Xavier University/PwC
Pharmaceutical Quality Metrics
White Paper

Xavier University
3800 Victory Parkway
Cincinnati, Ohio 45207

PricewaterhouseCoopers Advisory Services LLC
300 Madison Avenue
New York, NY 10017

January 31, 2016

Table of Contents

Executive Summary

The Food and Drug Administration Safety and Innovation Act (FDASIA)[1] of 2012 gave FDA authority to request data and information from the industry in advance of or in lieu of an inspection to identify potential risk for drug supply disruption, improve the efficiency and effectiveness of establishment inspections, and improve FDA's evaluation of drug manufacturing and control operations. This authority allows FDA to develop a process for resource allocation based on operations of greatest risk. As a result, FDA announced its Quality Metrics Initiative in February 2013 to ascertain data the industry could submit to FDA that would provide an indication of risk to product quality. FDA worked with the industry throughout 2013 – 2015 to identify metrics it could request from drug firms under its authority and issued its proposal in the 2015 "Request for Metrics" draft guidance.[2] Reviewing company-specific data out of context, however, could lead to false conclusions. In contrast, reviewing this data during an inspection could provide critical contextual information as it relates to the company itself, facilities, products, and importantly, risk to patients.

In support of FDA's intent to allocate its resources based on risk, Xavier University and PwC launched a Metrics Initiative in August 2014 to identify product quality risk metrics linked to patient safety that could be viewed during an inspection. Xavier University and PwC led a team of 30 industry professionals that developed a framework of 11 metrics across the Total Product Life Cycle (TPLC). The proposed metrics framework was designed to help offer a tool that stakeholders across the industry could use to inform decisions and trigger action. It is built upon driving a mindset of continual improvement that includes feedback loops across the entire enterprise to design quality into the product proactively at the source, instead of reactively catching inadequate quality after manufacture. The team recommends that this framework of metrics be incorporated into FDA's inspection protocol as a roadmap for investigators to evaluate drug manufacturing and control operations during an inspection.

[1] FDA Safety and Innovation Act: http://www.gpo.gov/fdsys/pkg/BILLS-112s3187enr/pdf/BILLS-112s3187enr.pdf

[2] The FDA "Request for Metrics" draft guidance and the intent of its use are described on the following FDA Voice Blog page: http://blogs.fda.gov/fdavoice/index.php/2015/07/quality-metrics-fdas-plan-for-a-key-set-of-measurements-to-help-ensure-manufacturers-are-producing-quality-medications/

Background

On July 9, 2012, the Food and Drug Administration Safety and Innovation Act (FDASIA) was signed into law, expanding FDA's ability to safeguard and advance public health.[3] The Act provides FDA with the ability to collect data and information from pharmaceutical companies prior to or in lieu of an inspection (FDASIA Title VII, Sections 704, 705 and 706). In February 2013, FDA announced its Quality Metrics Initiative[4], in which it engaged the pharmaceutical industry to develop a list of data FDA should request from pharmaceutical manufacturers to assess product quality risk and, therefore, aid in its risk-based resource allocation decisions. Additionally, the data requested could provide an indication to FDA of risks to drug supply disruption and can assist investigators in defining where to focus inspectional time spent in the manufacturing plants for more efficient and effective inspections.

As a result of FDA's outreach to the industry, several initiatives were undertaken to define, collect, and analyze a wide array of quality metrics that could be used by FDA. Several organizations, including: the Pharmaceutical Research Manufacturers Association (PhRMA), the Parenteral Drug Association (PDA), the Generics Pharmaceutical Association (GPhA), and the International Society for Pharmaceutical Engineering (ISPE), proposed information and metrics for FDA consideration.

During the March 2014 FDA/Xavier University PharmaLink Conference, Russ Wesdyk from FDA's Center for Drug Evaluation and Research (CDER) presented the following potential data FDA could request from the industry: lots attempted, lots rejected, lots reworked, out of specification results, and lot release results invalidated due to laboratory error or anomaly.[5] Although FDASIA gives FDA the authority to review the data collected in lieu of an inspection, Xavier University expressed in an April 2014 proposal to FDA[6] the importance of reviewing the data during an inspection in order to ensure proper context.

---

[3] FDA Safety and Innovation Act: http://www.gpo.gov/fdsys/pkg/BILLS-112s3187enr/pdf/BILLS-112s3187enr.pdf

[4] The 2015 FDA "Request for Metrics" draft guidance and the intent of its use are described on the following FDA Voice Blog page: http://blogs.fda.gov/fdavoice/index.php/2015/07/quality-metrics-fdas-plan-for-a-key-set-of-measurements-to-help-ensure-manufacturers-are-producing-quality-medications/

[5] FDA/Xavier University PharmaLink Conference, March 14, 2014 presentation by Russ Wesdyk: http://xavierhealth.org/wp-content/uploads/3.-Wesdyk_Next-Steps-for-the-CDER-Challenge.pdf

[6] "FDA/Industry Collaborative Approach to Quality: With the Patient in Mind", A Proposal submitted by Xavier University for FDA and Industry consideration. April 12, 2014. http://xavierhealth.org/wp-content/uploads/Xavier-Proposal-for-CDER-Metrics-program.15-April-2014.pdf

In May of 2014, the Engelberg Center for Health Care Reform at the Brookings Institution hosted a discussion among industry representatives and FDA officials entitled "Measuring Pharmaceutical Quality through Manufacturing Metrics and Risked-Based Assessment"[7] to assess the compilation of data proposed by various organizations.

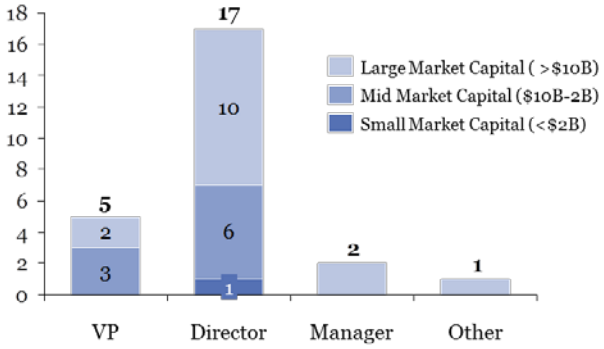
In June of 2014, Xavier University and PwC launched an initiative that could inform decisions and trigger action with the following three goals:

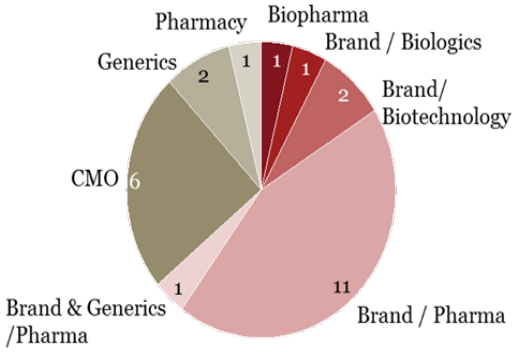
1. Identify metrics that would enable the industry and FDA to understand proactively the risk to product quality
2. Assess risk to product quality across the total product lifecycle to drive a mindset of designing quality into products at the source
3. Provide a framework that could be used by FDA to assess data gathered during an inspection and therefore, within the context in which it was generated

The Xavier and PwC team believes that accomplishing these goals will produce metrics that will prove to be meaningful both to the industry and FDA. The backgrounds of the 30 industry representatives on the team (see team list in Appendix A) can be viewed in the graphic below (not included below is additional consultant support that represented 2 owners, 2 founders and 1 president):

Size	Type	VP	Director	Manager	Other	Total
Large Market Capital	Brand / Biologics		1			1
	Brand / Pharmaceutical		6	2	1	9
	Brand and Generics / Pharmaceuticals	1				1
	CMO		1			1
	Generics	1	1			2
	Pharmacy		1			1
Large Cap Total		**2**	**10**	**2**	**1**	**15**
Mid Market Capital	Brand / Biotechnology	1	1			2
	Brand / Pharmaceutical		2			2
	CMO	2	3			5
Mid Cap Total		**3**	**6**			**9**
Small Market Capital	Biopharmaceuticals		1			1
Small Cap Total			**1**			**1**

[7] Engelberg Center for Health Care Reform at Brookings meeting, May 1-2, 2014: http://www.brookings.edu/events/2014/05/01-measuring-pharmaceutical-quality





On July 28, 2015, FDA issued a draft guidance to the industry entitled "Request for Quality Metrics,"[8] which stated that the metrics collected by FDA would be used to: (1) help develop compliance and inspection policies and practices; (2) improve the Agency's ability to predict and, therefore, possibly mitigate future drug shortages; and (3) encourage the pharmaceutical industry to implement state-of-the-art, innovative quality management systems for pharmaceutical manufacturing. The following metrics were proposed in the draft guidance:

- Lot Acceptance Rate
- Product Quality Complaint Rate
- Invalidated Out-of-Specification Rate
- Annual Product Review on Time Rate

Xavier University and PwC maintain the position that it is important to assess risk to product quality data during an inspection and, thus, in context. The remainder of this recommendation reflects the work conducted through the Xavier University/PwC Metrics Initiative and its recommendation to FDA and the industry on how to use the output of the metrics to inform decisions and trigger action.

