Defect Types and Surveillance Strategies for One-Shot Items

XXXX X. XXXXXXXX, XXXXXX XXXXXXXX XXXXXXXXXXXX
XXXX X. XXXXXXXXX, XXX, XXXXXX XXXXXXXX XXXXXXXXXXX

Key Words: one-shot items, surveillance, defect taxonomy

SUMMARY & CONCLUSIONS

This paper will describe some of the challenges and strategies for sampling and testing of complex one-shot systems. A taxonomy for defect types will be offered that informs the nature of the testing and analysis that should be done. In addition, some options for balancing and articulating risk will be described for the various surveillance programs described.

1 DEFINITION OF ONE-SHOT

This paper will describe some of the challenges and strategies for sampling and testing of one-shot systems. Key attributes of these systems are that they typically stay in dormant storage until called upon for one-time use. Common examples of one-shot devices are air-bags in vehicles, fire suppression systems, certain types of safety features in nuclear power plants, missiles, thermal batteries, and some stand-by systems. Some of the components of the system may be capable of multiple operations, but the fundamental usage is as a one-shot item. This paper will focus on complex one-shot systems which may be difficult or expensive to test.

2 DEFINITION OF SURVEILLANCE OBJECTIVES

Unlike continuously operating systems, one-shot systems typically do not reveal defectiveness until tested. This poses a major challenge for surveillance, especially if the testing is destructive. In many cases, one is limited to making inferences regarding the entire population from testing of samples.

This paper will focus in particular on systems that have long periods of dormant storage, potentially decades. One must thus manage one's assets recognizing that over such a time period, there can be a variety of defectiveness that could be manifested over the life of the system. The objective of surveillance is thus to look for unanticipated defectiveness through the use of sampling and testing.

The risk of not doing surveillance is that defects cannot be found and fixed and more importantly, one's presumptions about the reliability of the one-shot system may be flawed. A history of few defects in a one-shot system may mean that the inherent reliability is good but it may also mean that an inadequate test program has been conducted. It is critical that the risks of not looking are understood and communicated in addition to understanding the assumptions made when interpreting results of tests that have been conducted.

3 DEFECT TAXONOMY

One of the most fundamental assumptions one must make is whether defects are catastrophic quality defects or margin insufficiency defects.

In Figure 1, a quality defect is shown notionally. A histogram of some parameter is plotted for a population of units. The performance of each population member is either well above the performance requirement L or fails catastrophically (e.g., no output). Understanding the distribution to the right of L tells us nothing about the probability of units to the left of L.

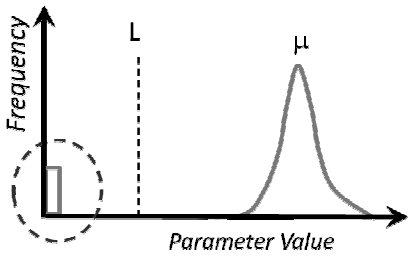

Figure 1: Quality Defect

This is contrasted with

Figure ***2***, a margin insufficiency defect. Here, it is presumed that knowledge of the population distribution characteristics allows one to make inferences about the probability of units to the left of L. Clearly one will plan surveillance differently depending upon which view one holds regarding the nature of the defects that are present.

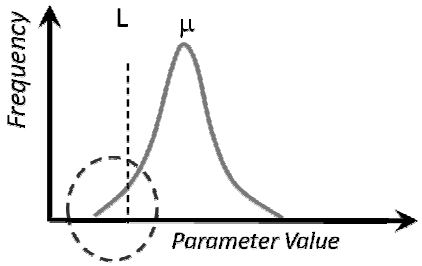

Figure 2: Margin Insufficiency Defect

For certain classes of complex one-shot systems, the historical experience has been that quality defects overwhelmingly predominate. This is not surprising when

considering the design process. Significant resources and effort are invested during design to ensure that margin to the requirement is large, and conservative requirements and product acceptance specifications are defined to ensure this. In short, rigorous design, development, and qualification ensure that few cases of margin insufficiency are present when a system is newly fielded.

