GOLDILOCKS FOR PROCESS VALIDATION STATISTICS

TARA SCHERDER

SOCAL PDA VALIDATION SYMPOSIUM

OCT 11, 2018

SynoloStats

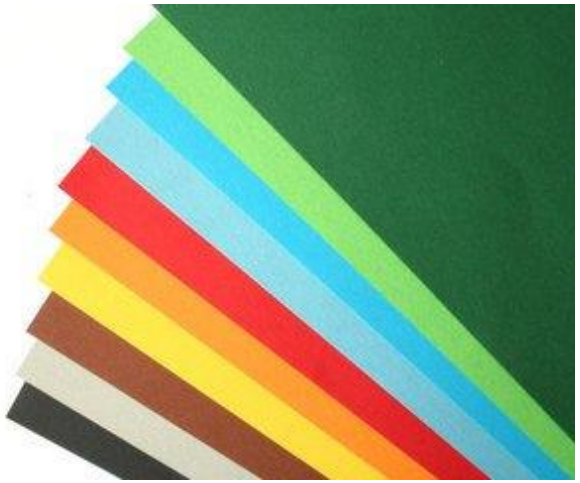
INTEGRATING TOTAL SOLUTIONS

CONCEPTS

- Goldilocks Principle
- Real Goals
- Statistical Decision Making and Lean Thinking
- Our Context
- Statistical Concepts by Example
- What's Best for the Patient?

THE STORY OF GOLDILOCKS AND PROCESS VALIDATION STATISTICS

Too little



JUST RIGHT!

Too much

THEN AND NOW

The Past

- Limited Use of Designed Experiments
- 3 Validation Batches, release testing only
- Annual Review (often done by non process SME)

Post ICH Q8, Q9, Q10, FDA Guidance on Process Validation, EMA Annex 15

- Designed Experiments
- Statistically designed sampling plans within lot and for number of lots
- Acceptance criteria that requires statistical confidence
- Review by Process SME
- Over concentration on statistical signals, normality, and trending
- Statistics for statistics sake

STATISTICS FOR DECISION MAKING AND LEAN THINKING

- Deliver what the patient needs
  - Right medicine (infers right quality)
  - Right time
  - Right cost
- Deliver what the business needs
  - Essentially enable delivery of what the patient needs!
  - Use resources effectively
  - Be compliant
  - Desirable employee environment

MANAGE RISK FROM PROCESS VARIABILITY AND SAMPLING

Variability

Sampling

Science and Statistics

Risk

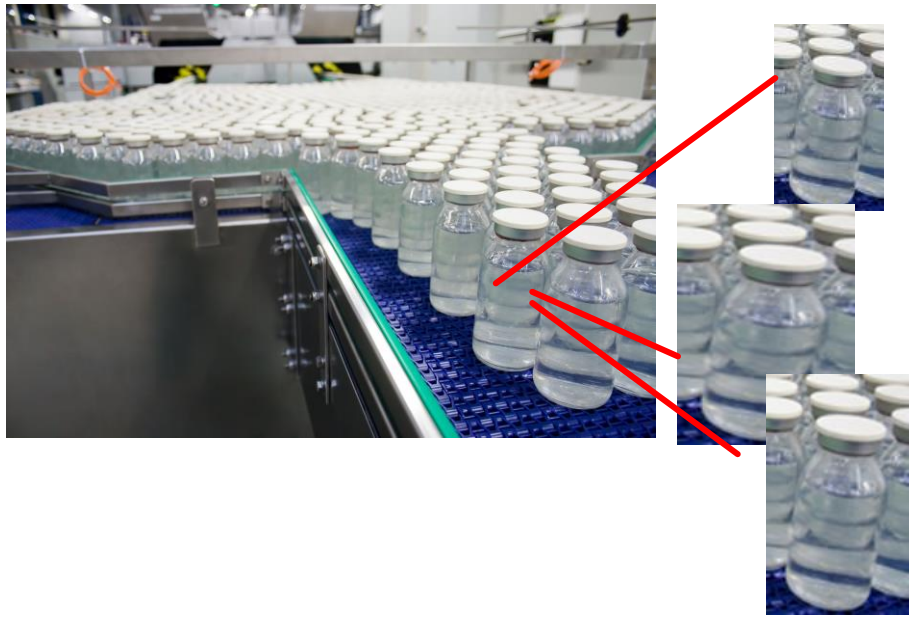
ASSURING EVERY DOSAGE UNIT MEETS QUALITY SPECIFICATIONS

2011 PV Guidance:

"Before any batch from the process is commercially distributed for use by consumers, a manufacturer should have gained **a high degree of assurance** in the performance of the manufacturing process … **Information and data** should **demonstrate** that the commercial manufacturing process is **capable** of **consistently** producing acceptable quality products within commercial manufacturing conditions."

ASSURING EVERY DOSAGE UNIT MEETS QUALITY SPECIFICATIONS

- Statistical tools *across the product lifecycle*, combined with process knowledge, provide assurance
  - DOE, predictive modeling
  - Statistical intervals
  - Sampling plans & acceptance criteria
  - Capability, Control Charts

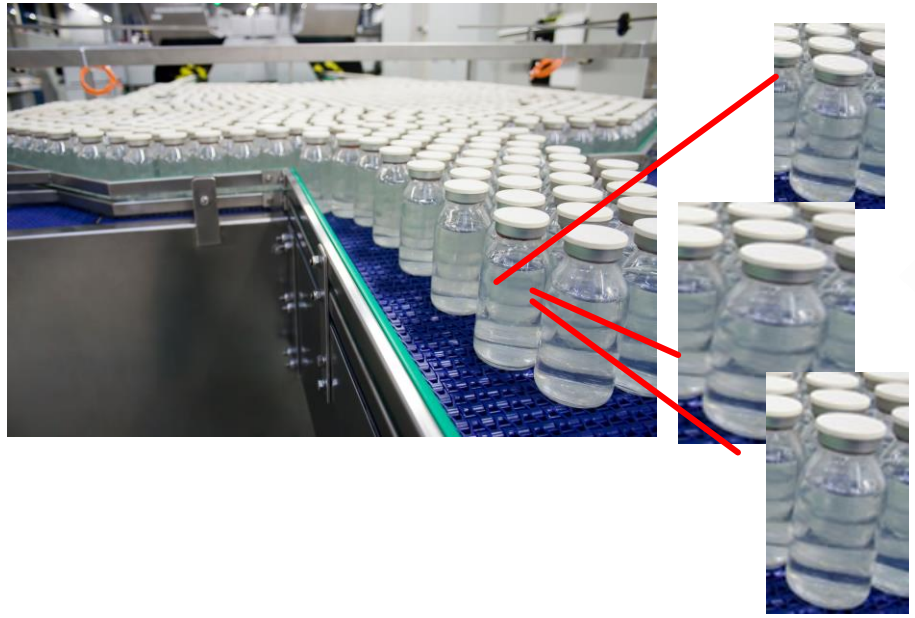


STATISTICAL ACTIVITIES MUST BE VALUE-ADDED

- T – Transport
- I – Inventory
- M – Movement
- W – Waiting and delays
- O – Overproduction
- O – Overprocessing
- D – Defects