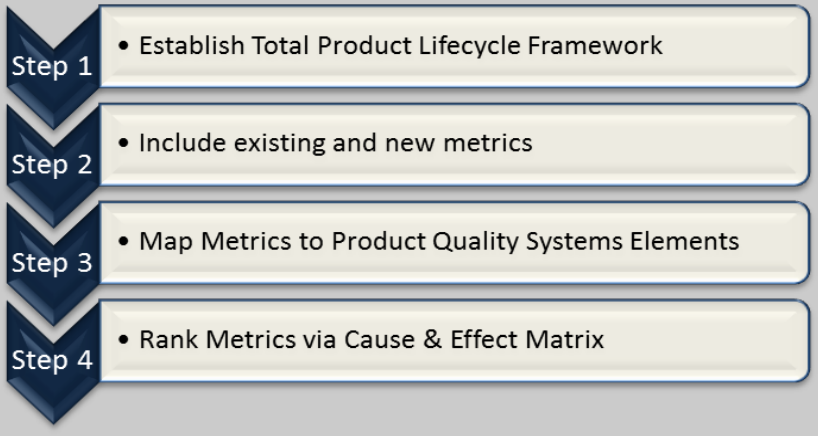
Methodology

The Xavier University/PwC Metrics Initiative involved a rigorous, four-step methodical process outlined in Figure 1 to ensure: (1) each phase of the total product lifecycle was explored, (2) existing and new metrics were considered, (3) metrics were mapped back to quality system elements, and (4) the proposed metrics were ranked against critical criteria for relevance and

[8] 80 FR 144 (July 28, 2015)

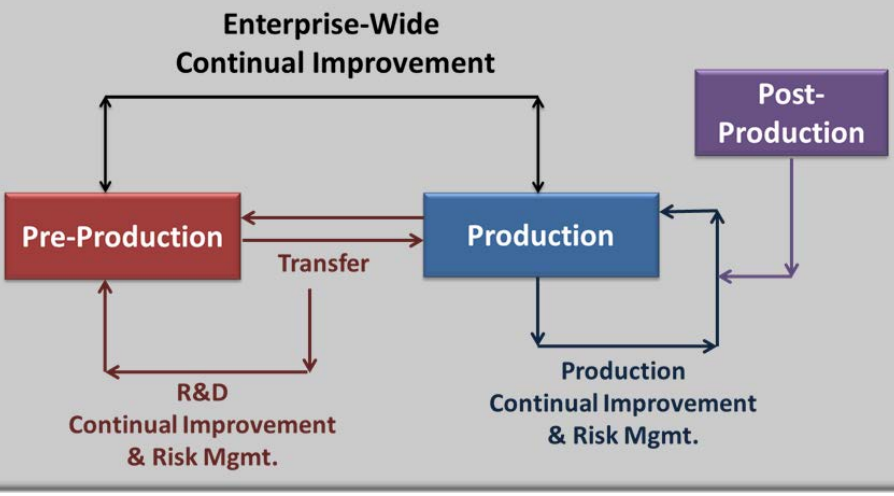
impact. Each step will be discussed in detail in this proposal, as each proved to have a significant impact on the quality of the output and rigor of the resultant recommendation.

Figure 1. Steps Involved in the Xavier University/PwC Proposal



Step 1: Establish Total Product Life Cycle (TPLC) Framework

Figure 2. Total Product Lifecycle Framework.



Traditional methods used to assess risk to product quality tend to focus on tracking and trending production and post-production metrics. However, in order to assess the cultural

mindset of designing quality into products at the source, the rigor of the development process needed to be explored in depth. Figure 2 depicts the TPLC framework developed through this initiative that incorporates continual improvement within each phase of production, as well as across the entire enterprise.

Xavier University and PwC divided the team of industry representatives who volunteered to participate (names listed in Appendix A) into three groups: pre-production, production, and post-production. By focusing on a single phase of the TPLC, each team was able to explore in depth what could be measured in a meaningful way, what information might be required from other phases within the framework, and how the output of the metrics assessed could inform decisions and trigger actions.

At the beginning of the initiative, identifying what to measure during the pre-production phase posed a challenge. The concepts of trial and error, testing to the edge of failure, and unknown-unknowns[9] made it difficult for the team to identify "failures" that would provide an indication of the success of the development process. The team asked itself at what point R&D says, "I believe the product and process are developed." By doing so, each member of the pre-production group could explore how to measure failures after that point and, importantly, how to feed that information back into R&D for continual improvement. The pre-production group identified specific design space elements that must be completed prior to transfer in order to improve the success rate of the product and process in production. Additionally, during transfer, any product- or process-related failure could be attributed to the rigor of the development process and could, therefore, be used to improve the overall system of product development in a way that would increase the success rate of future similar projects.

> "As we know, there are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns—the ones we don't know we don't know." [9]

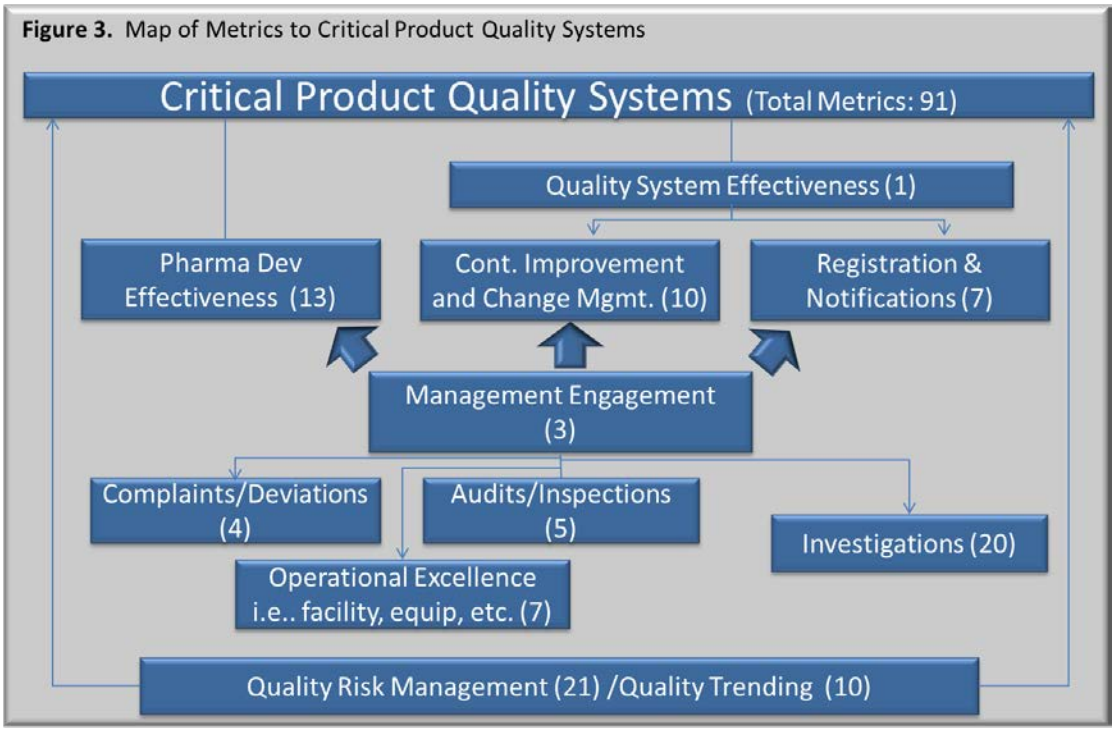
[9] Quote from United States Defense Secretary, Donald Rumsfeld, to United States Department of Defense. February 12, 2002.

Step 2: Include Existing and New Metrics

Each group started with a consolidated list of existing metrics from the May 2014 FDA/Brookings meeting (Appendix B) to ensure that a wide range of industry input was included. They then gathered additional ideas within their own organizations as well as ideas they themselves generated. Each group identified new metrics by narrowing their focus on ways to measure product quality risk within its assigned phase of production, determining what information would be needed from other phases of production, and assessing how the output of any new metrics could be used to inform activity in other phases of production. High-level definitions were assigned to each metric so that the metrics identified from all three groups could be compared and consolidated into one list. This consolidation was accomplished through cross-group discussion and understanding of the interdependence of the metrics across phases, and resulted in 101 total metrics (Appendix C). All 101 metrics provided in Appendix C include a high-level definition and are linked to the appropriate phase of production.

Step 3: Map Metrics to Product Quality System Elements

Figure 3. Map of Metrics to Critical Product Quality Systems



Despite the vast array of metrics identified, the team wanted to ensure that metrics were associated with each of the quality systems that are critical to reducing product quality risk. In order to identify potential gaps, the team agreed upon the 11 critical product quality systems shown in Figure 3, and mapped 91 of the metrics to those systems (the team determined that 3 of the 101 metrics appeared to be duplicative, and also removed 7 metrics that were specific to sterile products). The number of metrics associated with each critical system is shown in parentheses next to each system. As a result of the exercise, the team ensured that all critical product quality systems were covered and spanned all three phases of production.

Step 4: Rank Metrics via Cause & Effect Matrix

Recognizing that "Not everything that counts can be counted, and not everything that can be counted counts,"[10] Xavier University and PwC developed a cause and effect matrix (C&E matrix) with the team through which the remaining 91 metrics were assessed against pre-defined critical criteria. This tool, and the subsequent Pareto analysis, allowed the team to determine which of the 91 metrics would provide significant linkage to critical risk factors.

