

Separation of Compliance and Performance Under FDASIA

Jason J. Orloff

The root cause of drug shortages is mismanagement of variation.

The seeds of the drug shortage crisis were sown when the bio/pharmaceutical industry as a whole failed to embrace Statistical Quality Control in the 1978 GMP revision, documented in the preamble to *Code of Federal Regulations (CFR) 21*. However, when congress amended the Food, Drug, and Cosmetic Act in July of 2012 it conferred new authorities to FDA for both safety and innovation. In light of the agency's request for comments on preventing drug shortages, there is a rare opportunity to correct this 30-year spiral. With the 2011 revision to the process validation guidance and 2012's Food and Drug Administration Safety and Innovation Act (FDASIA) Title VII (705/706), industry and government are uniquely poised to correct and prevent practices that undermine process improvement, decrease supply, and make manufacturing more difficult with no benefit to patients. Key to a "maximally efficient, agile, and flexible industry" could be a single meaningful metric to focus attention on process variation and separate regulatory oversight into distinct departments for compliance and performance. This metric is the out-of-specification (OOS) rate.

From change inertia to enabling improvement

All too often specifications are set by limited experience at the time of filing. Then, when a process is improved and variability reduced, the specifications are tightened to maintain this positive change. Rarely are specifications therapeutically and toxicologically relevant

Performance against specs is the cost by which product becomes too expensive to manufacture.

boundaries to patient risk. Instead industry has what a senior FDA representative described as "a risk envelope through which drugs are delivered to the market and government protects the public health." Until International Conference on Harmonization (ICH) Q9 in 2005, there was no objective criteria preventing this envelope from tightening because as proxy for the patient, the world health authority's safest risk was "Zero Risk" and the mechanism to control variability was ever shrinking specifications. Rather than a return on investment, improvement became a disincentive.

FDA's approach to FDASIA 705 is remarkable in that the agency recognizes drug shortages as harmful to patients and fundamentally tied to industry's manufacturing capability. Poor performance is now a risk to be factored into the inspection schedule and an opportunity to check and balance a system of pure compliance. Unlike approaches that used the recall rate as evidence lagging a complete systems failure, today FDA is asking for leading indicators. Modeling such risks means weighing potentials instead of simply acting upon evidence existing in fact. It also means that inspection must go beyond mere compliance and address such questions as: How could this be better? Why is this important? What is enough? To answer such questions requires a different set of skills from compliance as detailed by Judith Malsbury (1). With the data collected under FDASIA 706 in lieu or in advance of an inspection, it is possible to craft a scheduling system where both compliance and performance matter.

Process capability as self audit

Consider one hypothetical critical quality attribute (CQA) for one product in one plant. Imagine that lots are occasionally OOS with respect to the lower regulatory specification limit (LSL) and upper regulatory specification limit (USL). Now instead of treating these as isolated events that either comply or not, consider the last 200

Jason J. Orloff is statistical and engineering consultant at PharmStat, jorloff@validationbydesign.com.