Note that one could argue that many problems that are first detected as catastrophic defects could have been tracked as margin insufficiency, if we just knew the right parameter to examine. Thus this characterization of defects as either quality or margin insufficiency hinges to a large extent upon the nature of the evaluations that we do. However from a practical standpoint, given that complex one-shot systems may include numerous individual components, each with a wide variety of materials and functionality, it is clearly not possible to explore every single parameter that might reveal margin insufficiency.

The question then arises: what happens to systems as they age? How do we most effectively "look for change"? One model of age-related defects is shown notionally in Figure 3. While the margin of the parameter value (e.g., the amount of active ingredient) may have been adequate at time of production (t_0), it may be slowly changing over time such that at some point in time (t_2) there is a defective fraction due to margin insufficiency.

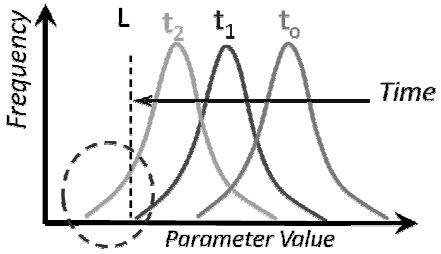

Figure 3: Margin Loss

Thus even though static quality defects might still be present, at some point in time the greater risk for the one-shot system becomes age-related defects. From a surveillance perspective, the concern shifts from detection of quality defects to emerging margin insufficiency defects ("margin loss"). Along with this comes increased emphasis on variables data analysis in order to efficiently detect margin loss as depicted in the graphic below.

It should be noted that there is a second aging model in addition to margin loss. This is noted in Figure 4 as a quality defect later in life – these are typically referred to as "latent quality defects". Latent quality defects are catastrophic defects that appear with time.

These, like quality defects early in life, are not predictable from knowledge of the distribution that lies to the right of L. Instead one generally relies on regular functional testing of a large number of units to characterize their prevalence. A good example is stress voiding, where the circuit path continues to conduct current as stress voiding is occurring until that critical moment when the voids progress to the point that the circuit path is opened. To outward appearances, the performance "drops off the cliff" with no parameter signaling prior degradation.

Note that latent quality defects may eventually affect the entire population (as stress voiding does) or only subpopulations (an example is corrosion that only occurs in units that were contaminated during production). Thus while it may be possible to predict some latent defects as margin loss, if one has access to the correct parameter, this will be impossible to do in the case of unknown problems that affect unknown fractions of the population. The difficulty is compounded by the huge scope of materials and potential degradation mechanisms present in complex one-shot systems. Given this, latent quality defects are generally the most common way that age-related issues are observed and assessed, rather than margin loss. As with early-in-life quality defects, latent quality defects cannot be characterized by the main population distribution; instead, one must rely on sampling data to estimate the prevalence of a defective subpopulation.