Table 1. Scoring Mechanism for Cause and Effect Matrix.

Requirement	Weight	Scoring
Patient Safety: Likelihood that a poor result from the metric would lead to patient harm resulting from poor product quality	5	Highly Probably = 9 Slightly Probable = 6 Slightly Improbable = 3 Highly Improbable = 0
Quality System Robustness: Likelihood that a poor result from the metric would provide an indication of larger systemic quality failures	4	
Process Reliability: Likelihood that a poor result from the metric would lead to a missed opportunity to avoid recurrence of failure	3	
Supply Assurance: Likelihood that a poor result from the metric would indicate a decrease in the probability of getting the right product and of the right quality to the patient when needed	2	
Failure Costs: Likelihood that a poor result from the metric would lead to increased Cost of Quality	1	

[10] William Bruce Cameron. "Informal Sociology: A Casual Introduction to Sociological Thinking". 1963

The problem statement used to establish the C&E matrix was, "We need to identify measures that provide an indication of the degree of product quality risk." As shown in Table 1, five critical customer requirements were then identified that provide insight to product quality risk: patient safety, supply assurance, process reliability, quality system robustness, and failure costs. Next, a weighting was assigned for each of the critical customer requirements based on its perceived importance. A simple rank of 1 through 5 was utilized with 1 representing the least important requirement and 5 representing the most important. The team ranked the attributes in the following order of decreasing risk: patient safety (5), quality system robustness (4), process reliability (3), supply assurance (2), and failure costs (1).

Each team member then assessed all 91 metrics against all five critical customer requirements using a four-tiered scoring system to determine the probability that a poor result of the metric would result in an impact to the critical customer requirement. If a poor result from the metric had a high probability of affecting the critical requirement, then it was scored with a 9. This and the remaining scoring possibilities are shown in Table 1.

Through the C&E matrix, each metric was scored by multiplying the weight of the critical customer requirement ("CCR") by the probability score given by each team member. The addition of the five subtotals generated a total score for each metric.

Figure 4. Cause and Effect Matrix Scoring Example.

Critical Requirements >>>>		Patient Safety	Supply Assurance (availability)	Process Reliability	Quality System Robustness	Failure Costs	
Importance Rating of CCR>>>		5	2	3	4	1	
Metric	Metric Definition	0 = Highly Improbably; 3 = Slightly Improbable; 6 = Slightly Probable; 9 = Highly Probable					Total
# of process changes as a result of inadequate process development	Measured as a lagging indicator during production phase, but can serve as a leading indicator for future process development in pre-production	9	9	9	9	9	135

In the example shown in Figure 4, the number of process changes due to inadequate development resulted in a high probability of impact to each of the five critical requirements, meaning that a poor result (i.e. many changes necessary due to inadequate development) would likely lead to risk to patient safety, lack of supply in the field, recurring failures in manufacturing, systemic quality issues, and high costs. The calculation for the overall score would be (9 × 5) + (9 × 2) + (9 × 3) + (9 × 4) + (9 × 1) for a total score of 135, which is also the maximum possible score.

A Pareto analysis of the results from 23 respondents was conducted based on average rank to generate the list of top 15 metrics shown in Table 2. These metrics ranked as the top 15 due to their strong correlation to the critical criteria of the C&E matrix.

Not surprisingly, the majority of those 15 metrics correlate most strongly with the critical customer requirements of the C&E matrix, since these metrics are associated with times when the product has already failed. As a result, there is certain impact to the patient, supply, process, quality systems and cost. Interestingly, since it is difficult to say that these metrics are not critical, industry groups and FDA officials have consistently identified one or more of these metrics as important measures of product quality risk.

Table 2. List of Top Metrics from Cause and Effect Matrix.

1. Total recalls	9. Critical investigations rate
2. Critical complaint rate	10. Confirmed OOS rate (drug substance & product; stability)
3. Media fill failure rates	11. Product quality complaint rate
4. Field Alert Report rate	12. Process capability assessment or rate (CPK; PPK)
5. % successful media fills	13. Health Authority inspections: # of inspections, and # of critical & major observations
6. Adverse event rate	14. Batch reject rate
7. # media fill failure requiring validation	15. Lot acceptance rate
8. Recurring deviation rate	

The C&E matrix is a powerful tool that enabled the team to recognize why the strong correlation existed (i.e., the product has already failed). Since the goal of the initiative was to identify metrics that could proactively provide an indication of product quality risk, the team could use the Pareto analysis of the C&E matrix results to exclude quality failure metrics and, therefore, dive to the next tier of metrics. The remaining metrics in the C&E matrix were sorted by phase of production (pre-production, production, or post-production) in rank order. Each team member was then given 20 points to vote on the metrics with the guiding principle of focusing on designing quality into the product throughout the total product lifecycle, as opposed to catching inadequate quality. The results were aggregated and the team met in person at the FDA/Xavier University PharmaLink Conference in March 2015 and at Xavier

University in June 2015 to finalize the results. The resultant top metrics were discussed in relation to the Total Product Lifecycle framework provided in Figure 2.

The team finalized the list of metrics shown in Table 3 and reiterated the importance of assessing each metric in the context of a larger system, instead of a single metric in isolation. The Total Product Lifecycle framework shown in Figure 2 was confirmed as critical in order to demonstrate the interconnectivity of the metrics, critical feedback loops, and mechanism by which lagging indicators become leading indicators – all of which are critical for the industry to understand fully any existing risk to product quality.

Table 3. Final list of Proposed System of Metrics.

Total Product Lifecycle Phases	Proposed Metrics
Pre-Production	1. Design Space
	2. Supply Chain Assurance
Technology and Knowledge Transfer	3. RFT of Analytical Method Transfer
	4. RFT of Process Validation
Production	5. RFT for Production
	6. CAPA Effectiveness Rate
	7. Commitment Index
	8. Supplier Risk Index
Post Production	9. Market Reliability Index
Enterprise-wide Continual Improvement	10. Quality by Design Effectiveness
	11. Root cause analysis of the Production RFT metric

Discussion of Results

The list of final metrics is provided in Appendix D, along with definitions, clarifications, formulas, and notes. It is important, however, that recognize that since the Xavier/PwC Initiative is focused on identifying metrics that industry can use to assess itself and reduce risk, there is not a need for universal definitions of the metrics. Definitions are provided in Appendix D in order to provide a starting point, but each company should work to define the terms in a way that is meaningful for its products and its business. In order to experience the value of the recommended metrics, companies should work to stay in-line with the intent of each metric identified through this initiative, and should not change definitions throughout the year, or

from year to year, in order to artificially make the trends look positive. Gaming the metrics is always possible, so the results of this initiative can best be used by companies that are truly interested in self-improvement and reducing product quality risk. The Total Product Lifecycle framework in Figure 2 provides a mechanism of continual improvement based on designing quality into the product at the source. Re-framing the final list of metrics into this system resulted in a better understanding of what to measure and how.

Discussion of Pre-Production Metrics

1. **Design Space Metric**

$$\frac{\text{\# of projects completed with scientifically justified predefined ranges}}{\text{Total \# of projects completed}} \times 100$$

The team was careful to recognize that "inadequate" development does not necessarily equate to poor development. Relative to what may be measured during the pre-production phase, it was identified that product and process development does not always include justification and data to support the critical process parameters (CPP), critical material attributes (CMA), and critical quality attributes (CQA) that are proposed. Development work does not always include experimental or statistical verification of the appropriateness of historical ranges used for other/similar products before adopting those ranges for the product in question. In order to decrease risk of product failure and patient harm, these design elements need to be scientifically supported by experiment or acceptable statistics. As a result, the team identified a Design Space metric that measures the number of projects completed with predefined ranges (with justification) versus the total number of completed projects. The team recognized that having predefined ranges does not in and of itself reduce failure if the ranges are not product- and process-specific and defined with scientific rigor. Therefore, the Design Space metric is ultimately used in conjunction with the RFT production and transfer metrics, as well as the QbD Effectiveness metric to provide a more holistic assessment of the effectiveness of the development and technology transfer processes.

2. **Supply Chain Assurance**

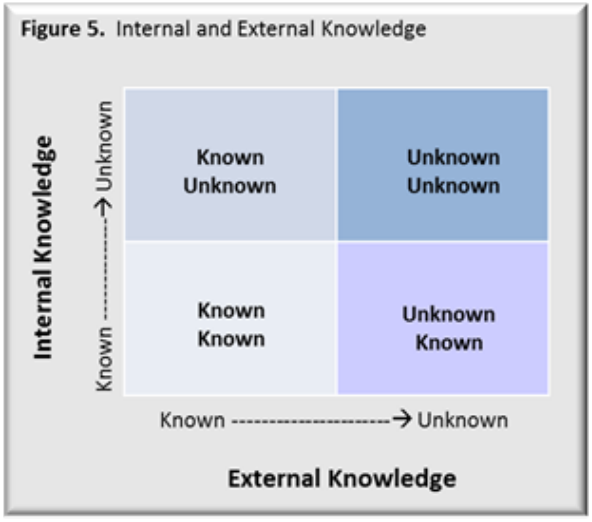
$$\frac{\text{\# of Tier 1 suppliers approved through cross-functional review}}{\text{Total \# of suppliers in the supply chain for the product in question}} \times 100$$

During pre-production, the R&D group has the opportunity to influence a low risk supply chain from the beginning (when possible). The Supply Chain Assurance metric proposed during pre-production is based on the assumption that product quality risk can be improved by putting in place a cross-functional review of the proposed suppliers as part of the supplier selection and qualification process. Alignment of requirements (i.e. quality, cost, capability, capacity, etc.) must be assured across key functional groups in order to identify a supplier that best fits the needs of the product and business. Although this concept is widely accepted, it is not often followed. Therefore, this metric provides visibility into how often cross-functional approval is included in the supplier selection process to help improve the systems that establish supply chains for each product. This same metric is to be used to assess the effectiveness of new suppliers needed throughout the lifecycle of the product.