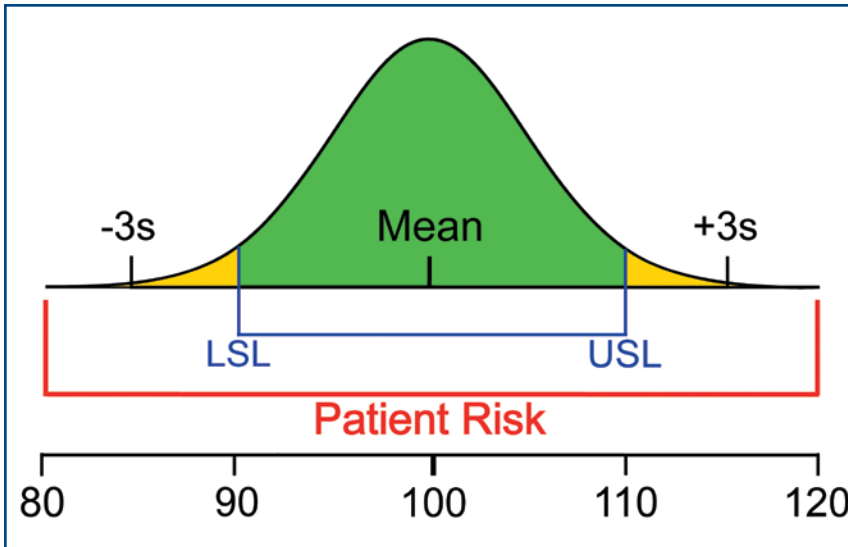


Figure 1: A model of critical quality attributes performance against specifications, assuming normality. LSL is lower specification limit. USL is upper specification limit.

Table I: Simulated plants representing quality system maturity, manufacturing volumes, and critical quality attributes complexity and capability. OOS rate is number out of specification on lot release/total number of tests run in a year.

| | #OOS, Average | Number of Tests, Average | OOS Rate, Avg. Per Year |
|---------|---------------|--------------------------|-------------------------|
| Plant-1 | 153 | 12649 | 1.21% |
| Plant-2 | 3 | 193 | 1.55% |
| Plant-3 | 96 | 12463 | 0.77% |
| Plant-4 | 1 | 131 | 0.76% |
| Plant-5 | 3 | 1566 | 0.19% |
| Plant-6 | 0 | 75948 | 0.00% |
| Plant-7 | 1 | 752 | 0.13% |
| Plant-8 | 1 | 1474 | 0.07% |
| Plant-9 | 1 | 71 | 1.41% |

ALL FIGURES ARE COURTESY OF THE AUTHOR

lots of release data as a model of the whole process (2). Shown in yellow in Figure 1, the previously isolated OOS events suddenly take shape. Instead of the exception, the process is regularly generating OOS material. OOS must be expected. In addition, the variability inherent to the process—here depicted by the +/- 3 standard deviations (s) rule of thumb—compared to the range allowed by the specifications determines a consistent average rate of OOS. More variability means less capability and a higher OOS rate.

Proposed by both the Parenteral Drug Association (PDA) and the International Society of Pharmaceutical

Engineers (ISPE) as a leading indicator preventing drug shortages, the OOS rate has some particularly useful properties (3, 4). Foremost, the definition of what is critical—the lot release specification—has already been negotiated, so there are relatively consistent definitions. These critical regulatory specifications are the boundary protecting patients from hazards of high severity. Combined with the class of pharmaceutical product or device, they rank risk to the patient. Consistent with ICH Q9 where Risk = Severity * Probability * Detectability, their OOS rate is a measurement of the probability of hazard for the patient. This rate ranks the effec-

tiveness of the quality system's detection of such hazards and the probability/capability of the process to manufacture product within specifications.

Capability impacts the cost of compliance but also the cost of manufacturing and risk to the patient. Accordingly, the 2011 process validation guidance has correctly set the emphasis on variation (5). Walter Shewhart's genius—as the father of statistical quality control—was recognizing that the trend expected is predicted by past performance and the variation threshold to detect a change is a balance point between the manufacturer's and patient's risk (6). Although there are many ways to set "trend" thresholds, the most familiar are +/-3σ, +/-4.5σ, or tolerance intervals with additional rules to check non-random looking patterns.

When a result falls outside the classical +/-3σ, we are roughly saying it looks different from 99.73% of the historical measurements—the process may have changed. The purpose is to recognize and counter such a trend before the process goes out of control. To put it simply, where "out of specification" limits guard the patient, "out of trend" (OOT) limits guard the company. OOS limits are the range allowed. OOT limits could be the variability expected in the process. The capability of a process for a given critical quality attribute is then:

Critical Quality Attribute Capability

$$= \frac{\text{Variability Allowed by Specs}}{\text{Variability of the Process}} = \frac{\text{OOS Range}}{\text{OOT Range}} \quad (\text{Eq. 1})$$