WASTE IS TRANSFERRED TO PATIENT IN COST, IN DELAYS, MISDIRECTED RESOURCES, etc

CONCEPT: STATISTICAL UNCERTAINTY

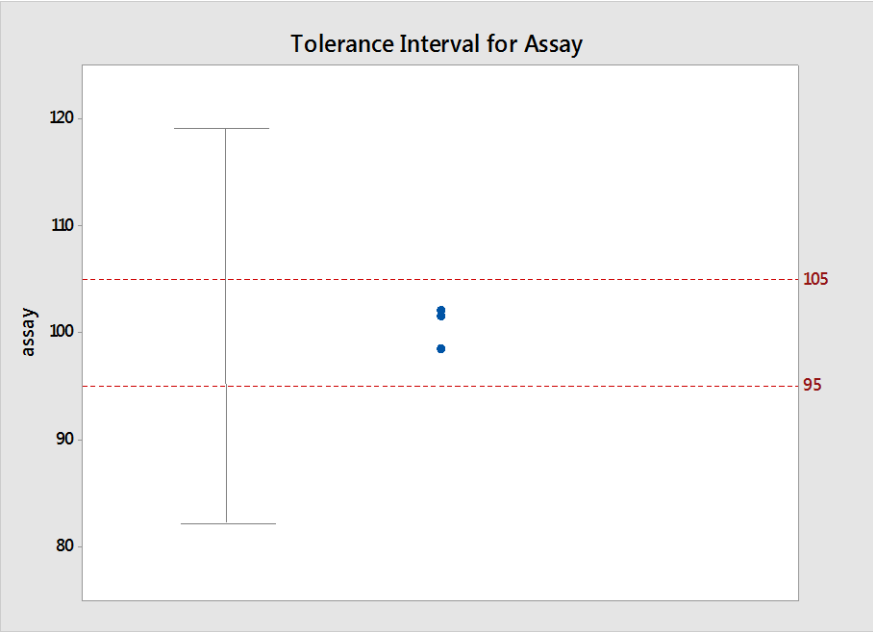


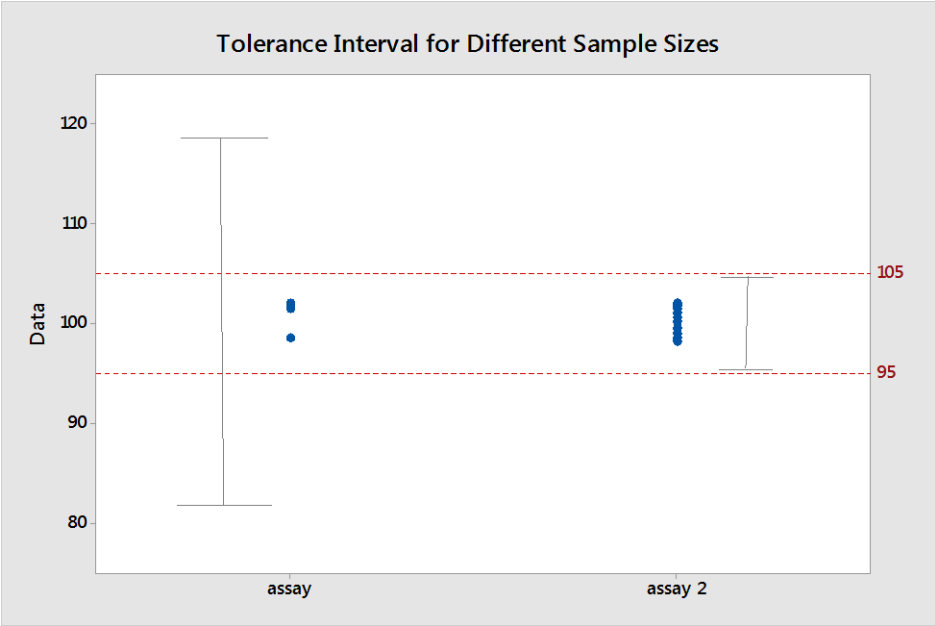
We use statistics to assure that the "next" tablet, syringe, cream, patch…..

meets patient needs

STATISTICAL UNCERTAINTY IN THE CLAIM

But you don't want to be wrong when you make the claim. Specifically, claim that most bottles will meet specification when they actually don't





STATISTICAL INTERVALS AS ACCEPTANCE CRITERIA

Tolerance Interval

Ex: We are 95% confident that 95% of units will be between 92.6-107.5%

$$\bar{X} \pm k\, s$$

Where $\bar{X}$ is the sample average

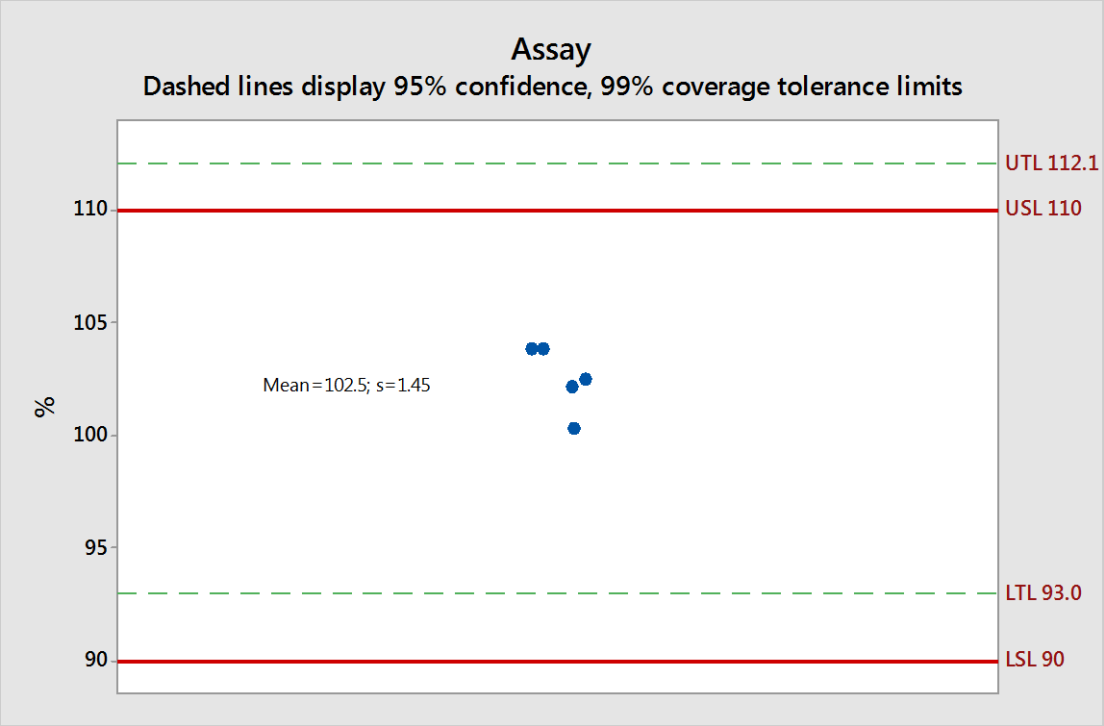
k is a factor that depends on sample size, confidence level and coverage