Discussion of Transfer Metrics

Utilizing the Total Product Lifecycle framework in Figure 2 resulted in the identification of the transfer phase as a critical step and provided an opportunity to think through how to measure Right First Time (RFT) as early in the development cycle as possible. Again, once the transfer of product, process, and analytical methods begins, R&D has essentially said to itself "I believe the product, process, and methods are developed." Therefore, any product-, process-, or method-related failures that occur during transfer can be attributed to inadequate development.

Again, "inadequate" development does not necessarily equate to poor development. As shown in Figure 5, some knowledge is just not known – internally or externally. The goal is to reduce the size of the unknown unknown box as much as possible but it will always exist. Continual feedback of new learnings into the development process for future product development work can help reduce the risk of unknown unknowns.



The team recognized that during product,

process and method transfer work, the measurement of process validation attempts and analytical method transfer success would provide an early indication of the rigor of the development process. By measuring the success of these activities, trends can be identified across products to identify areas of opportunity for improving the initial product development process through the R&D Continual Improvement loop. Traditionally, transfer failures are resolved for the success of the product in question but are not always cycled back to improve the overall development process for future products. Additionally, the failure rate at the transfer stage is not typically trended and assessed at the Senior Management level.

The use of the following metrics during transfer work will help highlight the rate of failures to assist in identifying root cause trends and improving the overall development system:

3. **Process Validation Right First Time**

$$\frac{\text{\# of process validation batches without product/process related deviations}}{\text{Total \# of validation batches attempted}} \times 100$$

4. **Analytical Method Transfer Right First Time**

$$\frac{\text{\# of analytical methods transferred with no method related deviations}}{\text{Total \# of method transfer attempts per product}} \times 100$$

Discussion of Production Metrics

5. **Production Right First Time**

$$\frac{\text{\# of batches/lots without deviations}}{\text{Total \# of batches/lots attempted}} \times 100$$

The goal of our work is to shift to the Right First Time mentality as early in the product development cycle as possible, so it is not surprising that the Production Right First Time metric is included in our final list of metrics. Importantly, however, the results from this metric need to be assessed for root cause trends and shifts in order to ensure the systems that govern production work are continually monitored and improved.

6. CAPA Effectiveness

$$\frac{\text{\# of successful effectiveness checks}}{\text{Total \# of effectiveness checks attempted}} \times 100$$

CAPA effectiveness was chosen to capture the "right second time" success and linkage to elimination of repeat failures.

7. Commitment Index

(Investigations × 0.2) + (Customer Complaints × 0.2) + (CAPA × 0.1) + (APR × 0.1) + (Stability × 0.1) + (Training × 0.05) + (Audits × 0.05) + (PM × 0.05) + (Reg. Commitments × 0.1) + (Revalidations × 0.05)

The team created the Commitment Index (each term in the above calculation is defined in Appendix D) to assess the commitment of the organization to follow through on deadlines across multiple expectations, in addition to viewing each commitment in isolation. A holistic view is taken of the organization's performance related to on-time completion of the following standard practices and FDA expectations: investigations, complaints, CAPA, Annual Product Reviews, stability testing, GMP training, audits, PM/Calibration, Regulatory Commitments, and revalidations. This metric can help reveal the strength of an organization's quality culture, the shared cross-functional ownership of quality, and the adequacy of staffing of key functional groups.

8. Supplier Risk Index

The Supplier Risk Index assesses supplier risk based on qualitative and quantitative factors, such as level of concern related to performance, audit findings, geographical risk, leverage[11], capacity, and status of necessary agreements.

The following factors (A through G) are to be measured using a scale: 0, 5, 10 (where 10 is good):

A. Level of confidence relative to performance of supplier, as measured by complaints related to the supply in a given time period based on the number of lots received
B. Level of confidence relative to audit/regulatory findings in a given time period (if no audit in given time period, then previous results apply)

[11] "Leverage" refers to the amount of negotiating power an organization has with its supplier

C. Necessary agreements (i.e. supply agreement, quality agreement) are in place
D. The supplier has sufficient capacity and or redundancy such that risk of a shortage is lowered
E. Level of confidence relative to geographical risk (e.g. under-regulated regions of the world)
F. Level of confidence related to leverage and supply stability — assessment of the % of supplier's bottom line attributed to our business
G. Level of confidence in track record of the supplier (previous materials supplied)

Formula: $A + B + C + D + E + F + G \leq 70$

Suggested Actions based on Score:

- 60 – 70: No action required, assuming all responses are 5 or higher
- 40 – 55: Cross-functional assessment of mitigation strategies, as well as meetings with suppliers to identify improvement opportunities
- 20 – 35: Cross-functional escalation of risk awareness, assessment of supplier alternatives and mitigation strategies, heightened involvement in supplier operations, and oversight.
- 0-15: Cross-functional escalation of risk mitigation requirements, identification of alternate source of supply, integral involvement with supplier operations, and oversight

Discussion of Post-Production Metrics

9. Market Reliability Index

(100 - % Customer Complaints) × 0.15 + (100 - % Adverse Events) × 0.15 + (100 - % Drug Shortages) × 0.30 + (100 - % field alerts) × 0.20 + (100 - % Recalls (will intentionally include those issues already captured in field alerts)) × 0.20

The post-production metric is a holistic picture of market reliability (each term in the above calculation is defined in Appendix D). Again, in addition to viewing any single metric in isolation, the above listed post-market signals are assessed in aggregate: complaints, adverse events, drug shortages, field alerts, and recalls. The team considered several other indicators, such as stability failures, but recognized that these would be captured as field alerts and/or recalls. This index can also feed into a scorecard

or heat map to identify areas of higher and lower risk quickly, thus informing future decisions and triggering action.

Discussion of Enterprise-Wide Metrics

10. Right First Time Rate for Production (root cause triage)

$$\frac{\text{\# of batches/lots without deviations}}{\text{Total \# of batches/lots attempted}} \times 100$$

Although Right First Time (RFT) in production is commonly measured, a methodical triage of root causes related to failures is the key to ensuring the RFT output is used to inform decisions and trigger action. The team proposes that failures with root causes related to inadequate product and process development are communicated back to R&D through the Enterprise-Wide Continual Improvement loop (refer to Figure 2 for the TPLC diagram). For this reason, the RFT production metric is listed again in the Enterprise-Wide phase of the TPLC. There is no difference in the formula from the production metric, however, the output in this instance is used to reduce the unknown unknowns (refer to Figure 5) and, therefore, improve the overall product development process.

11. Quality by Design Lifecycle Effectiveness

(100 - % Customer Complaints)* 0.25 + (100 - % Process Capability) * 0.25 + (100 - % Stability Failures)* 0.25 + (100 - % Product Failures)* 0.25

All terms above are defined in Appendix D. The Enterprise-Wide Continual Improvement loop provides a mechanism by which production and post-market failures related to product and/or process design can be tracked, trended, and communicated back to development in order to improve the overall development process. For example, signals received through post-market surveillance can be triaged in production to determine root causes related to development. Those root causes can then be communicated back to development as a Quality by Design Effectiveness metric through the Enterprise-Wide Continual Improvement loop and, therefore, become a leading indicator for future products. This process in effect shifts the “right first time” mindset to one of designing

quality into the product at the source. By doing so, unknown unknowns, failures, and costs can be reduced, while speed to market can be increased.

The team proposes that each company modifies/customizes each of the eleven metrics in a way that makes sense for its business, yet maintains the spirit of the intent. For example, some companies might measure batches versus lots, collect the data quarterly instead of monthly, or include additional factors in the indices due to known issues they are working to resolve. It is critical that the definitions are not altered once implemented in an effort to avoid demonstration of false improvement. Since the team recommends that FDA view these metrics in context, company specific modification of the metrics is not an issue.

Recommendation and Conclusion

Through the system of metrics proposed in Table 3 and in recognition of the importance of the use of these metrics in the TPLC framework proposed in Figure 2, all three goals of the Xavier University/PwC Metrics Initiative were accomplished:

1. Identify metrics that would enable the industry and FDA to understand risk to product quality proactively
2. Assess risk to product quality across the total product lifecycle to drive a mindset of designing quality into products at the source
3. Provide a framework that could be used by FDA to assess data gathered during an inspection and therefore, within the context in which it was generated

The Xavier University/PwC Metrics Initiative employed a methodical approach to identify proactive metrics of product quality risk that can be used by industry representatives and FDA officials. The team assessed metrics for each phase of production (pre-production, production, and post-production), mapped the metrics against critical quality systems, compared the metrics against critical customer requirements through the use of a cause & effect matrix, created a Total Product Lifecycle framework for the metrics, and determined appropriate calculations for each metric.

