based on prior statistics [facts/evidence] and – in particular – how well the facility carries out and uses the results from it's own internal audits and management reviews. A further conclusion is that the NZ Health and Disability Standards fit well into an ISO 9001:2000 framework for what is - in effect - a management system of best practice, provided we benchmark, measure customer satisfaction in a consistent way and proactively seek "business improvement" at all management level as well as with clinical/medical caring processes. The JAS-ANZ recommended times are a good guide in the case of large multi site Healthcare facilities.

INTRODUCTION

Risk in a clinical/medical/surgical and caring environment requires a multiple approach to analysis. Risk is more generally a combination of events, few of which in their own right will cause death or even minor injury. There is a need to separate those combined risks whose individual probability of occurrence is low from those risks that are due to absolute incompetence and for which personnel selection and peer review processes have to take full responsibility for adverse events.

Against a background of Healthcare experiences in the UK, Malaysia, Japan, Australia and New Zealand ranging from full surgical and medical through to Age and Mental Care as well as support functions like Air Ambulance/Flying Doctor, the authors present the case for Confidence in the results of an audit in a diverse range of Healthcare settings within the current NZ Health and Disability Standards framework.

1. SOME BACKGROUND ISSUES

Financial Auditors have had available to them, for many years, a set of statistical tools, sometimes called "Discovery Sampling". The Audit Process may well be a “Journey of Discovery”, hence one of the reasons why an Audit needs to take into account the 24/7 requirements of the Healthcare sector. How well one might find out the use of seclusion practices at 5am or other alleged infringements of the Mental Health Standard during an audit visit is uncertain, especially if patient notes do not indicate such practices are occurring. Nevertheless it is essential that under the Health and Disability Standards such visits do occur. Whilst surprise audits are often viewed as unfair and against the spirit of “Value Add”, the fact still remains that in many industries, “modified” work practices outside 0700 to 1800 hrs, together with a reduced understanding of the implications of the changes is common. The earlier apparent issues in a mental Health Unit in Auckland are one such case for significant audit exposure both internally and to external agencies such as Designated Auditing Agencies. [The NZ Herald – Thursday May 8 2003].

Issues such as apparent “Burn Out” of midwifery and A and E staff through lack of staff has been reported on in a number of centres in recent months. Clearly when there are pressures on such staff to work extra or extended shifts, there are risks to patient safety. The Audit process will of course probe rosters, and the staff capability [skills] within those settings. Clearly the auditors can make the call to MoH if the team sees clinical risk at the time. The message here is that the Audit team may be satisfied as to current levels and skills; the long term scheduling of staff to meet projected maternity demand may also appear to be adequate, yet at that time due to illness or staff resignation the service comes under extensive pressure. Auditors need to probe forward and back as to skills and availability in order to project possible risk issues in the future based on prior knowledge.

If we are to believe the statistics of the prestigious Harvard School of Medicine’s latest report on 750 New Zealand patients, then one in four of those with serious health problems have been the victims of medical error. The sample size appears to be statistically sound and in effect large compared to later comments in this paper. So our auditors would have to probe records to see how many clients have seen say more than five clinicians, had duplicated tests, been given conflicting advice or had poorly handled co-ordination of care? No wonder that many Doctors have said privately that they view the audit process as another potential peer review. *Yet if the audit team does not probe in these areas, we are according to the Harvard School data, missing a one in four chance of perhaps preventing such a high error rate.*

It should not be forgotten, that auditing is intended to be positive process whenever possible. However remember the four key audit steps that a Designated Auditing Agency is striving for:

- Test that a management system is in place.
- Does it meet the NZ Healthcare Standards (or other) standard/legislation?
- Do they do what they say they do?
- Is the management system effective?

We will open up the debate with a series of examples and make reference to the Risk Management standards in Healthcare. A series of scenarios to set the scene!

2. RISK and AUDITS – THE WIDER ISSUES

The following scenarios are not untypical of many large international healthcare organisations bearing in mind that the bulk of complaints centre on communication, or its lack of or clarity, including waiting times. Add to this the diverse cultures within NZ and the communication barriers these bring at consumer and staff levels, we have a need for some accurate statistics if we are to manage risk in a meaningful, understandable, repeatable and value added way. The "silo" culture can still lead to changes in probability of success or failure due to lack of understanding/communication over interactions between specialisations.

a). **Patient** files contained irregular use of language to describe levels of pain [i.e. in depth description not complete, number of times complained] thus a risk of incorrect/inaccurate assessment/evaluation of patient pain management process. From say 4,000 care giving staff and any one patient can see typically in a weeks stay five clinicians and up to 10 nursing staff, small wonder that both continuity of care and interpretation is at risk.