s is the sample standard deviation

From Development Data Expected Mean = 100.0; Expected Standard Deviation = 1.0; Confidence = 95%; Coverage = 99%	
n	Tolerance Interval Assay
3	87.4 – 112.6%
4	91.8 – 108.2%
5	93.4 – 106.6%

Development data → 5 intra-batch samples

PPQ RESULTS

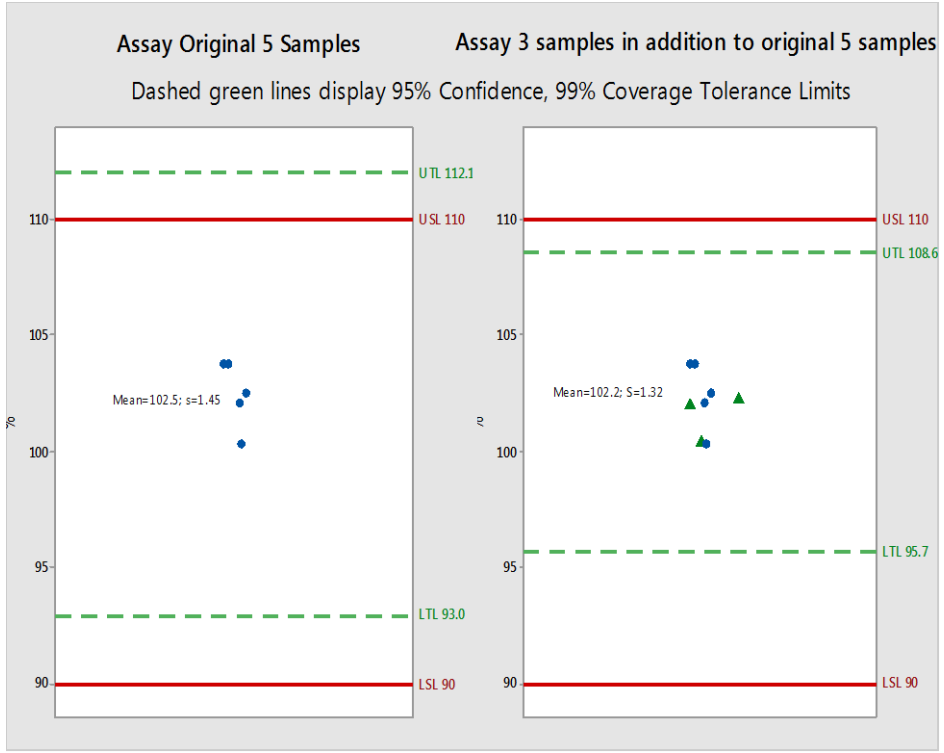


Slightly higher mean and standard deviation than expected. All results well within specification, but UTL beyond specification

Should the batch and possibly the entire PPQ fail?

What's best for the patient?

INFLUENCE OF SAMPLE SIZE



With an additional three samples the interval falls within specification.

<u>Performance is the same; additional samples provide more information,</u> decreasing uncertainty in the mean and standard deviation (and hence k). This results in a ***more accurate interval.***

The batch should not fail because of the limitations of a statistical interval.

The answer is not simply over-sampling to avoid this situation at all costs. Over-sampling is waste, and in some cases very costly

PPQ protocol can include a risk based approach to evaluate an interval outside of specification which would include the influence of sample size, outliers, etc. Staged sampling and Stage 3A enhanced sampling are also options.

PPQ NUMBER OF BATCHES

"….The number of samples should be adequate to provide sufficient statistical confidence of quality both within a batch and between batches…"

- Requiring Statistical Confidence
  - Necessitates a large number of batches
  - Is typically a complex, situationally unique, error prone computation

What Is Best for the Patient?

PPQ NUMBER OF BATCHES – PATIENT FOCUS

- Want to avoid delays
- Want to minimize costs (from additional samples and activities, etc.)

Without introducing risk to quality

- Number of batches depends on results of risk analysis evaluating process knowledge/control strategy
- Intra-batch sampling
- Other Inter-batch statistical analyses; variance components, graphs confirm variability is as expected

Remember, this is a continuum

- Should have substantial information due to the use of more evidence in process design stage, and continue collection of evidence during stage 3b

- The evidence of inter-batch control has to support claims of reproducible commercial manufacture, but should not need to be strict statistical claim as in the conclusion of a clinical trial