Xavier University and PwC recommend that the industry and FDA utilize the proposed system of metrics in Table 3 in conjunction with the TPLC framework provided in Figure 2 in order to establish a more holistic view of product quality risk. The final proposed system of 11 metrics can assist organizations in identifying risk to product quality. By using all 11 metrics in a

systemic approach, an organization can self-assess, gauge, and continually improve the risk to the quality of its products, while striving to design quality into its products at the source. The metrics allow companies to self-identify issues and potential issues for appropriate action, and to communicate potential shortages to FDA in a more timely fashion. Since it is proposed that the metrics be reviewed by FDA during an inspection (and therefore, in context), FDA can utilize the system of metrics as a roadmap to identify data linked to patient safety, quality system robustness, process reliability, and supply assurance.

Through the system of metrics approach provided within this white paper, both the industry and FDA may gain a better understanding of risk to product quality.

Appendix A: List of Team Members

This white paper represents the individual professional opinions of the following industry contributors, and not necessarily the position of the companies they represent. Only team members who gave their consent to be listed are provided below.

The Xavier/PwC Initiative was led by Marla Phillips (Director, Xavier University) and Sam Venugopal (Partner, PwC)

First	Last	Title	Company
Dee	Abelha	Director, Global Compliance Audits	Bristol Meyers Squibb
Kanshit	Bheda	Manager, Quality Systems Operations	PwC
Grace	Breen	VP Corporate Quality	Impax Laboratories
Laura	Cannon	Senior. Director, Quality Intelligence & Knowledge Management	Teva Pharmaceuticals
Brian	Carlin	Director Open Innovation	FMC Health & Nutrition
Dee	Carri	Founder and Director	Torque Management Limited
Patrick	Crowley	Owner	Callum Consultancy
James	Horger	Senior Director, Quality Systems & Compliance	Mallinckrodt Pharmaceuticals
Daniel	Jordan	Quality Director	Shire
Anil	Kane	Executive Director, Global Head of Formulation Development	Patheon
Jonathan	Lee	Senior Associate, Quality Systems Operations	PwC
Kimberley	Mandrell	Senior Project Manager	Mallinckrodt Pharmaceuticals
Mike	Markham	Associate Director Analytical Research	Adare Pharmaceuticals, Inc.
Sean	McCrossen	Director	IEXA100 Consulting
Andrew	McNicoll	VP Quality Systems and Compliance	Patheon
Joseph	Northington	Head of Quality	Purdue Pharma L.P.
Christophe	Pamelard, Pharm. D.	Associate Director Quality	Purdue Pharma L.P.
Peter	Pitts	President	Center for Medicine in the Public Interest
Anil	Rattan	Associate Director	Shire

Kathy	Regelski	Quality Systems Coordinator	Adare Pharmaceuticals, Inc.
Prakash	Savarirayan	Head of Operations Procurement	Shire
Kevin J.	Slatkavitz	President and Founder	ThinkQuality, LLC
Snehal	Srikrishna	Manager, Quality Systems Operations	PwC
Jamie	Wilson	Director, Quality Assurance	Navidea Biopharmaceuticals
Bob	Zinser	VP, Commercial Technology, N.A.	Patheon

Appendix B: List of Initial Metrics

#	Metric Name	#	Metric Name
1	Development / Supplier Quality Risk Management (QRMs)	28	Cycle Time (disposition and end-to-end) Rate
2	Risk Management & Mitigation Profile Changes	29	Supply Chain Cycle Time
3	Adherence Preventative Maintenance Level	30	Upside Supply Chain Adaptability
4	% of Overdue PM for Critical Equipment Rate	31	Order Fulfillment by Line
5	Maintenance and Calibration of Measuring Equipment	32	Plant Capacity
6	Unplanned Downtime (due to unplanned maintenance including utility failures)	33	Plant Utilization
7	Recapitalization as % of Asset Value Rate	34	Forecast Accuracy
8	PM as % of Asset Value Rate	35	# of Direct Material Suppliers
9	Lot Acceptance Rate	36	# of Mfg. locations (third party)
10	Confirmed OOS Rate	37	Inventory (days of supply)
11	Invalidated OOS Rate	38	Supply Chain Adherence
12	Confirmed OOT Rate	39	Redundant Capacity
13	Stability Failure Rate	40	Potential Stock-out or Drug Shortage Rate
14	Right First Time	41	Product Quality Complaint Rate (Total, Critical)
15	Mean time between failures (MTBF)	42	Adverse Event Rate
16	Lot Disposition Rate / Time	43	Recall Rate (Class I/II)
17	Lot Yield	44	Annual product quality review (on time performance)
18	Rework Rate	45	Health Authority inspections # of Inspections # of critical & major observations
19	Reprocessing Rate	46	Audit / Inspectional Commitment On-Time Completion Dates Rate
20	Deviation Rate	47	Supplier Complaints
21	Recurring Deviation Rate	48	Lead time for Investigations (cycle times, ability to close)
22	Investigation Free Lots Rate	49	CAPA Effectiveness Rate
23	Process Capability Assessment (cpK)	50	Quality Culture
24	Right Second Time	51	Training Effectiveness
25	Medial Fill Failures / Success Rate	52	% Quality Assurance (QA)/ Quality control (QC) staffing
26	% Lots with Environmental Monitoring (EM) Excursions	53	Quality Trending
27	% Lots Rejected for EM Excursions	54	Organizational Health Metric (percentage of temporary workforce, employee satisfaction %, safety, employee turnover rate)

Appendix C: List of 101 Metrics

Metric	Description	Phase of Production
Risk Profile Mitigation Rate	Number of completed (product quality, compliance, validation, etc.) mitigation plans in the period/(Number of risks identified in the period + Open Risks from previous periods) x 100%	Production and Post-Production
% Known Risks with Mitigation Plans	Number of mitigation plans currently active/the number of items that require mitigation in the risk management profile x 100.	Production and Post-Production
Risk Mitigation Plans from Quality Trends	Number of mitigation plans initiated as a result of a proactive Quality Trend in a given time period.	Production and Post-Production
Adherence Preventative Maintenance and Calibration Level	Percentage of Preventive Maintenance & Calibrations completed on schedule in the defined period expressed as number of completed on schedule/total number scheduled x 100%	Production
% of Overdue PM for Critical Equipment Rate	Percentage of Overdue Preventive Maintenance activities for critical equipment in the defined period expressed as number of overdue/total number scheduled x 100%	Production
Unplanned Equipment Downtime (due to unplanned maintenance including utility failures)	Percentage of lost productivity due to unplanned maintenance of equipment, utilities and support systems expressed as hours lost/total hours planned x 100	Production
Re-capitalization as % of Asset Value Rate	Dollars spent on replacement capital for the facility/process during the time period (Quarter or Year)/Total Dollar of Asset Value at end of time period * 100	Pre-Production and Production
Preventative Maintenance as % of Asset Value Rate	Dollars spent on facility/process Preventative Maintenance during the time period (Quarter or Year)/Total Dollar of Asset Value at end of time period * 100.	Production
Batch Reject Rate	Number of rejected batches/number attempted batches x 100% An attempted batch is after the first production step is initiated. Only includes entire batches that are rejected. Partial rejections are not considered rejects. If the application allows reprocessing, this would not be considered a rejected batch.	Production
Lot Acceptance Rate	Number of batches released/number of batches attempted x 100%	Production

	An attempted batch is after the first production step is initiated. Does not include entire batches that are rejected. Partial rejections are not considered rejects. If the application allows reprocessing, this would not be considered a rejected batch.	
Confirmed OOS Rate (Drug Substance & Product; stability)	Confirmed OOS results attributed to the product. Number of Confirmed OOS results/Total Number of tests performed x 100%	Production
Invalidated/Unconfirmed OOS Rate	Invalidated OOS results are those found not to be product related post investigation. Ruled out during Phase I investigation or during Phase II laboratory investigation. Number of invalidated OOS results/Total Number OOS investigations in the period x 100%	Production
Confirmed OOT Rate	Number of confirmed Out of Trend test results/batches produced x 100% - may be calculated on a per product basis.	Production
Right First Time	Number of batches dispositioned without potentially product impacting exception/deviation, investigation, OOS, rework, or rejection/total lots dispositioned for the defined period x 100% Recommend using batch instead of lot — meaning processing steps, production trains/department or Batch Records. Focus on events that could potentially impact the product.	Production
Mean time between failures (MTBF)	Total up time divided by number of breakdowns. See: http://world-class-manufacturing.com/KPI/mtbf.html	Production
Lot Disposition Rate/Time	Number of Lots Released/Number of Lots Attempted x 100% per defined timeframe. Definition of Lots Attempted: Finished Dose begins when first critical raw material (defines Exp. Date) is added to process vessel. Exclude pre-weighs. API processing definition: Multi-step process, once the critical intermediate is engaged.	Production
Lot Yield	Actual manufacturing yield/theoretical yield x 100%	Production
Rework Rate	Number of reworked (in order to disposition) batches /total number of batches dispositioned during time period x 100	Production
Reprocessing Rate	Number of reprocessed (in order to disposition) batches /total number of batches dispositioned during time period x 100	Production

