[redacted]

What *Is* My Failure Rate?

by GM Samaras

Introduction

The quality of a product not only includes fitness for intended use, but also the reliability of that product attribute. In the August 2015 issue, we discussed medical device life-cycle risk management with emphasis on the similarities of pre- and post-market risk management and discriminating system use vs. human user errors[1]. In both pre- and post-market risk management we have the opportunity to acquire failure event data (so-called "numerator" data). These are values of observed events that may or may not be statistically valid or scientifically useful depending upon how, when, where, and from whom we acquired the data. In the pre-market setting, limited testing results in limited cost, but also limited data; in the post-market setting, uncontrolled "natural experiments" yield real-world results, but in a manner that makes the data difficult to analyze and generalize (i.e., to justify external validity). This is often used to minimize the utility and value of the FDA-required complaint-handling[2] process, because of the perceived quality of the data. This is both technically incorrect and wasteful of corporate resources. Post-market data have enormous value for sentinel event (safety signal) recognition and health hazard evaluation, but only if they are properly analyzed and the failure rate is properly computed. Ignoring or mishandling these data undermines competiveness, profitability, and perceived corporate excellence; it also violates federal regulations in the United States and exposes the firm to product liability.

Failure Rate Computation

Whether you are tasked with investigating reliability or survivability, what you are fundamentally concerned with is (repairable or nonrepairable) failure rate. The rate can be based on time, product units in use, product uses, or some other denominator. Very often, those unskilled in reliability engineering will choose the largest possible denominator, so as to make the failures look good with a small rate value. From an engineering perspective, this is both nonsensical and counter-productive, as it masks the triggers necessary to justify making changes to design, manufacturing, or distribution of the product. The central objective of reliability engineering is to identify product defects, so that they can be corrected; obfuscating those defects protects nothing, except possibly sales commissions. Obfuscating defects will cost the enterprise far more in the long-term. To help identify product defects, you need to choose the correct denominator.

In the pre-market phase, most medical device manufacturers seem to rely on subjective estimates of probability of occurrence, rather than quantitative failure rate data, for their construction of a design (or device) failure mode and effects analysis (DFMEA). Furthermore, they typically do the analysis with a single unit in mind, rather than the aggregate of units that are expected to be sold.[3] As you transition to the post-market phase, expected frequency of occurrence of an event is typically derived from your failure mode and effects analysis (FMEA) and incorporated into your health hazard analysis (HHA). But,

1. Samaras, GM. Medical Device Life Cycle Risk Management. *BioFeedback*, 43(2): 15-19, 2015.
2. 21 Code of Federal Regulations 820.198: Complaint Files.
3. If you estimate your individual product will fail only once every million uses (10^{-6}; "improbable" per ISO 14971), but you plan on selling at least 1,000 units, then your probability estimate has to be somewhere close to once every thousand uses (10^{-3}; "frequent" per ISO 14971, assuming independence, same epochs, etc.).

cont. on p. 6

post-market medical device risk management[4] requires you to update your pre-market predictions with post-market data, which then changes the results of your DFMEA and your residual risk analysis. To do this correctly, you need a robust complaint management system that actively seeks product failure information and correctly determines the root cause.

We can think in terms of two types of medical device product failures: (a) those that interfere with user workflow, requiring the product to be serviced or replaced, and (b) those that result in reportable[5] adverse events. The former are typically noncritical, but potentially costly defects related to design, manufacturing, or distribution; the latter may be quite costly and have protracted effects on competitiveness and profitability. How you discriminate these and apportion resources for corrective or prevention action, in part, depends on your calculation of the failure rate. Your choice of denominator is an indicator of your expectation (or implicit assumption) regarding the underlying root cause. It also drives the occurrence value in your updated DFMEA and your HHA. In both we are focused on likelihood of harm given exposure to the hazardous situation—not on likelihood of product failure.

Consider the following scenario. You have a medical device with an accessory and a disposable; the device has no utility without the accessory and the accessory cannot function safely without the disposable. This configuration can be found throughout many medical device classifications and in multiple products outside the medical device domain, driven in part by contemporary marketing strategies. Now consider multiple complaints of reportable adverse events that were potentially caused by, or contributed to, reported product "failures." And, of course, your job is to figure out whether the most recent complaint is a sentinel event (safety signal) or "just" human error; your enlightened members of management are relying on you to tell them what is actually occurring, so they can take appropriate action and allocate appropriate resources. The complaints group gives you its information (information about the person injured, the injury, and some information about your product that was being used at the time).

Table 1 – Denominator Examples and Putative Root Causes

Example Denominator	Implicit Assumption of Root Cause
Number of units	Design, design transfer, or manufacturing
Number of users	Instructions for use, or training and credentialing
Number of uses	Wear, cleaning, or maintenance degradation

Table 1 offers some example denominators and implicit assumptions of the root cause of the failure. You could choose the number of uses (no. uses) as the denominator, get a really low failure rate, and get on with more important things, since it is obviously human error; but, what if someone asks you to prove that wear, cleaning, or maintenance degradation is not the root cause and the problem is simply a design defect resulting from your choice of materials or the manufacturing process? You could choose the number of users (no. users), which is a smaller denominator than number of uses (no. uses). Sure, you get a larger failure rate, but you still can blame it on human error; but, what if someone asks you why the root cause is not a defective IFU[6] or inadequate training protocol[7] and, oh by the way, can we see your labeling comprehension and usability validation study final

4. ISO 14971:2007 §9.
5. 21 CFR 803.3: Death or serious injury (life-threatening **OR** results in irreversible harm to body structure/function **OR** reversible harm to body structure/function that requires surgical/medical intervention to prevent it becoming irreversible).
6. Instructions for Use, part of FDA required medical device labeling and subject to human factors validations since at least as early as 1997.
7. It could be that the user is violating the labeling (resulting in failures indicative of deficient training) or the user is following the IFU and still resulting in failures (indicative of deficient labeling).

cont. on p. 7

report for the IFU and/or the training? You could conclude that the denominator should be the number of units (no. units), because you suspect a—as yet to be uncovered—defect in design, production, or maybe even distribution. But, then you realize that is going to be a serious headache; R&D, manufacturing, and marketing are all going to be very skeptical. Furthermore, are we talking about the disposable, the accessory, or the base device? Or, is it the interface between the disposable and the accessory or is it the interface between the accessory and the base device?

You obviously need more information … much more information, but the complaints group gave you everything they have and claim that is everything it could get.

Conclusion

Incorrect failure rate estimation in the pre-market phase is the result of defective risk management and a failure of design control. However, in the post-market phase, incorrect failure rate estimation typically arises from deficient complaint handling and/or choice of an inappropriate denominator. This is, in effect, censoring and can severely limit data validity. Notwithstanding the famous work of Kaplan-Meier [8], ignoring or mishandling complaints of product failures is an effective censoring mechanism that denies corporate management the information necessary to make correct business decisions. Defective failure rate estimation undermines competiveness, profitability, and perceived corporate excellence; it also unnecessarily exposes firms to failed inspections, unwelcomed audits, and product liability. One approach to ameliorating this is allocating more resources to the complaint-handling process and insisting on greater rigor in the complaint investigation and data collection. You cannot manage what you cannot control, and you cannot exert effective control without necessary and sufficient information.

8. Kaplan, EL and Meier, P. (1958). Nonparametric estimation from incomplete observations. J. Amer. Statist. Assn. 53 (282): 457–481.