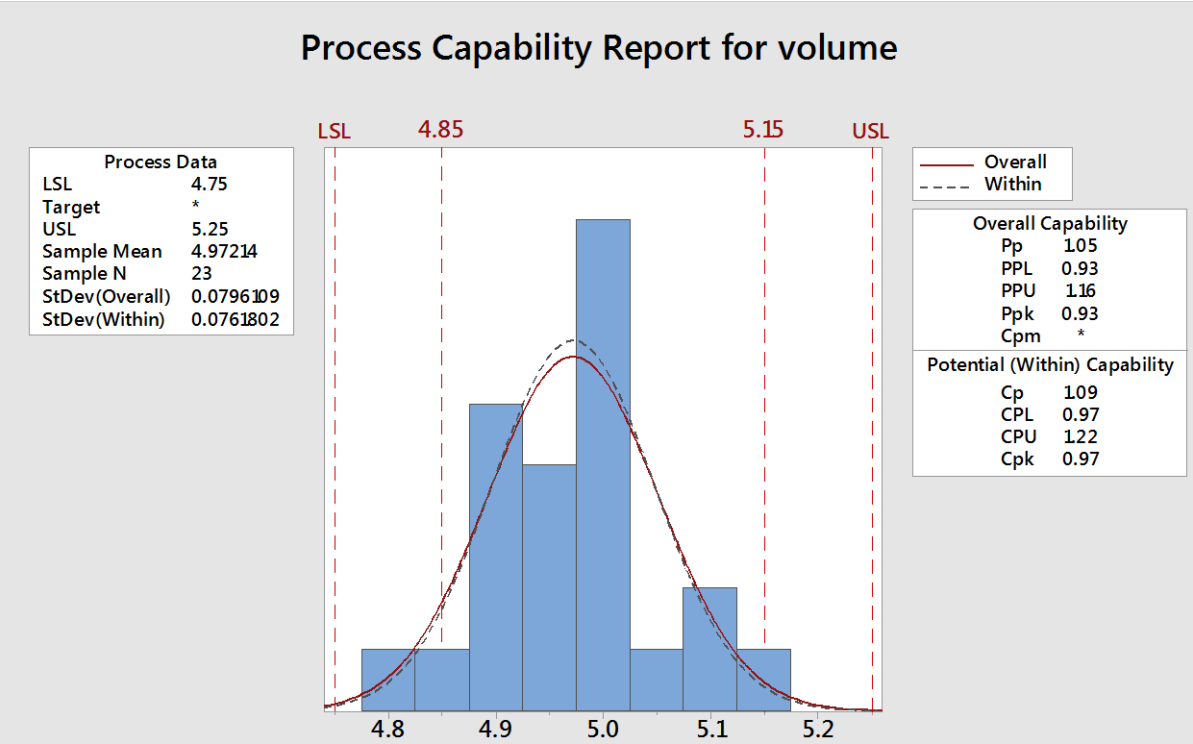
CONCEPT: I.I.D

"Whenever you fit a model to your data you are assuming that those data are homogeneous. If they are not homogeneous, all of your statistics, all of your models, and all of your predictions are going to be wrong" [1]

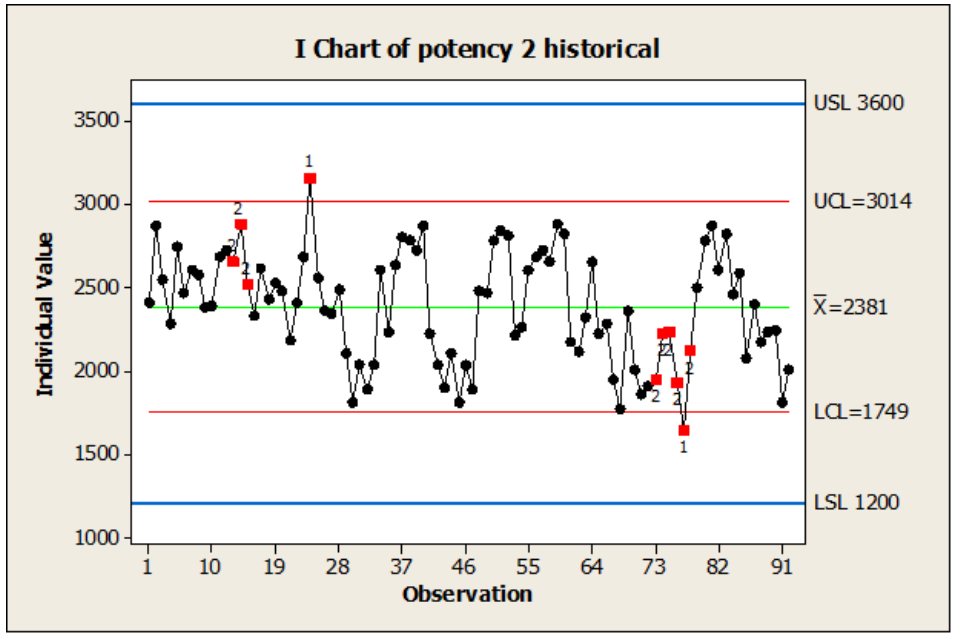
- Random samples are **Independent** and **I**dentically **D**istributed
  - Requirement for typical statistical tests (e.g., hypothesis, regression, control charts)
- Independent: successive observations are not related to each other
- Identically distributed: drawn from a single distribution (single mean and variance)
- Neither are likely true in pharmaceutical manufacture because common sources of variability are not used randomly

[1] Donald Wheeler, Quality Digest, 30 July 2012

IPC



CONTROL CHARTS



- In typical applications of SPC (Phase 2 SPC), observations in a Shewhart control chart are assumed to be I.I.D

They will not be….

- Common sources of variability are not used randomly, resulting in multiple sub-populations
- Can’t interpret statistical signals as if observations were i.i.d.
- Statistical signals are actually rich in information. Need a business process to evaluate appropriately

CONCEPT: THIS IS NOT REAL-TIME SPC

- Woodall [1] describe two phases of SPC. CPV most often aligns with Stage 1; data is viewed retrospectively, large sources of variability still exist, and immediate action to adjust the process is not sought
- A state of control is not synonymous with statistical control. Statistical control, exhibited by the absence of statistical signals, should not be expected
- Identifying sources of variability is more important than reducing so-called false alarms, or increasing detectability of small changes
- There is no regulatory requirement to initiate an investigation for statistical signals

  "..Not all signals are created equally. …Magnitude of reaction depends on the severity of the signal…" [2]

- If red dots are always the enemy, change your mindset and the business process

[2] Alex Viehmann, FDA/CDER/OPQ, ISPE PV Statistician Forum April 2015

ABOUT THE VARIATION….

2011 PV Guidance:

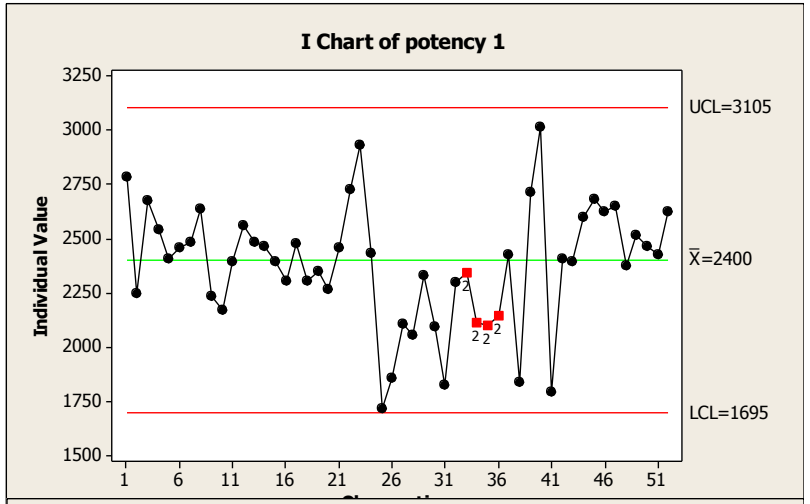
"Manufacturers should:

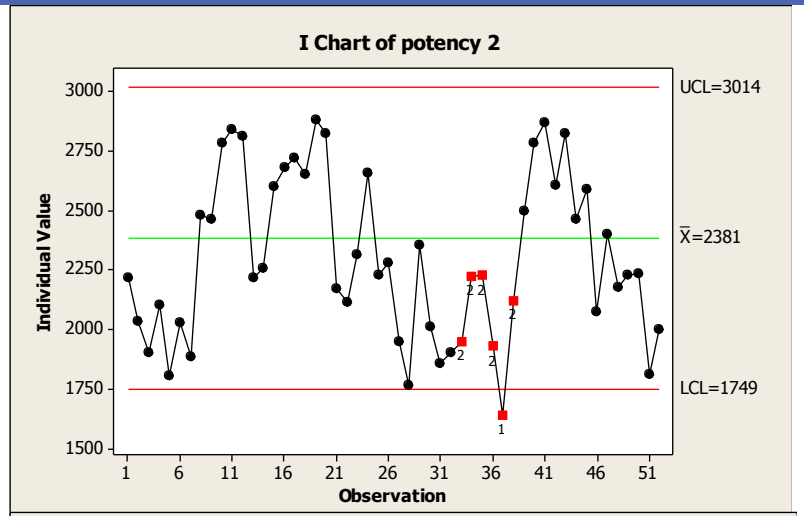
- Understand the sources of variation
- Detect the presence and degree of variation
- Understand the impact of variation on the process and ultimately on product attributes
- Control the variation in a manner commensurate with the risk it represents to the process and product"

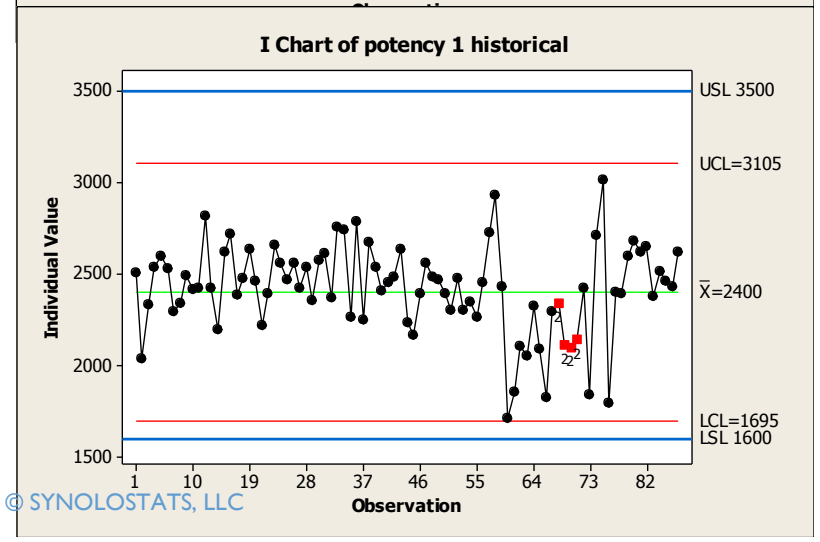
OVERPROCESSING AND DEFECTS IN MONITOR/TREND

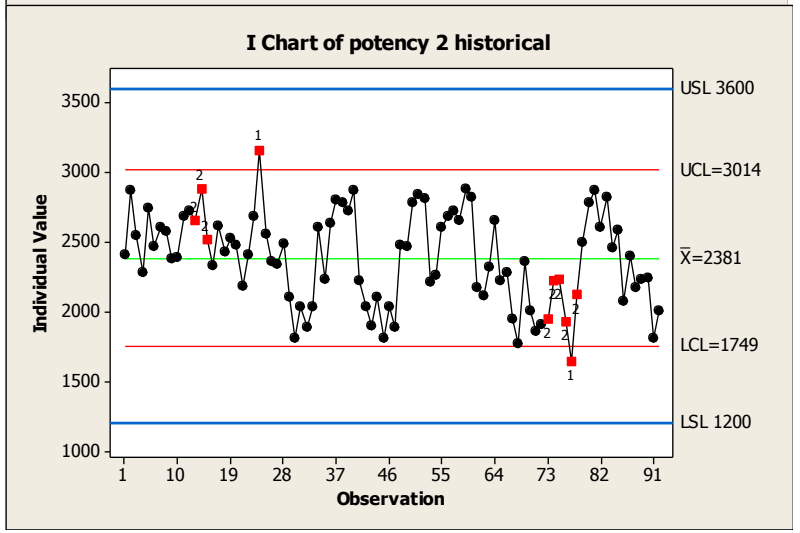
- Too many parameters (choose using a risk based approach)
- Evaluating all processes equally instead of on a risk basis (one size fits all approach)
  - Highly capable performance does not need the same review as lower performance
  - Note however, too seldom can lead to defect waste
- Too many charts (create meaningful charts)
- Too many people "touching" the same data
- Too few data points to establish control limits or process capability
- Too often updating limits, capability, distribution – control chart tells the story

OVERPROCESSING IN TAKE ACTION





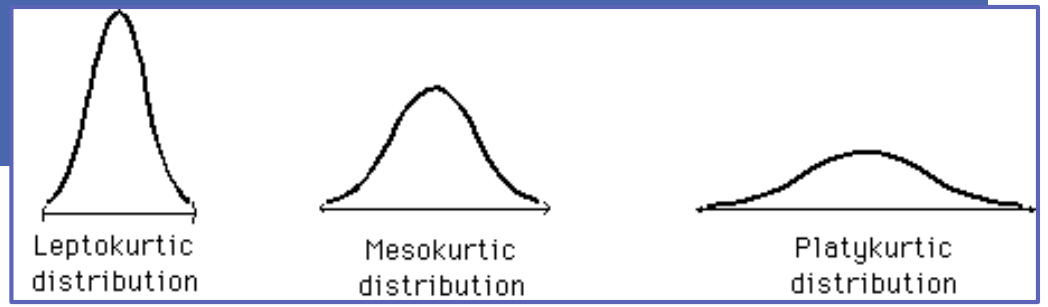




Consider:

- Attribute Severity
- Capability
- Magnitude and duration of excursion
- Historical behavior
- Process and measurement knowledge
- Can develop decision tree or matrix

OVERPROCESSING IN EVALUATE NORMALITY



leptokurtophobia - an irrational fear of using non-normal data in your analysis

Acknowledge the non-normality, and assess the practical effect on interpretation and use.

Transform only when warranted by an underlying distribution that is non-normal due to physical/chemical/biological reasons.

"Whenever you fit a model to your data you are assuming that those data are homogeneous. If they are not homogeneous, all of your statistics, all of your models, and all of your predictions are going to be wrong" (1)

(1) Donald Wheeler, Quality Digest, 30 July 2012

OVERPROCESSING AND TRANSPORTATION IN REPORT

- Too frequent reporting
  - Can choose on volume, events (e.g. campaign), time
  - Attributes with lower capability warrant more frequent evaluation and reporting
  - As capability increases, reporting can decrease
  - Consider alerts for even highly capable processes
- Consider other reporting mechanisms; e.g., product review, integrated team meetings
- Align organizations that are reporting (lab, CPV, APR)
- Better for SME to look more frequently at CQAs, than less frequently with more parameters and extraneous information

OVERPROCESSING AND TRANSPORTATION IN REPORT

- Optimize a format/template
  - Streamline!
  - Every graph does not need to be explicitly discussed. If all parameters are "in control", discuss only the ones worth noting.
  - State of control needs reference to historical behavior; current data should be compared to previous behavior
  - Every statistical exception does not need to be repeated (for example your treatment of non-normality)
  - Don't recompute metrics unless appropriate, e.g, process capability indices (stable after sufficient sample size of 60-90). If the control chart looks "good", that's sufficient!