	Reprocessing steps that are defined in the batch record are not included in this calculation.	
Deviation Rate	Number of batches with potential product impacting deviation investigations/total number of manufacturing batches x 100 Total Number of deviation investigations/month Need to define the criticality of deviations to be included in this metric.	Production
Critical Investigations Rate	number of batches with critical investigations/total number of manufacturing batches during time period x 100	Production
Critical Complaint Rate	Number of critical complaints/batches shipped.	Post-Production
Recurring Deviation Rate	Number of batches with potential product impacting recurring deviation investigations/total number of manufacturing batches x 100 Total Number of recurring deviation investigations/month Need to define the criticality of deviations to be included here. Recurring is defined as same root cause within the production train	Production
Investigation Free Lots Rate	Number of batches investigation-free/total number of batches dispositioned x 100	Production
Process Capability Assessment or Rate (CpK, PpK)	Process Capability is measured for each of the critical processing steps. Cpk is an index (a simple number) which measures how close a process is running to its specification limits, relative to the natural variability of the process. The larger the index, the less likely it is that any item will be outside the specifications.	Production
Is Process Capability Measured? (Y/N)	Yes/No question. Yes would be selected if process capability is measured for at least one critical processing step of the process.	Production
Right Second Time	Number of reworked and reprocessed batches which are acceptable after the rework and reprocessing /total number of batches dispositioned during time period x 100 This metric in combination with RFT, will show how close we can get to RFT after a second processing step.	Production
Number of media fill failures requiring revalidation	Number of media fills failed requiring validation/ total number of media failures Failure is based on failing protocol acceptance criteria.	Production

Media Fill Failures Rate	Number of media fills failed/total number of media fills attempted during time period x 100 Failure is based on failing protocol acceptance criteria.	Production
% Successful Media Fills	Total number of media fills passing/total number of media fills attempted during time period x 100 Failure is based on failing protocol acceptance criteria.	Production
Environmental Monitoring Rate (excursions in A/B areas)	Number of EM action level failures in AB areas/total number of samples in area during time period x 100	Production
% Lots with Environmental Monitoring (EM) Excursions	Number of batches with EM excursions/total number of manufacturing batches during time period x 100 Does not include EM alert limit excursions. Include personnel monitoring.	Production
% Lots Rejected for EM Excursions	Total number of lots rejected because of EM excursions/total number of lots manufactured X 100	Production
Environmental Monitoring (number of sterile lots with investigations related to action limit excursions)	Number of sterile lots with investigations related to action limit excursions vs. total number of lots manufactured x 100	Production
Cycle Time (disposition and end-to-end) Rate	Disposition: Average number of days from completion of production to Quality Unit release of produced material (average is calculated using all products produced by company or site) End-to-End: Average number of days from start of the process (first processing step) to Release of finished product. - Reported on product by product basis. Note: This is complicated by campaign/multi -step processes.	Production
Upside Supply Chain Adaptability The quantity of increased production an organization can achieve and sustain in a period of time	Total available, unused, annual finished product output potential of all plants within the supply chain. (calculation is completed for each production train or operation) Where contract manufacturers are used, the calculation should only include capacity that the contract manufacturer is capable of providing given their other commitments. Includes how much you can do with the existing work-force and allowable work day length.	Production
Number of Direct Material Suppliers	Number of suppliers whose materials are incorporated into final products	Pre-Production and

		Production
Number of Manufacturing. locations (third party)	Number of third party manufacturing locations for drug substance, drug product and final packaging. (calculate by product)	Pre-Production and Production
Material Inventory (components, API drug product)	Number of days manufacturing can be completed with existing inventory.	Production
Lots on hold/inventory on hold	Number of produced lots on hold for reasons other than normal release evaluations vs. total number of lots manufactured	Production and Post-Production
Supplier Supply Chain Adherence	% of materials delivered on time according to the Master Service Agreement (number of orders delivered on time based on PO delivery date for each supplier/total number of PO orders delivered for each supplier x 100). This is tracked for materials critical to the supply.	Production
Drug Shortage Notifications	Number of drug shortages in past 12 months	Post-Production
Product Quality Complaint Rate	Number of complaints vs. batches shipped	Post-Production
Adverse Event Rate	Number of adverse medical events vs. batches shipped	Post-Production
FAR/BPDs	Number of field alerts vs. lots batches shipped	Post-Production
Total Recalls	Number of recalls per year	Post-Production
Annual product quality review (on time performance)	APR's completed on time vs the number of APR's at the site	Post-Production
Health Authority (Audit) inspections number of Inspections number of critical & major observations	Number of Inspections Number of critical & major observations (total and per inspection)	Post-Production
Audit/Inspectional Commitment On-Time Completion Dates Rate	Number of audit remedial actions completed on time vs the total number of remedial actions	Post-Production
Supplier Complaints	Number of complaints issued to suppliers (includes materials and service providers) vs total number of orders received.	All
Investigation closure time	Average time required to close all deviations	Post-Production
Lead time for Investigations (cycle times, ability to close)	Number of open deviations at the end of period (monthly)	Post-Production

Lead time for Investigations (cycle times, ability to close)	Number of investigations closed on time vs total investigations closed	Post-Production
CAPA Effectiveness Rate	Number of relevant deviations after a CAPA had been implemented	Post-Production
% Quality Assurance (QA)/ Quality control (QC) staffing	Number of Quality Assurance personnel/Total personnel at site including temporary and contract personnel x 100 in the reporting period. Number of Quality Control personnel/Total personnel at site including temporary and contract personnel x 100 in the reporting period. Include contract and temporary personnel supporting processing/testing/disposition activities.	Production
CAPA cycle time	Average Number of Days to Close (from Date Opened to the Date Closed, Including Effectiveness Check Time period).	Production
Outstanding CAPA's %	Number of open CAPA's at end of a period (monthly) vs the Number opened. (Supplier Corrective Action Reports SCAR, excluded).	Production
% of Products under CPV (Continuous Process Verification)	Number of products under CPV/number of products at the site	Post-Production
Excursions (Temp, Time) during Transportation	Number of Excursions (Temp, Time) during Transportation vs. number of shipments	Pre-Production and Post-Production
CAPA Rate (APR)	The number of corrective or preventative actions that were initiated due to an APR, divided by the total number of APRs generated.	Production
% Supplier audits completed to Schedule	% of supplier audits completed at the end of the month that were scheduled for the month. For Pre- Production: % of Supplier audits performed prior to manufacture of commercial product	Pre-Production and Production
% First Pass Yield - Incoming Inspection	Ratio of Number of Receipts Accepted First Time/Total Number of Receipts at Incoming Inspection at end of a period (monthly).	Production
% CAPAs Currently Overdue	% of CAPAs that are open at the month end close date and are currently overdue, regardless of the CAPA stage. If a CAPA has an approved active extension, it is considered on time.	Production
CAPAs Initiated	Total number of CAPAs opened during the period (monthly)	Production
CAPAs Closed	Total number of CAPAs closed during the period (monthly)	Production

% CAPAs Open with due date extensions	Total Number of CAPA with Extensions that are currently within their extension timing at period end (monthly)/the total number CAPA open at month end.	Production
% CAPAs Open more than 1 year	Total number of CAPAs open > 1 Year at period end (monthly)/divided by total number of CAPA open at month end.	Production
% Deviations (Exceptions) closed in 30 days or less	The number of deviations (exceptions) which are closed in 30 days or less following the issue's discovery date/total number of deviations (exceptions) closed x 100.	Production
% Deviations (Exceptions) Open more than 45 days	The number of deviations /exceptions which are open more than 45 days following the issue's discovery date/total number of deviations (exceptions) open x 100.	Production
% OOS Investigations Open more than 45 days	The percentage of OOS reports which are open more than 45 days following the issue's discovery date.	Production
Changes Initiated	Total number of Process/Product/Equipment/Facility change requests opened during the period (monthly). Sites do not need to include procedure changes in count.	Production
Changes Closed	Total number of Process/Product/Equipment /Facility change requests closed during the period (monthly). Sites do not need to include procedure changes in count.	Production
Changes Open at Month end	Total number of Process/Product/Equipment/Facility change requests open during the period (monthly). Sites do not need to include procedure changes in count.	Production
% Employees with Overdue Training	% of employees who have one or more overdue training items at the end of the month. Calc: number of employees with overdue training/Total number of employees at the site.	All
Number of Sterilization Non-Conformances	Number of non-conformance events (deviations) related to sterilization processes (i.e. Autoclave cycles, EtO cycles, etc.)	Production
Cost of Poor Quality	The costs associated with events as a result of poor quality. Includes costs for investigations, reject, complaints, recalls, downgrading material, yield loss, etc.	Production and Post-Production
Batch Record RFT (right first time)	Number of batch records that have no documentation errors/the number of batch records that were reviewed and dispositioned during the review period x 100	Production
% Internal Audits completed to schedule	Number of Internal audits conducted in month (or audit schedule period)/number of scheduled for month (or audit schedule period)	All
Number of unscheduled Work Orders	Number of unscheduled maintenance work orders that were initiated during a review period (month, quarter). This includes work orders for GMP areas. It excludes office areas and other non-GMP areas. Unscheduled work orders do	Production