b). The inwards referral and discharge process fails to clearly define the symptoms or the urgency of the need to have a specialist review or ongoing treatment that minimises risk in the public environment.

c). Assume 40,000 patient assessments in an acute setting per year, what is the statistical likelihood of a percentage of these assessments being inaccurate?

d). In a residential care setting [long term care] with multiple sites, staff turnover amongst care givers is 20%, high turnover will immediately jeopardise standard of care for the long term consumer due to lack of personal knowledge of the individual.

e). Procedures dealing with Treaty of Waitangi issues in a multi site Health Care environment with a significant number of Maori clients have not been developed to meet the requirements of the Health and Disability Standards in a number of aspects of the service.

f). The adverse clinical events discussed in HB 228 Page 95 Case Study 6 are valuable from a viewpoint of demonstrating what internal audit and management reviews could show in terms of performance and benchmarking.

g). Gaps that were identified in the Alfred Hospital Melbourne Clinical Risk Management Process study in 1998 should be seen as a clear benchmark for any other provider of similar services. Size should not be a major reason as to failing to have a system that has focussed in on the key issues found at the Alfred Hospital such as: -

- Much data collected, little accurately measures risk
- Inconsistency in definition of events, reporting follow up

- Development of pathways and protocols, their impact in risk unknown
- Uncertainty over data collection due to legal issues
- Data not disseminated to interested parties in the Hospital

2.1 Expected Cost of an "event"

In simple terms the "Probable cost" of a series of events is the sum of the expected cost of each event, if it occurs, multiplied by the probability of that event occurring? This is sometimes called "Expected Monetary Value [EMV], but works just as well with costs. Classic studies in this area include Space travel, aircraft design and the notorious Ford Pinto scandal.

Probable Cost of Event 1 = its probability of occurrence * the true cost of the event

Probable Cost of Event 2 = its probability of occurrence * the true cost of the event

Sum of probable costs of events 1…n becomes a simple addition of the EMV for each of the events. So long as we have a reasonably robust statistic relating to the Probability of the event, and an honest appraisal of the costs we can budget for adverse events. One small point is that we can only have a total probability of 1.

A non-trivial example follows in budgeting for a surgical procedure:

Pr of success is 0.7 and the cost is $10,000
Pr of failure is 0.3 and the cost is $30,000

The likely cost of the procedure is 0.7*10,000 + 0.3*30,000 = $7,000 + $10,000 = $17,000

Clearly true cost would be labour, materials, time, and compensation to customer, to name a few in Healthcare. Even without counting costs in a litigious society of legal fees and adverse judgments, the costs of rescheduling patients due to a need to carry out remedial surgery or other medical treatment assumes a significant sum. The wider availability of software and statistics such as the NHS [UK] "Safecode" outlined in Case Study 10 of HB 228 [See p115] would probably help to dispel much of the mystique in NZ Healthcare issues.

In recent months there have been models developed on Clinical /Surgical practice to enable administrators to make more informed decisions about when to take surgical and medical intervention using classical risk models.

2.2 The Current Risk Management Process in Healthcare

In the majority of situations we are faced with the use of both the Guidelines for managing risk in healthcare [HB 228:2001] as well as the individual element evaluations such as A 2.2.4 in SNZ 8134:2001 in the Hospital workbook. In the latter we are invited to make value judgments against a variety of options such as Figure 1.

Attainment level and consumer safety?

	√	Cr Hi **Mo** Lo Neg **Actions by whom and when**	**Possible evidence trail examples**
CI			
FA			
PA	√	The system is new and largely provided through templates. Little "buy-in" from staff, no previous history, provider not yet in full control of changes under new ownership. To re audit this area in 3 months by the DAA. Owner to also monitor progress.	System sighted in use. Discussed progress with Admin as well as staff at all levels. System access is through one staff member per shift
UA			

Figure 1 Typical Clause Outline for evaluation H and D Standards

3. A THEORETICAL MODEL

3.1 An overview

Currently many Management and Technology audits are carried out where the Repeatability (by the same auditor) and Reproducibility (another auditor using the first auditor's notes/CAR) would have a very low chance of success. Much of this failure is because the CAR'S and working notes often lack enough detail to repeat the finding. Many current audits could be just as well decided by tossing coins, so maybe confidence is less than 50%! One might also question just how many "populations" we have to sample from in a Health Provider environment; the following is a **very short** list: -

- Patient Records
- Staff Schedules
- Training Records
- Professional Development Records
- Menus
- Maintenance Records
- Procedures [How Many?]
- Instructions [How Many?]
- Medicine Round Records
- Pharmacy Orders and identity
- Physiotherapy Records
- Occupational Therapy Records
- Multi site variations

The presentation slides include additional ideas and data.