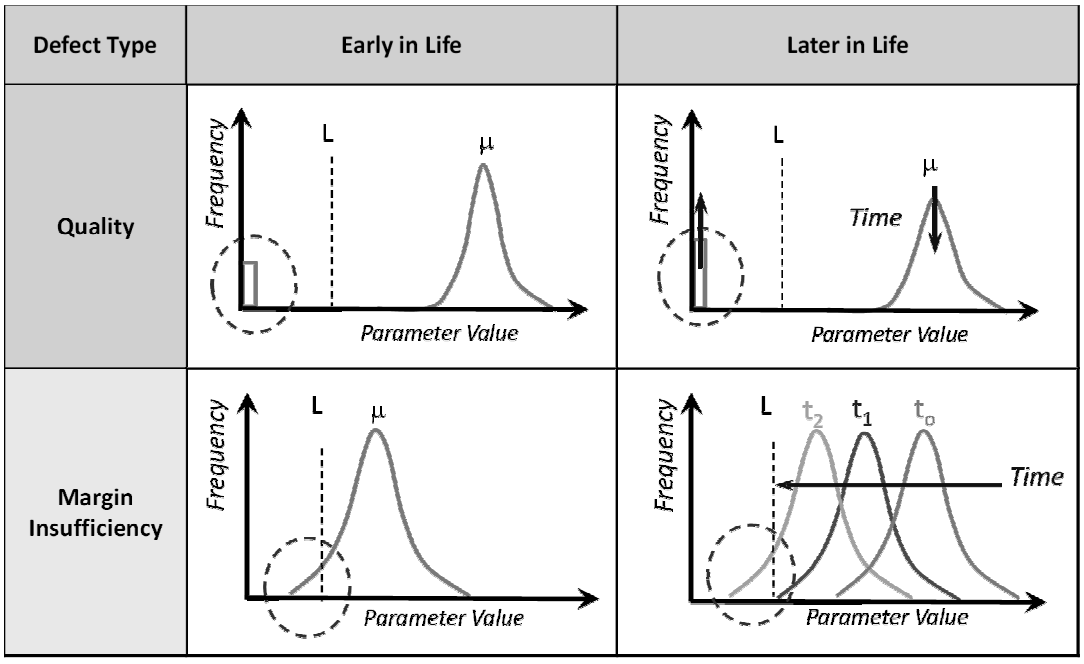

Figure 4: Defect Taxonomy for One-Shot Systems

4 SELECTION OF TESTS AND PARAMETERS

One of the strengths of variables data analysis is that it offers the opportunity to detect stockpile change through the monitoring of various parameters. However this is a double-edged sword. If the wrong parameters are monitored, one may be misled into believing that change is not occurring when it actually is. There is a major difference in the nature of inferences that can be made from functional tests compared to variables data analysis on a handful of parameters – each has its strengths and weaknesses. And certainly different parameter selections give us different types of information.

There is a clear paradox in selection of parameters to be analyzed. Parameters measured at higher levels of assembly often tend to be easier to relate to overall performance and also reveal a broader range of failure mechanisms than parameters at lower levels of assembly. This is particularly true if realistic functional tests can be conducted. However if something is changing, that change may not be manifested in measurements taken at a higher level of assembly but could be detected by monitoring parameters at lower levels of assembly.

This underscores the interesting dichotomy noted earlier. The very same issue can appear to be either a latent quality defect or a margin loss defect, depending upon the particular parameter one is monitoring. The dilemma is that in order to detect margin loss, one must (1) have access to the relevant parameter in order to measure it, (2) be smart enough to choose that parameter for monitoring, and (3) be careful enough in making the measurements over time such that a change can be differentiated from changes in testers or test procedures.

5 SURVEILLANCE STRATEGIES FOR DEFECT TYPES

One's assumptions about the nature of the defectiveness and types of evaluations available thus have important implications for sampling and data analysis for complex one-shot systems. Methods to characterize the distribution of a parameter are an excellent tool to identify and understand margin loss, but they cannot answer questions regarding latent quality defects. In order to estimate reliability, it is not sufficient to collect data to characterize the population distribution to the right of L. Instead, one must sample to ensure that the quality defects (as shown on the left of L) can be detected. In the absence of identified parameters, one must rely on functional testing to make this judgment. Defect discoveries from functional testing often lay the foundation for understanding the parameters that describe physical degradation in these complex one-shot systems.

It is thus critical to keep both the strengths and limitations of variables data in mind when considering how well we understand complex one-shot systems. A similar comment can be made for sampling strategies: they must take into account the possibility that any of these defects may be present in the population with sample sizes chosen appropriately to enable detection with reasonable level of risk.

In this section, some potential strategies for sampling and analysis for each of these defect types will be described.

5.1 Quality Defects

Looking back at Figure 1, recall that analysis of the variables data to the right of L does not help understand the probability of catastrophic failure. The most reasonable approach is to avoid making a distributional assumption and treat the functional evaluations as Bernoulli trials. One can size the sampling program by deciding the level of risk one is willing to live with (cast in terms of "what is the size of the defect that may be undetected?" along with a confidence level that the defect will be detected) and using the hypergeometric distribution or binomial approximation to calculate number of samples. Of course there are many important assumptions that must be kept in mind, foremost being that this presumes that one is able to detect all defects present if one is lucky enough to sample a defective unit. This requires careful test design for complex one-shot systems, particularly if the tests are destructive and there is only one chance to exercise the hardware.