SAMPLING PLANS

- Sampling plans support statistical performance evaluations
  - Statistical intervals (PPQ)
  - Heightened sampling/monitoring (Initial CPV)
- Optimal sampling plans
  - Reduce risk of :
    - Weak claims of process capability
    - Failure to understand major sources of variability
  - Consider the CQA and how it is influenced by the process and control strategy
  - Are representative of the expected sources of variability
  - Consider lifecycle stage and why sampling is being done
  - Minimum sample size to meet acceptance criterion if performance is acceptable & as expected
  - Based on historical performance, process knowledge

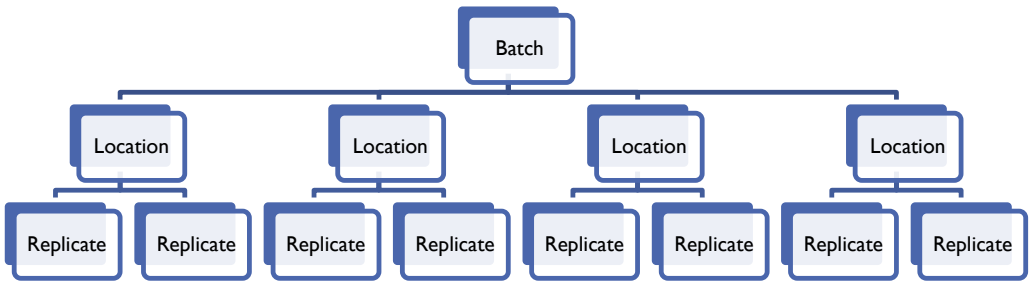
SAMPLING PLANS

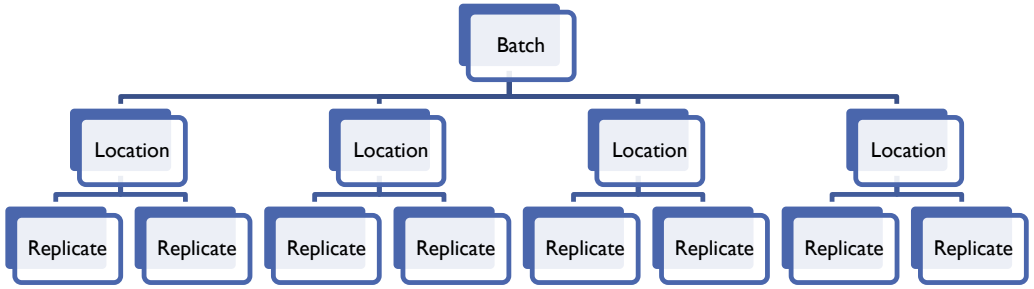
- Stage 2 (PPQ) Planning
  - Target level of statistical assurance and type of statistical interval
  - Use development (and historical) performance data to estimate expected PPQ performance
  - Choose sample size so that statistical interval will be within acceptance criteria if performance is as expected
- Statistical intervals provide assurance by describing performance with a measure of uncertainty

  Useful as part of PPQ performance *evaluation*
- *Because* they incorporate uncertainty, they are quite conservative

  Risky as stringent PPQ acceptance *criteria*

VARIANCE COMPONENTS



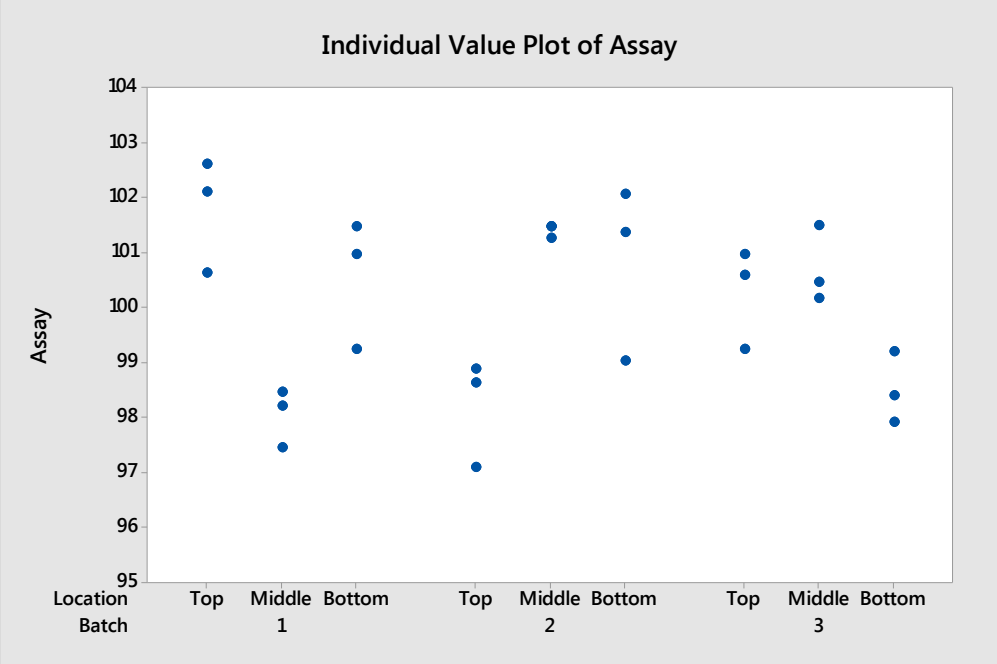


Compute amount of variability contributed by each source

Batch to Batch
Location to Location
Sampling and Measurement

Examples:
Bottle filling using multiple fill nozzles and torque heads
API material separated into multiple drums
Vessel separated into top, middle and bottom
Tablets across multiple time locations and two press sides

RELATIVE CONTRIBUTIONS OF SOURCES OF VARIABILITY



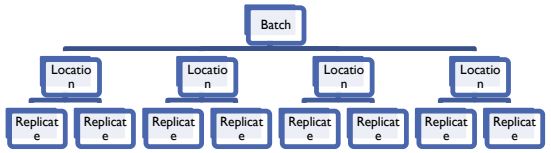
3 Batches
3 locations per batch
3 samples per location

Variance Components, using Adjusted SS

Source	Variance	% of Total	StDev	% of Total
Batch	-0.848198*	0.00%	0.00000	0.00%
Location(Batch)	2.34570	72.79%	1.53157	85.32%
Error	0.876744	27.21%	0.93635	52.16%
Total	3.22244		1.79512	