3.2 Some Fundamental Sampling

(a) If we randomly sample (say) 20 pieces of Medical Equipment and find 10 do not meet the requirements defined [e.g. calibration, maintenance], we can assume that approximately 50% of the population will show the same defect(s).

(b) If we randomly sample (say) 20 Patient Records and find 5 that do not follow the procedure/are incomplete, we can assume that approximately 25% of the population will show the same problem(s).

(c) If we randomly sample (say) 20 staff training records and find no people untrained/unqualified, we cannot say that the defect rate is 0%. There may be some in the workforce population that are still not trained fully or Not Competent without supervision; our sample did not detect them.

(d) If we **randomly** sample (say) 20 clinical test reports and find one with incorrect conclusions, we **cannot say** that 5 % of all test reports will be suspect! We may just have got lucky and found the only one! Nevertheless it is a defect (e.g. CAR).

(e) If we randomly sample (say) 20 procedures/ instructions in use and find that two, three, four or five are not being followed correctly, then we probably do have an ***increasingly/ cumulative knowledge*** of a base defect rate. i.e. two may still be by chance but three, four or five is certainly indicative that the real percentage defect is 3/20 [15%], 4/20 [20%] or 5/20 [25%].

The next section examines the more theoretical approach.

3.3 Sampling Approach to Auditing. (How confident are you)?

Where there are a very large number of procedures or items that require auditing (be they statistical procedures, equipment requiring calibration, clinical notes, work instructions or any of the other parts of the Health and Disability quality management system), is it necessary to check every possible item, with utmost care? This will often be impossible to achieve economically, so what compromise is made?

Statistical sampling of the items is the only sensible approach. To select some and report that no non-conformances have been found is inadequate unless ***some level of confidence*** can be expressed. One approach follows, with acknowledgement to the CSA Guide on Auditing (CAN3 - Q395-81) as well as Charles Mills [page 212]. It is also possible to use other sampling plans, bearing in mind that most were designed for high volume repetitive production processes, which Healthcare is certainly not!

If there are a large number of QA system items to be checked, some of which are critical to product/service conformances while the remainder are desirable to achieve good product/service performances, check all the critical items.

For example 7 clients on DDA, it could be feasible to check the records of all seven clients wrt dosage, application, reactions, and outcomes as well as informed consent issues.

For the non-critical items see Figure 1, select randomly a sample of “n” items where n is calculated from the formula

n = Log (1- C) / Log (1 - P)

Where **P** is an unacceptable high proportion e.g. > 5% non-complying clinical records expressed as a decimal (e.g. .05) and

C is the confidence (or probability) you wish to have detecting that level of non-conformance (expressed as a decimal). (e.g. 90%)

Clearly sample size "n" is going to impact on the cost of the audit through the time taken; the proportion defective "P" is also going to need to be low in these sectors where non complying procedures, records, test results can lead to patient recall and other remedial activities. [NOTE: It is important that n/N is < 10% where N is the population for that parameter]

For Health Care audit requirements, we would like a confidence level of 90%. You may argue that this should be 100% or at least 99.8%, try out the numbers from the table below.

If you like, you are 90% confident that you, or someone else as skilled and given the same information, could find the part of the quality management system in question and the same problems and come to the same conclusions [The latter comment assumes that "Corrective Action" had not occurred immediately!]. There is also the possibility that on 10% of occasions you could be wrong/ not find the problem.

Example 1

If when checking a large number of hospital work instructions you wish to be 90% confident you will pick up at least 1 non conformance if there are 5% or more non conforming work instructions, then you will need a sample size of

n = Log (1 - C) / Log (1 - P)

n = Log (1- 0.9) / Log (1 - 0.05) = Log 0.1 / Log 0.95

n = 44.89 = 45 Items

A sample of 45 will need to be checked. Still a high sample to simply show that work instructions are (i) up to date, (ii) approved and (iii) being used.

Clearly it is still important to compare similar populations when drawing the sample, for example in Quality Records in different departments we must clarify which *sub population* we are sampling from.

Example 2. By increasing confidence to 95% and proportion defective to 2% we have

n = Log (1 - 0.95) / Log (1 - 0.02) = Log (0.5) / Log (0.98)

n = 150 Now a totally unrealistic size of sample?

An alternative approach is to say that if - in a sample of size "n"- ***no non-conformance's are found***, we are C% confident that P% or fewer non-conformance's are in the system where n, C, P are given by the following table, P is the value in the table.