	not include Preventative Maintenance.	
Number of Expired open Temporary Change Controls	Number of open temporary change controls that have exceeded their approved temporary change window. This number is tracked monthly and includes events currently open at the end of the month. Expired change = change beyond the approved completion date. Approved extensions would extend the completion date. Temporary Change = Making a change for a period of time and then changing back vs Planned Deviation putting it in place for future	Production
Number of Open Change Control Action items open > 1 year	Number of open change control action items > 1 year old that are open at month end. This number is tracked monthly.	Production
% Damaged containers	Number of finished product containers that were damaged while being stored at the facility/the number of finished product containers produced during the time period.	Production
% lots rejected for Key Assay	Comparison between the numbers of confirmed failures for the key assay/number of lots dispositioned.	Production
% of procedures/test methods that are beyond their periodic review date)	Percentage of procedures, test methods and other documents that go through a periodic review cycle that have not completed their review at month end. This would compare the number of documents that are currently overdue vs. the total number of documents that go through a periodic review process.	Production
Validation schedule attainment	Number of scheduled validation final reports and revalidation assessments approved on schedule vs. total number of scheduled validation final reports and validation assessments.	Production
% QbD elements completed during API development	**% of the following activities completed (including risk assessments):** 1. Critical Material Attributes and relationship to material and product quality 2. Critical Process Parameters and relationship to API and drug product quality 3. Design, control, and knowledge spaces for key stages 4. Risk assessments performed and mitigation plans constructed 5. Control strategy established based on QbD findings.	Pre-Production
Risk ranking of Cross-contamination potential based on process developed	Have a pre-defined risk grid for process risk complexity. This ranking could trigger the need for additional process steps, facility requirements, cleaning validation, etc.	Pre-Production

% QbD elements completed during drug product development	**% of the following activities completed (including risk assessments):** 1. Critical Quality Attributes and relationship to material and product quality 2. Critical Process Parameters and relationship to API and drug product quality 3. Design, control and knowledge spaces for key stages 4. Risk assessments performed and mitigation plans constructed 5. Control strategy established based on QbD findings.	Pre-Production
Risk ranking of package functionality related to drug product protection.	Have a pre-defined risk grid for package criticality and shipping controls needed. This ranking could trigger the need for additional component testing, supplier controls, temperature studies, etc.	Pre-Production
Risk ranking of shipping controls needed.	Have a pre-defined risk grid for package criticality and shipping controls needed. This ranking could trigger the need for additional component testing, supplier controls, temperature studies, etc.	Pre-Production
Number of investigations related to method failures during validation and technology transfer	Measured as a lagging indicator during production phase, but can serve as a leading indicator for future method development in pre-production	Pre-Production
Number of method changes as a result of inadequate method development	Measured as a lagging indicator during production phase, but can serve as a leading indicator for future method development in pre-production (change is not due to a proactive enhancement, but rather due to inadequate method development)	Pre-Production
% of Suppliers from the entire product supply chain that are in the high risk category	Have a pre-defined risk grid for supplier risk ranking. This ranking could trigger the need for audit frequency, material qualification, elements in quality agreement, etc.	Pre-Production
% of Suppliers listed in the DMF/Regulatory Filing that have Business Continuity plans prior to Tech Transfer	Collect data that can enable correlation between product quality during commercial manufacturing. and supplier selection during development stages	Pre-Production
Number of investigations related to process failures during validation and technology transfer	Measured as a lagging indicator during production phase, but can serve as a leading indicator for future process development in pre-production	Pre-Production
Number of process changes as a result of inadequate process development	Measured as a lagging indicator during production phase, but can serve as a leading indicator for future process development in pre-production (change is not due to a proactive enhancement, but rather due to poor process	Pre-Production

	development)	
% of Products with On Time Regulatory Filings (e.g. NDA, ANDA, PAS)	% of Products with On Time Regulatory Filings	Pre-Production
% of Products with On Time Regulatory Approvals (e.g. NDA, ANDA, PAS)	% of Products with On Time Regulatory Approvals	Pre-Production
% of products with CMC-related delays (days) to first pass approval	% of products with CMC-related delays (days) to first pass approval	Pre-Production
Risk ranking of supply chain based on number of Manufacturing Locations Qualified by Product	Indicator of complexity of tech transfers to multiple locations and process variation introduced that may have an impact on product quality and safety on account of this decision	Pre-Production
Number of Stock-outs Attributed to Supplier Related Issues and/or Quality of CMC Development	Lagging indicator of insufficient/ineffective planning or development practices employed (assess within the first 2 years of Commercial Manufacturing)	Pre-Production

Appendix D: Final Proposed Metrics

A. Pre-Production Metrics:

Design Space	
Definition	Design Space (as defined by ICH Q8r) represents the "the multidimensional combination and interaction of input variables (e.g. material attributes) and process parameters that have been demonstrated to provide assurance of quality." The intent of capturing this metric will be to demonstrate the robustness of the development activities, specifically, the ability to properly characterize the products and help enable the principles behind Quality by Design (QbD): "A systematic approach to development that begins with predefined objectives and emphasizes product and process understanding based on sound science and quality risk management." – ICH Q8r
Clarifications	It was identified that product development does not widely include scientific justification and data to support the critical process parameters (CPP), critical material attributes (CMA) and critical quality attributes (CQA) that are proposed during product transfer. Development work has been found to include historical ranges used for other/similar products without experimental verification of the appropriateness for the product in question. In order to decrease risk of product failure and patient harm, these design elements need to be scientifically supported. The effectiveness of this program is measured by other metrics, such as RFT Production, RFT Transfer, QbD Lifecycle Effectiveness Project definition: project is "completed" within the scope of this metric when the product is ready for tech transfer
Formula	# of projects completed with scientifically justified predefined ranges[1] x100 Total # of projects completed [1] for CPP, CMA, and CQA

Supply Chain Assurance	
Definition	Number of Tier 1 suppliers approved through cross-functional approval to ensure internal alignment against all critical success factors
Clarifications	By measuring the effectiveness of a company's supplier base, companies can get a better understanding of the steps needed to improve Quality. For this metric, the team agreed to look at what it labeled as "Tier 1" suppliers:

	• API contract manufacturers • Excipient contract manufacturers • Primary packaging components The metric proposed is based on the assumption that product quality can be improved by putting in place a cross-functional review* of the proposed suppliers as part of its supplier qualification process. It is critical to ensure there is alignment of requirements across key functional groups (i.e. quality, cost, capability, capacity) in order to identify a supplier that best fits the needs of the product. *Cross-functional review: Includes key business functions and stakeholders, such as quality, supply chain, regulatory, marketing, legal, planning, procurement, finance, etc.
Formula	# of Tier 1 suppliers approved through cross-functional review x 100 Total # of Tier 1 suppliers in the supply chain for the product in question

B. Transfer Metrics

Process Validation Right First Time	
Definition	Percentage of process validation batches without deviations related to product and process development. • Deviation: any deviation from the process validation protocol related to product or process development; includes any incident that results in a failure of the ability of the product or process to meet protocol requirements and/or product specifications (such as critical process parameters, critical material attributes, process ranges, in-process controls) • Batch/Lot: As defined in the validation protocol
Clarifications	The Team concluded that by measuring the success of the process validation batch runs, companies could get a better sense of how well the product and processes had in fact been characterized and in turn, the level of quality that would be achieved once the product was released commercially. Metrics are to be captured by product.
Formula	# of process validation batches without product/process related deviations x 100 Total # of validation batches attempted

Analytical Method Transfer Right First Time	
Definition	Measure of the percentage of analytical methods transferred without analytical method development deviations. This includes planned and unplanned deviations. • Planned Deviation: any deviation from the proposed analytical methods (API or Drug Product) deemed necessary to provide a meaningful measure of product quality, process performance or stability of commercial product. (Note – will count deviations that were due to inadequacies of the methods) • Unplanned Deviation: only includes unplanned deviations that were determined to be related to inadequate method development. Therefore, it does not include unplanned deviations that were due to inadequate laboratory execution (i.e., equipment failures, out-of-calibration equipment, human error not related to inadequate method instructions, etc.)
Clarifications	This metric was proposed as one that will help organizations assess the rigor of the analytical method development process for the current and future products. Often, analytical methods are improved during the transfer process without assessing the rigor of analytical method development across the board. This metric will help highlight the success of Right First Time in analytical method development, which can trigger continual improvement that reduces opportunity for analytical error.
Formula	# of analytical methods transferred with no method related deviations x 100 Total # of transfer attempts per product

C. Production Metrics

Right First Time during Production	
Definition	A measure of the percentage of batches without potentially product impacting deviations, investigations, out of specification results, or unplanned rework or rejections. Recommended to calculate by product and by site. Deviation: Any incident that would result in the following: • Negative impact on product and/or results in stopping production or testing, or results in quarantine of any portion of the batch (which includes in-process and finished product) • Any incident that delays the disposition of the batch due to a potential quality issue • Deviations that need to be investigated to confirm the impact on the product