** Value is negative, and is estimated by zero.*

RELATIVE CONTRIBUTIONS OF SOURCES OF VARIABILITY



3 Batches
4 containers per batch
2 samples per container

Nested ANOVA: Particle Size vs batch, container

Analysis of Variance for D10

Source	DF	SS	MS	F	P
Batch	2	5.3988	2.6994	8.061	0.010
container	9	3.0140	0.3349	1.747	0.182
Error	12	2.3002	0.1917		
Total	23	10.7130			

Variance Components

Source	Var Comp.	% of Total	StDev
Batch	0.296	52.89	0.544
container	0.072	12.81	0.268
Error	0.192	34.30	0.438
Total	0.559		0.748

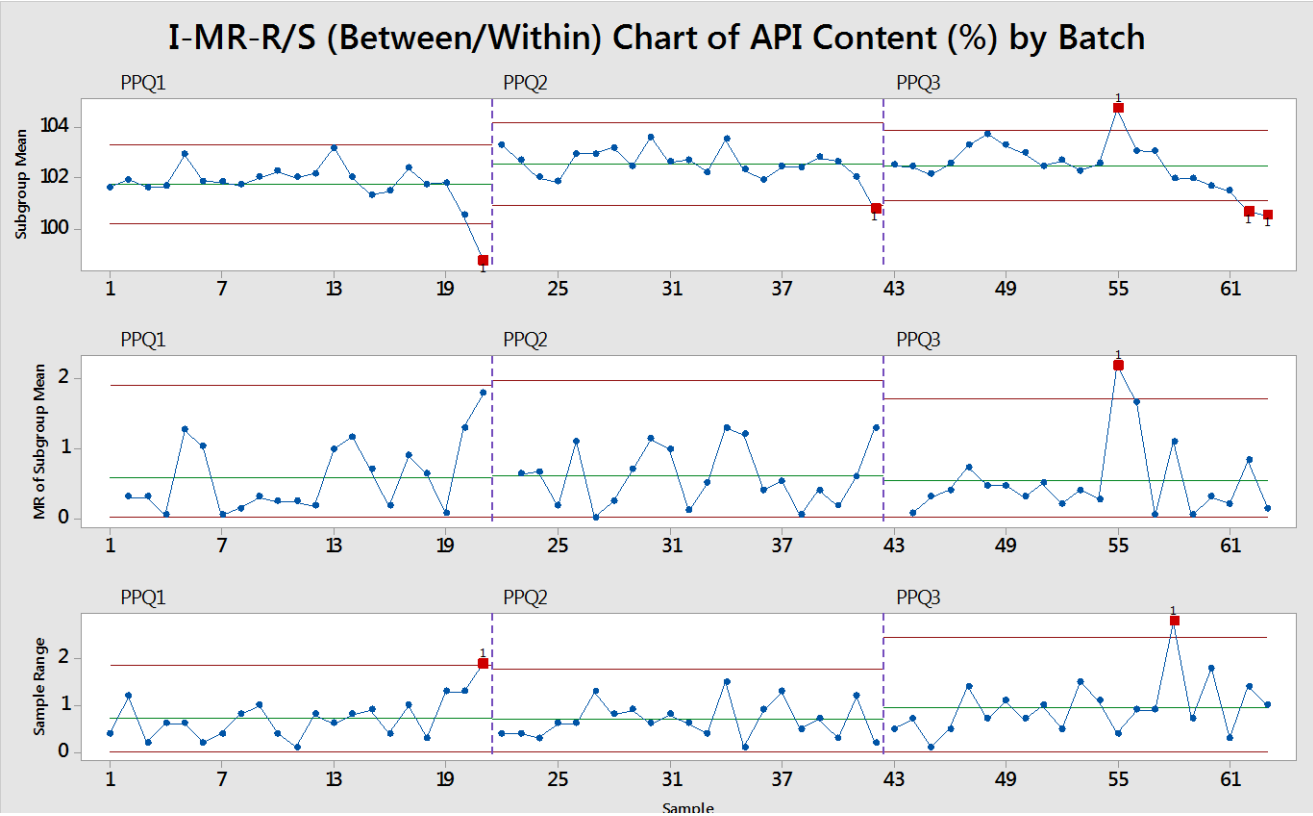
- Are the batches homogenous?
- Is the batch to batch contribution too high?

Don't apply strict acceptance criteria to %s. Use to support relevant claims of assurance of quality

ENHANCED SAMPLING IN CPV

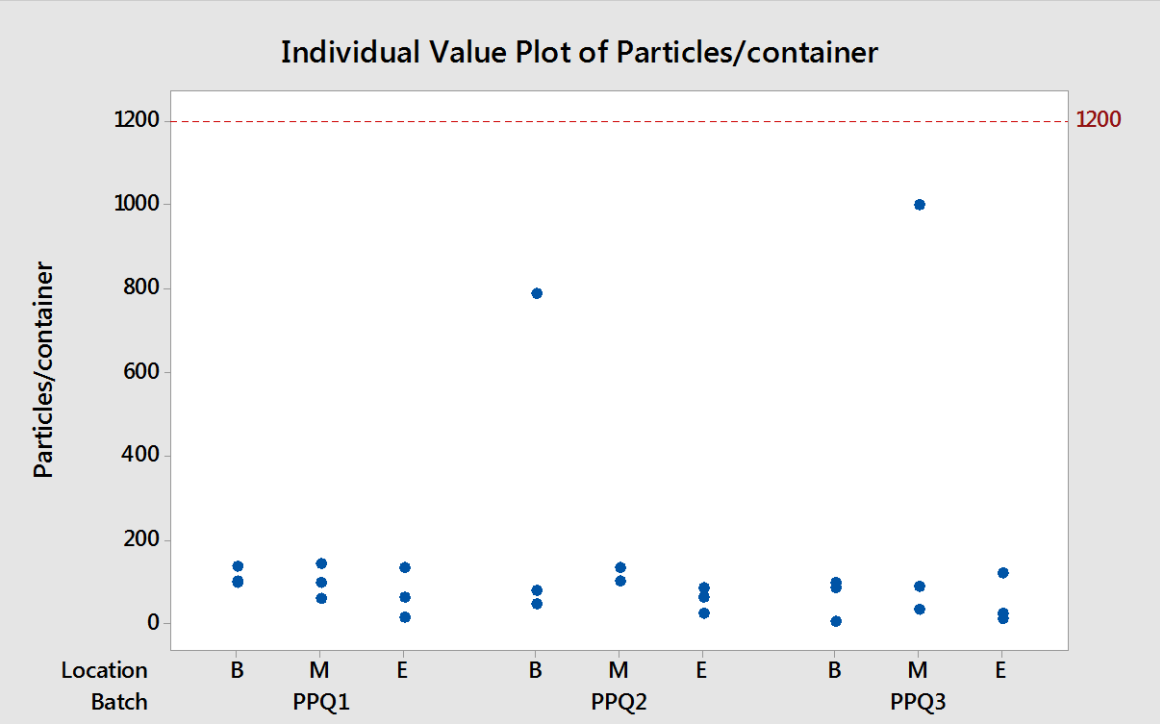
- Stage 3 (CPV) Enhanced Sampling
  - Usually not used to calculate statistical intervals (but may be)
  - Focus heightened sampling to address variability/risk identified in PPQ
    - Not same for all CQAs/CPPs, all processes
  - More process understanding ⇒ less sampling