Sample Size	Confidence **C = 90%**
n	**P%**
100	2.3
50	4.5
30	7
20	11
15	14
10	21
8	25
6	32
4	44
2	72

Note: The figures given in the table are only approximate as they are intended as a guide. (Correction factors not applied in the calculations). They are also based on the assumption that the sample **(n) represents < 10% of the total number of items. (N) Note that the confidence level is 90%!**

However, given both of these criticisms they do provide some idea of the confidence associated with decisions made when no non-conformances are found. In particular the Table clearly shows that for a sample of 20 we are 90% confident, ***given no defects in the sample*** that the population could still be up to 11 % defective.

Of course without checking the whole "population" we do not know the underlying trend; it may be 5.9% "defective" but this is still too high for our needs, yet the sample did not detect it.

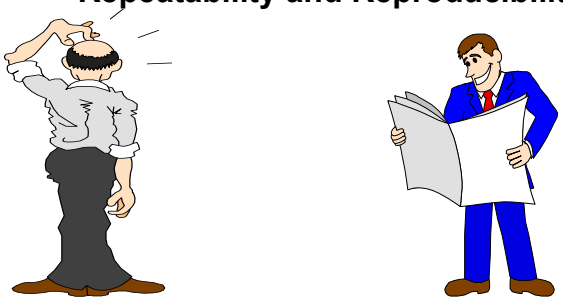
Example. 3

If 6 items are sampled and no non-conformances are found, we can report that we are 90% confident that fewer than 32% are non-conforming!! Now you can see why in the auditor should ask how many documents you have, how many instruments, and how many records in each category?

Note: Ideally these questions should be asked at the Counselling or Preliminary visit

So when auditors report any two non-conformances is it based on a random sample? Is it the only two errors they could find (and looked for)? Clearly if we only sample two features of the QMS (say Document Control) and find no errors we can only be 90% confident that there are less than 72% defects in the system! **There may be none**, but there could easily be 50% - a non-satisfactory result indeed!

Repeatability and Reproducibility?



These are two key points in the Audit process that need to be considered if the results of an Audit are to be seen as defensible in the legal sense. Can the Auditor use his/her notes to re-check the issues observed at the time [Repeatability] or can the Lead Auditor use the Auditor's notes to check that the information is correct and the appropriate decision made [Reproducibility]? The analogy with – for example – any measurement process and particularly say the "calibration" of pathologists against standard slides is a valid issue.

CONCLUSIONS

The authors conclude that the Audit can only have some valid significance in Healthcare settings if it is possible to set a level of confidence on the audit findings. The cost of the audit will be linked to the level of confidence based on the time taken to sample the various processes in the service(s) being audited. The level of confidence will be based on prior statistics [facts/evidence] and – in particular – how well the facility carries out and uses the results from its own internal audits and management reviews. A further conclusion is that the NZ Health and Disability Standards fit well into an ISO 9001:2000 framework for what is - in effect - a management system of best practice, provided we benchmark, measure customer satisfaction in a consistent way and proactively seek "business improvement" at all management level as well as with clinical/medical caring processes.

The reader should now be acutely aware that time for the audits are often too short so 'samples' are very small indeed! Hence the confidence we can honestly have on the results of many audits may typically be less than 60% due to the size of sample taken and the appropriateness of that population to be part of the overall risk package.

Confidence will generally be much higher in the **Single item Audit Trail** e.g. MRI Scanners [i.e. Trace forward or trace back technique], since the overall calibration confidence will be around 99.98%.

ACKNOWLEDGMENTS

To Gordon Smith, formerly of Massey University, New Zealand, to Malcolm Phipps of Canadian Standards (who first pointed out to me their standard CAN3 - Q395 - 81 on Quality Audits) and the many constructive comments of auditors from certification bodies around the world.

To my Auditing colleagues and co-authors in the Health sector, a sincere thank you for some considerable insight into often buried facts and figures, as well as highlighting the positive issues in an industry that is seemingly always stretched.

BIBLIOGRAPHY

1. Mil Std 105 E Sampling Procedures and Tables for Inspection by Attributes. Replaced by Mil Std 1916 April 1 1996

2. Charles A. Mills “The Quality Audit – A management Evaluation Tool” ASQC-McGraw Hill 1989

3. Lead Audit Training Course – International Management and Technology Ltd - IATCA Approved Chapter 11 “Confidence in Auditing”

4. SNZ HB 8134.2 2001 Health and Disability Standards (Hospital) Audit Workbook

5. HB 228:2001 Guidelines for managing risk in healthcare.

6. Mullock. Dr B. and A. Stephenson “Quality in Healthcare” IQA Conference Oxford 1985

NOTE: Copies of the Slides as PowerPoint may be obtained from

tony.stephenson@imat.co.nz