Using the hypergeometric distribution, one can calculate the confidence, γ, that at least one defective unit will be found in a random sample from a fixed population as:

$$\gamma = 1 - \frac{\binom{NP}{0}\binom{N-NP}{n}}{\binom{N}{n}} \tag{1}$$

where:

n is the number of samples,
P is the defective fraction, and
N is the population size.

If no defective units are found in the sample, γ can be interpreted as the confidence that the defective fraction is no more than P. This uses the reasoning that if more than NP defective units had been present, at least one of them would have been sampled with probability, γ. Since the defective fraction is unknown, the calculation is interpreted in terms of both the confidence and the defective fraction that is precluded. Thus for example, a sample of 25 out of a population of 500, with no defects found, would give 93% confidence that the defect fraction is no more than 10%, but only 70% confidence that the defect fraction is no more than 5%, and 23% confidence that the defect fraction is no more than 1%. Similar calculations can be done to examine the risks associated with other choices of sample size. Figure 5 shows example calculations for sample sizes of 10, 25, and 50

out of a population of 500 units. Note that 100% sampling is required to have 100% confidence in zero defects.

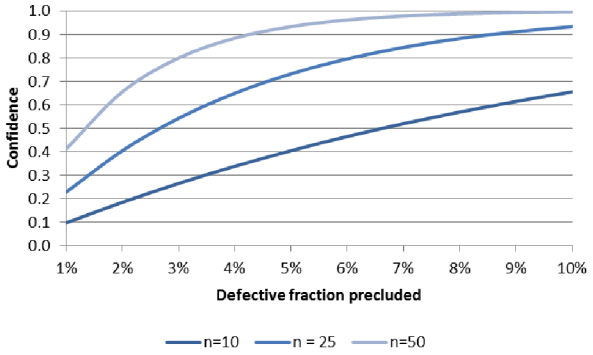

Figure 5: Confidence vs. Defective Fraction Precluded

5.2 Margin Insufficiency Defects

A tolerance limit approach provides a good basis for thinking about sampling quantities for margin insufficiency [1],[2]. From a theoretical standpoint, the number of samples required for determining margin depends upon the desired population coverage (the proportion of the distribution above the performance requirement that one wishes to demonstrate) and the desired confidence level. It also depends upon the inherent margin of the component – a component with more margin will require fewer tests compared to a component with less margin to show with the given level of confidence that the given coverage proportion has been achieved. However this inherent margin is not known a priori and is the performance property to be estimated from the data. Hence although critical, it cannot enter into determination of sample size until at least some data are taken.

Another key factor that affects one's estimate of margin, also unknown, is the distribution of the population. Because the distributional assumption is pivotal in estimating margin, validation of this assumption may be the driver for setting a minimum sample size. In cases of small margin, many samples may be needed to achieve the desired confidence.

Figure 6 shows a set of operating characteristic (OC) curves for the probability of demonstrating a specified margin as a function of true margin. Margin can be defined by a z-score, which is equal to the number of standard deviations from the population mean to the value that defines the defective fraction. If the distribution is normal, the defective fraction, P, is a known function of the z-score, z_P. For example, only about 2% of a normal distribution falls more than two standard deviations below the mean, so a z-score of two translates into a defective fraction of 2%. The confidence limit is equal to one minus the risk of falsely concluding that the z-score is greater than two. Calculations such as those shown on Figure 6 illustrate the risk of failing to demonstrate an adequate margin. The risk is reduced by increasing the number of samples, but the number of samples that may be required to have a high probability of demonstrating margin becomes very large as the true margin approaches the margin to be demonstrated. When true margin is equal to margin to be demonstrated, the probability of demonstration is limited to 10% by the nature of the tolerance limit calculation.