Where OOS guards against risk to the patient, OOT guards against OOS guarding against risk to the patient. Business needs space between OOS and OOT to operate; this is the key to performance and ensuring supply. Where OOS is evidence submitted in field action reports, OOT and capability are a form of self-audit. Their purpose is to identify potential problems and act before harm reaches a patient. Only OOS is reportable lest OOT become the new OOS. To report both creates a legal double jeopardy, like being tried for the same crime twice. However, a perfor-

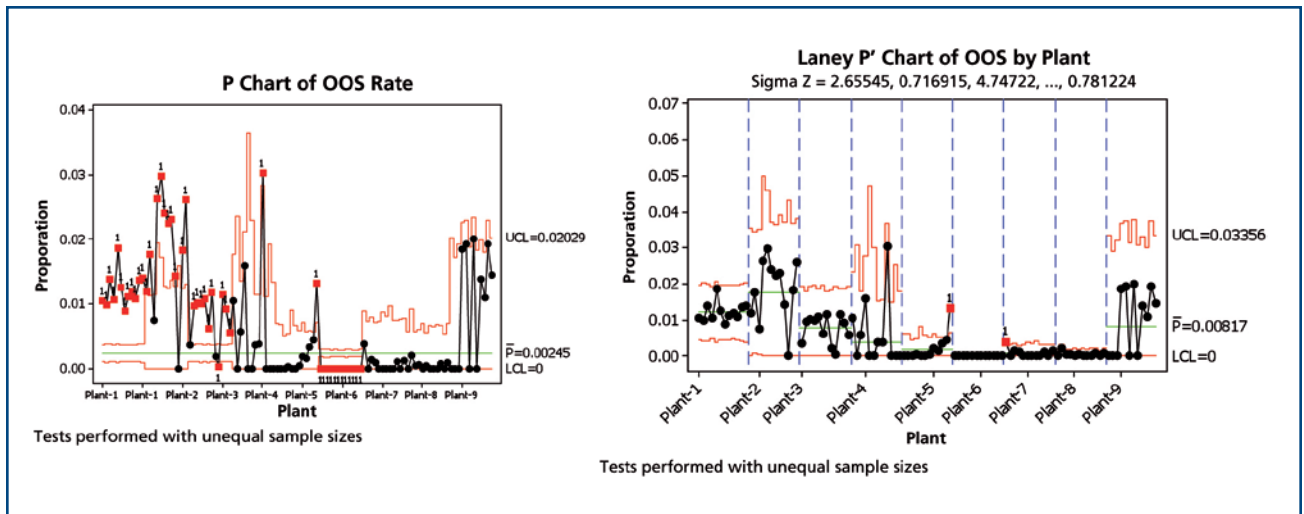


Figure 2: Data from nine simulated plants is shown in two different control charts. On the left, industry appears to be out of control from operation size and critical quality attributes complexity alone. On the right, standardized Z-scores reveal signals from the noise.

mance audit of rejecting lots on OOT would be evidence of a firm’s lack of understanding of statistical control and its proactive nature.

The role of OOS rate

After almost a decade of quality-by-design (QbD) review, there is now a group of specialists in FDA capable of auditing performance. Credit goes to Janet Woodcock’s CDER, which now has all the elements necessary to enable an agile pharmaceutical industry for the 21st century. Yet, to investigate a metric of performance such as capability directly without the goal of deeper understanding and the requisite expertise to drive improvement would be a mistake. Better instead would be to standardize with industry on the related metric of OOS rate to drive performance audits and schedule compliance inspections. The number of OOS could be divided by the number of lots in distribution reports to calculate this rate (7). Or perhaps the OOS and total number of tests run in a year could be requested under FDASIA 706, providing insight into both the state of global quality and which products or plants stand out for further scrutiny.

The key to implementing the OOS rate as a monitoring tool is to recognize and account for the radical differences between plants and products. Some products have only half a dozen specifications while others have 50 or

more. Some plants manufacture less than a hundred lots a year while others make tens of thousands. **Table 1** has nine simulated plants meant to represent the breadth of quality system maturity, manufacturing volumes, critical quality attributes (CQA) complexity and capability.