TARGET ENHANCED SAMPLING TO ADDRESS UNCERTAINTY/RISK



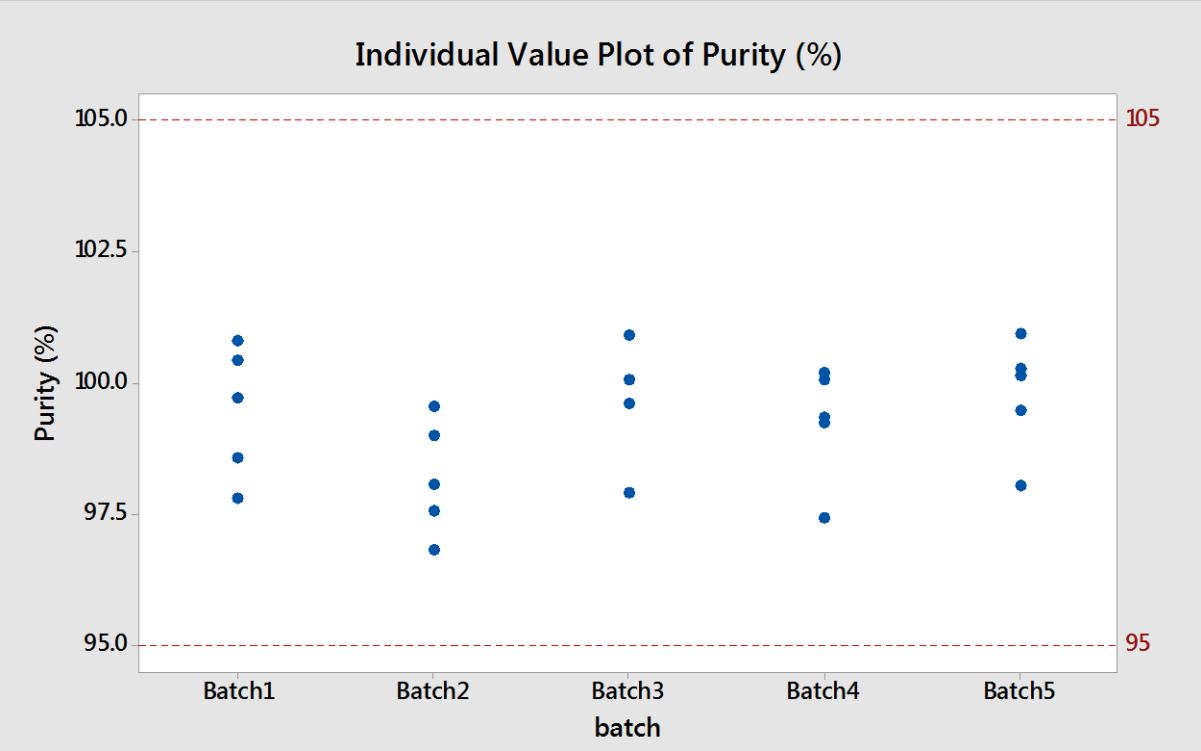
- API Content (%) drops at the end of each PPQ batch
- Is it expected?
- CU easily meets release requirements and provides statistical confidence that future samples will also meet release requirements
- Enhanced sampling? If so, what sampling plan, and for how long? Why?

SOMETIMES COMPLEX STATISTICAL ANALYSES ARE MORE COMPLICATED THAN NEEDED TO MAKE A DECISION



- Tolerance intervals? Variance components analyses?
  - Multiple sources of variability (between batch, between locations within a batch, within location) to account for – TI difficult to calculate
  - Most values are very far below USL of NMT 1200, with a couple of higher individual values
  - Hard to assess distribution to model performance – **may not be worth it** (yet, if at all)
  - What would you do?

SOMETIMES COMPLEX STATISTICAL ANALYSES PROVIDE KEY INFORMATION REGARDING RISK



- No values outside specifications, but high intra-batch variability with respect to specifications of [95.0 -105.0%]
- All *individual batch* 95/99 TIs are outside of specifications
- Enhanced sampling?
- This test is known to be highly variable and process SMEs believe *process* risk is low
- In a case like this, it may be worthwhile to use more complicated statistical methods to calculate a TI or probability of out of specification results incorporating within & between batch variability to assist with the risk assessment

HOW LONG TO CONTINUE ENHANCED SAMPLING?

- 2011 PV Guidance: "Considerations for the duration of the heightened sampling and monitoring period could include, but are not limited to, volume of production, process complexity, level of process understanding, and experience with similar products and processes"
- Statistical criterion for # of batches not recommended
  - Focus on *process understanding*
- Decision to discontinue should not be rigid, but allow for interim analysis after specified # of batches (not all of 3A), based on a risk assessment similar to post-PPQ

UNDERSTANDING DATA

No matter what statistical methods you use, the most important step to perform is to ***understand your data***

- Unless you understand …
  - How the data were collected
  - How the data relate to the process & control strategy
  - Potential sources of variability in the data
  - Distribution of the data

… Any statistical analyses you do with the data may be

- Inaccurate
- Misleading
- Waste of time, or
- Flat out wrong

UNDERSTANDING DATA

The best ways to understand the data include

- Process and sampling knowledge, that is, what exactly does the result represent? For example
  - What physical process created it?
  - Is the result an average, or an individual
  - Is the result a manipulated variable like an In Process Check (IPC)
  - When was it taken and measured
- *Graph, graph, graph!!* First!
- Put data in relevant order or separation to see potential *unknown* sources of variability
  - Manufacturing time order
  - Batch location
  - Dosage strength, manufacturing site, pre-/post-change, etc..

"JUST RIGHT" PV STATISTICS

- More is not necessarily better for the patient
  - Optimize samples
  - Optimize analyses
  - Optimize complexity
- Required to Optimize
  - Fundamental agreement on why we use statistics
  - Real recognition that waste is detrimental to patients
  - Fundamental understanding of essential statistical concepts (probability models, independence, random variables, variance components, etc.)
- Analyses are a continuum…not a final conclusion
- Concepts translate to other quality and operations activities

SynoloStats

INTEGRATING TOTAL SOLUTIONS

tara@synolostats.com