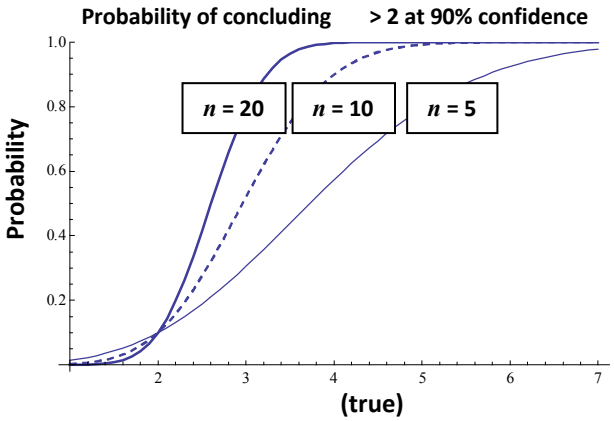

Figure 6: OC Curves for Demonstrating Margin

Comparison of Figures 5 and 6 shows the relative efficiency that may be achieved by monitoring variables data. For example, from Figure 5 one sees that a sample of 50 is needed to demonstrate a defect fraction of less than two percent at the 60% confidence level. In contrast, from Figure 6 one sees that one may be able to demonstrate a margin defect of less than two percent (*i.e.*, z-score greater than two) at the 90% confidence level with as few as 5 samples, if the true z-score is six or more. However, as noted above, the margin demonstration does not rule out the possibility of quality defects.

5.3 Latent Quality Defects

This case is very similar to that described earlier for quality defects. The difference is that since latent defects emerge over time, inclusion of early tests in the calculation may result in overly optimistic interpretation of the test history. Including only recent tests will give a much better picture of the ability of a given sampling program to identify latent quality defects. The time window to choose depends upon the components in the system, but it should be noted that some problems do have rapid times of onset and thus regular sampling is always prudent, particularly if the consequences of failure are high.

5.4 Margin Loss Defects

One can extend the tolerance limit calculational approach to include regression in order to consider sampling needs to detect margin loss over time. Like the margin insufficiency case described earlier, there are numerous assumptions that must be made, and effort should be expended to confirm those assumptions. In particular, great care should be taken in the case of limited data to ensure that the caveats surrounding performance conclusions are well communicated.

6 BALANCING RISK

The dilemma of course is that one typically does not know *a priori* the nature of the defects that may be present

when a system is produced or even more, that may emerge over time.

As can be seen from the above examples, monitoring variables data can be more efficient at identifying change in performance compared to attributes data but there is an important presumption that (1) there is a variable that is associated with performance and which manifests any age-dependent performance degradation that may occur and (2) that variable is being monitored during testing. Defects often don't meet both of these conditions, so functional testing must be the primary vehicle to detect them.

The point to underscore is that variables data analysis is not sufficient by itself to estimate reliability, because it does not address quality defects. In addition, one should not be lulled into complacency (and low sample rates) by high margin; there may be quality defects present that would then not be detectable.

One must thus craft a program that recognizes both types of risk and attempts to balance them. Even if one presumes excellent margin (and a concomitant small sample quantity), one must take into account the possibility of quality defects (either early in life or latent) when sizing one's sampling program. The planned program will tend to be a maximum of these two, in cases where functional test units can yield to subsequent margin testing.

7 SUMMARY

There are different classes of defectiveness that may be present in complex one-shot systems. Unlike continuously operating systems, defects will not be detected unless testing is done explicitly. One's ability to find defects (and consequent risk of undetected defects) thus depends upon sampling and testing. Different assumptions about the nature of the defects that are present or that may emerge over time can lead to very different sampling strategies. It is thus wise to consider each potential defect type separately and then select a program that gives an acceptable level of risk integrated across these defect types.

REFERENCES

1. A. Kiermeier et al. A new approach to estimating shelf-life. Pharmaceutical Statistics 3: 3-11, 2004.
2. D.B. Owen. A survey of properties and applications of the noncentral t-distribution. Technometrics 10: 445-478, 1968.
3. Kramer et al. Robust Surveillance Plans for Monitoring Shelf Life Systems with Degradation. Quality Engineering 11(2): 306-319, 1998-99.

BIOGRAPHIES

Name1
Address 1
Address 2
Address 3
City, State or Province, Postal Code, Country

e-mail:

Text.

Name2
Address 1
Address 2
Address 3
City, State or Province, Postal Code, Country

e-mail:

Text.