Unlike the recall rate where the expected state of catastrophic failure is zero, the OOS rate has neither the simplicity nor the pitfalls of a perfect state. Rather, the following must be asked:

- Is the absence of OOS in Plant-6 deserving of audit relief or is it too good to be true? (i.e., an easy two-sided control limit test that would discourage falsification of OOS reporting).
- Are the 153 OOS results in Plant 1 out of control or the result of using a process analytical technology (PAT) system and indicative of effective detection?
- Or worse yet for drug shortages, are Plant-1, 2, 3, & 9’s specifications overly tight and in need of review per ICH Q6A, section 2.5 that “... could involve loosening, as well as tightening, acceptance criteria as appropriate” (8)?

Certainly, direct comparison of such a table is inadequate. Just as industry must go beyond simple averages, so must government in an attempt to understand the system as a whole. The common tool is the control chart. **Fig-**

ure 2 shows two graphs of 12 reporting periods for each of the nine plants. The graph on the left charts the OOS rate as a proportion of the number of OOS results over the total number of tests run in a year. The proportions (P) chart puts everything on the same scale without accounting for the differences between plants. The industry appears out of control with red dots everywhere, but the smaller manufacturers are being penalized by the math of volume alone. In manufacturing, as in life, stuff happens, and within this industry subgroup it appears to happen at least 0.245% of the time regardless of complexity or size.

The chart on the right of **Figure 2** shows a comparison of within-group variation and between-group variation to see what stands out. One approach to this is the Laney P-prime (P’) chart where we see both the current level of quality in the industry and how the nine plants compare. What’s more, it is possible to see without the distraction of noise:

- That a problem has been developing over time in Plant-5.
- Plant-7, on the other hand, may have just been a false alarm or they proactively got a problem under control.

Without going into mathematical detail, plants might be directly ranked by a standardized Sigma Z score or reviewed for non-random changes over time with relevant rules. Alternatively,

industry might be segregated by class and volume using survey statistics so that a performance auditor is blinded to the individual plant risk score; concerned instead with understanding what is happening to inform both the plant and health authorities of performance.

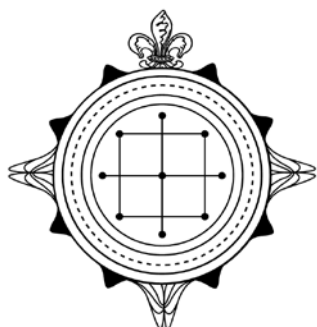
Such details have yet to be worked out. But using standardizing techniques such as Laney's P', every metric on every senior management quality scorecard could become a test case to compare product lines and plants for the purpose of prioritizing variability reduction and pro-active process improvement. Now is the time for us to begin such self-audit and action. A shift to managing variation by performance

Now is the time for us to begin such self-audit and action.

against specifications returns the investment on improving process capability. Should government adopt senior management's control charting to monitor plant and product performance, then we might yet develop a common language and rightfully devote ourselves to the common cause of reducing variation to improve quality, minimize costs, maximize profits, and reduce risk to the patient. Everybody wins.

References

1. J. Malsbury, "Performance Based and Compliance Based Auditing"(ASQ, 1995).
2. ASTM, *E2281-08a Standard Practice for Process and Measurement Capability Indices*.
3. ISPE, "Proposals for FDA Quality Metrics Program—Whitepaper" (Dec. 20, 2013).
4. PDA, *Points to Consider: Pharmaceutical Quality Metrics* (PDA, 2013).
5. FDA, *Guidance for Industry, Process Validation: General Principles and Practices* (Rockville, Md, January 2011).
6. W.A. Shewhart and W. Edwards Deming, *Statistical Method from the Viewpoint of Quality Control*, Dover Edition (1986).
7. CFR 21 Section 600.81.
8. ICH Q6A, *Specifications: Test Procedures And Acceptance Criteria for New Drug Substances And New Drug Products: Chemical Substances* (ICH, Oct. 6, 1999). **PT**



PHARMSTAT