	• "Minor" deviations (for example, any departure from an approved instruction or established standard/specification) that may have an impact on the product Batch/Lot: a specific quantity of material produced in a process or series of processes so that it is expected to be homogeneous within specified limits. Note – Team agreed that minor documentation errors would not be included unless they impacted the batch/lot as described above.
Clarifications	The metric is intended to offer another data point to help determine the robustness of a product and/or its process. A company finding itself with a low "Right First Time" rate might need to work to identify improvements across its Quality System. In addition, a low first time rate over time might indicate that the overall product development efforts might not have been robust and as a result, yielded a product that was not able to meet specifications in a consistent and regular manner. • This metric needs to be a snapshot in time. It is not intended to go back and re-calculate the RFT numbers based on information learned post-disposition. • It does not include planned deviations. • It is critical that a trend analysis is conducted for this metric to assess root cause for continuous improvement. • This metric is to be reviewed at the corporate level to foster enterprise-wide continuous improvement.
Formula	# of batches/lots without deviations x 100 Total # of batches/lots attempted

CAPA Effectiveness	
Definition	A measure of whether actions taken as a result of problems/issues encountered have effectively addressed the deficiency and prevented their recurrence. The CAPA must have an effectiveness check to be counted towards the metric.
Clarifications	The Team wanted to understand the ability of companies to close out CAPAs in a complete manner – one that led to effective (sustainable) solutions. The Team felt that by looking at this metric, a company could help confirm whether improvements were being realized as a result of the corrective and/or preventive actions put in place. An positive CAPA effectiveness measurement for example might yield a lower number of deviations, repeat issues, etc.
Formula	# of successful effectiveness checks x 100

	Total # of effectiveness checks attempted
Notes	Since CAPA effectiveness checks can occur over a long period of time, the metric is to be measured at a given time period. For example, if the cadence for reporting the metric is every quarter, then the metric would count only the effectiveness checks that ended in the most recent quarter. This prevents having a large denominator if there are multiple effectiveness checks that are running concurrently. In addition, if an originally successful CAPA effectiveness check is later deemed to have failed, then the failure would count in the most recent quarter. This highlights a present issue instead of fixing old data.

Commitment Index	
Definition	A score which measures the commitment of a company/site to a culture of quality through capture of performance related to the on-time completion of requirements associated with regulatory/industry expectations.
Clarifications	The index includes the assessment of a core group of metrics that can be modified (including the timeframes) or removed per company/site if that activity does not apply to that company/site: • Investigations – The number closed in 30 days or less per month vs. total number closed per month, times 100. • Customer Complaints – The number closed in 45 days or less per month vs. total number closed per month, times 100. • CAPA – the number of corrective actions completed on time per month vs. total number of corrective actions due per month, times 100. • APR – The number of APRs approved within 30 days of annual establishment due date per month vs. the number of APRs approved during the month, times 100. • Stability Testing — Number of samples pulled and tested by due date per protocol during the month vs. total number of stability samples pulled and tested during the month, times 100. Includes all stability testing — annual, validation, etc. • GMP Training – The number of employee training assignments completed on time per month vs. the total number of training assignments completed per month, times 100. (Alternately, we look at overdue training at month end.) • Audits – Number of Internal, Supplier, CMO, CRO and distribution audits completed as scheduled per month vs. the total audits scheduled per month, times 100. Ad hoc audits are excluded from this number. • PM/Calibration – Number of PMs and calibrations completed on or before the originally scheduled due date per month vs. the number of PMs and calibrations scheduled to be completed per month, times 100. • Regulatory Commitments – Number of commitments completed on or before the original commitment due date per month vs. the total number

	of commitments completed per month, times 100. This number is not to exceed 100. A commitment that is completed early will be considered 1/1. Commitments here are related to regulatory observations (not commitments made as part of a filing.) • Revalidation – Number of revalidations completed on or before the revalidation due date per month vs. the total number of revalidations completed per month, times 100. This number is not to exceed 100. Includes a review of the qualification status based on historical review of equipment performance.
Formula	See Clarification for definition of each term. (Investigations × 0.2) + (Customer Complaints × 0.2) + (CAPA × 0.1) + (APR × 0.1) + (Stability × 0.1) + (Training × 0.05) + (Audits × 0.05) + (PM × 0.05) + (Reg. Commitments × 0.1) + (Revalidations × 0.05)
Notes	This metric has flexibility on how to weight each term and which terms to include in the index to allow for flexibility in a way that makes sense for the products and business of each company.

Supplier Risk Index	
Definition	An assessment of supplier risk based on qualitative and quantitative factors, such as level of concern related to performance, audit findings, geographical risk, leverage, capacity, and status of necessary agreements.
Clarifications	Includes existing Tier 1 suppliers, such as: • API • Excipients • Primary packaging components • Contractors (manufacturing, laboratory, packaging, logistics)
Formula	All will be measured using a scale: 0, 5, 10 (where 10 is good): A. Level of confidence relative to performance of supplier, as measured by complaints related to the supply in a given time period based on the number of lots received. B. Level of confidence relative to audit/regulatory findings in a given time period (if no audit in given time period, then previous results apply). C. Necessary agreements (i.e., supply agreement, quality agreement) are in place D. The supplier has sufficient capacity and/or redundancy such that risk of a shortage is lowered. E. Level of confidence relative to geographical risk (e.g. under-regulated regions of the world). F. Level of confidence related to leverage and supply stability — assessment of the % of supplier's bottom line attributed to our business.

	G. Level of confidence in track record of the supplier (previous materials supplied) Formula: A + B + C + D + E + F + G ≤ 70 Suggested Actions based on Score: • 60 – 70: No action required, assuming all responses are 5 or higher • 40 – 55: Cross-functional assessment of mitigation strategies, as well as meetings with suppliers to identify improvement opportunities • 20 – 35: Cross-functional escalation of risk awareness, assessment of supplier alternatives and mitigation strategies, heightened involvement in supplier operations and oversight. • 0-15: Cross-functional escalation of risk mitigation requirements, identification of alternate source of supply, integral involvement with supplier operations and oversight.

D. Post-Market Metrics

Market Reliability Index	
Definition	An overall product confidence score established on a roll-up of the post-market surveillance data including: complaints or unexpected trends that triggered action (such as field alerts, corrective actions, changes), adverse events (not included in product labeling, or triggered a safety signal), stability failures, drug shortages, field alerts, and recalls.
Clarifications	The index includes the assessment of a core group of metrics that can be modified (including the timeframes) or removed per company/site if that activity does not apply to that company/site: • Number complaints or unexpected trends that triggered action per month vs. the total number of units released per month, times 100. • Number adverse events (not included in product labeling, or triggered a safety signal) per month vs. the total number of units released per month, times 100 • Drug shortage: Number of days a product is on back-order/365 days, times 100. Assess each month. (backordered: product is not available to fill PO to wholesalers/pharmacy) • Number of batches with field alerts per month vs. the number of batches released per month, times 100. • Number of batches with recalls per month vs. the number of batches released per month, times 100.
Formula	(100 - % Customer Complaints) × 0.15 + (100 - % Adverse Events) × 0.15 + (100 - %

	Drug Shortages) × 0.30 + (100 - % field alerts) × 0.20 + (100 - % Recalls (will intentionally include those issues already captured in field alerts)) × 0.20
Notes	This metric takes into account double counting of certain items. For example, an Adverse Event will have a complaint associated with it. By allowing Adverse Events to be double counted, it places more importance and weighting on Adverse Events, essentially weighting it more than the proposed 15%. The same applies to Recalls with regards to Field Alerts, Customer Complaints and Adverse Events. While it has similar weighting as the previous indicators, by double counting, Recalls is effectively weighted more than the proposed 20%.

E. Enterprise-Wide Continual Improvement Metrics

Right First Time Rate for Production	
Definition	This metric is to be taken directly from the production metric.
Clarifications	See "Right First Time during Production"
Formula	$\frac{\text{\# of batches/lots without deviations}}{\text{Total \# of batches/lots attempted}} \times 100$
Notes	An assessment of the root causes and trends is critical for enterprise-wide learning, and is to feed back into R&D and Production as appropriate with triggers for decisions.

Quality by Design Lifecycle Effectiveness	
Definition	An assessment of all failures related to product, process and supply chain that are attributed to development and transfer. Examples include: changes, complaints, FARs, recalls, poor Cpk, poor yield, stability failures, inadequate material characterization, and product failures. This metric is an index score that is weighted based on failure criticality.
Clarifications	• Number of confirmed, critical customer complaints and adverse events related to failures associated with the product design per month vs. the total number of units released per month, times 100. • The Cpk of the process. (1.33 - actual process capability)/1.33, times 100. If actual is greater than 1.33, then the CpK factor is reported as 0 in the calculation • Number of lots implicated with confirmed stability failures related to inadequate product design per month vs. the total number of lots released per month, times 100 • Number of confirmed product failures (OOS/OOT) related to inadequate product design per month vs. the total number of lots dispositioned per month,

	times 100
Formula	Sum the values of the % score for each follow through score times its weight factor. Maximum score for the index would be 100 if 100% effective. (100 - % Customer Complaints)* 0.25 + (100 - % Process Capability) * 0.25 + (100 - % Stability Failures)* 0.25 + (100 - % Product Failures)* 0.25
Notes	An assessment of the root causes and trends is critical for enterprise-wide learning, and is to feed back into R&D and Production as appropriate with triggers for action.