

Content Uniformity (CU) testing for the 21st Century: CDER Perspective

Richard (Rik) Lostritto, Ph.D.

**Acting Deputy Office Director for Science and Policy and
Biopharmaceutics Lead**

Office of New Drug Quality Assessment (ONDQA)

Office of Pharmaceutical Science (OPS)

Center for Drug Evaluation and Research (CDER)

Outline

- Introduction
- USP <905>
 - Summary
 - Features
- Case study 1: Levothyroxin
- Case study 2: Split Tablets
- Alternatives to USP <905>
- Parametric Tolerance Interval Testing (PTIT)
- Case study 3: PTIT
- Summary and conclusions

Introduction: Setting the Stage

- Content Uniformity (CU) testing is an important assessment of unit dosage form performance.
- Because pharmacological responses are dynamic, variable, and their manifestations easily confounded, clinical response alone cannot serve as an arbiter of adequate CU performance.
- CU testing is usually destructive and consumes resources (wet chemical analyses)

Introduction: Setting the Stage

- In some cases CU testing may not be destructive
 - e.g., Near Infrared (NIR) and weight variation (WV)
- Process Analytical Technologies (PAT) and Real Time Release Testing (RTRT) allow for more frequent and convenient sampling
 - Greater product knowledge
 - More efficient use of resources / less off line analysis
 - High assurance of product quality

Introduction: Launch point

- Is USP <905> perfect? No.
- Is USP <905> adequate? Not all the time.
- What are the weaknesses in USP <905> ?
- Are there viable alternatives?
 - Parametric Tolerance Interval (PTIT)
 - Large N
- Cases of interest
 - Narrow therapeutic index drugs
 - Drug device combination products
 - Specialized dosage forms (e.g., inhalation)

Summary of USP<905>

- Applies to many dosage forms
- Weight variation (WV) versus content uniformity (CU) for tablets (as an example)
 - If < 25 mg or $< 25\%$ then CU
 - If ≥ 25 mg and $\geq 25\%$ then WV
- Sample sizes for tablets are
 - 10 units (L1) and
 - 30 units (L2 as L1 + 20 more)

Summary of USP<905> (as CU for tablets)

- Acceptance value, **$AV = |M-X| + ks$**
 - X = sample mean as percent of label claim (%LC)
 - K = 2.4 for L1 and K = 2.0 for L2
 - S = standard deviation of the SAMPLE
 - M depends on the sample mean
 - If X is between 98.5 %LC and 101.5 %LC then M = X
 - If X is < 98.5 %LC then M = 98.5
 - If X is > 101.5 %LC then M = 101.5
 - **This scenario creates a zone of indifference to the mean**
 - AV cannot exceed 15.0 at L1 or 25.0 at L2
- **There are also limits on individual values**

Summary of USP<905> (as CU for tablets)

- Limits on individual values in L2
 - No unit $< 0.75 M$ and
 - No unit $> 1.25 M$
 - That means no unit outside 75 -125 %LC for on target mean (if the mean is within 1.5% of LC)
 - For off target means ($> \pm 1.5\%$) the limits on individuals vary

Features of USP <905> as CU for tablets

- Fixed sample sizes as L1 and L2 (10 / 30)
- L1 (especially) and L2 are relatively small and may not provide for a confident estimation of the batch compared to larger sample sizes.
- There are limits on individual values.
 - Is that valid in a statistically relevant model?
- There is a 1.5% zone of indifference around 100%LC.
 - Rewards off-target sample mean
 - Also changes M which changes individual limits

Case Study 1: Levothyroxine

- Potency
 - Drug product specifications allow up to 10% loss of potency over expiry
 - Intermediate strengths (112-150 mcg) are separated by less than 10% of dose.
 - One strength has degraded to contain less active than a lower strength tablet.
- Tighter controls necessary for potency to assure accuracy of individualized dosing which is titrated per patient.

Case Study 1: Levothyroxine

- CU
 - Variability is also observed within batch
 - Loss of potency and inter batch variability are linked
 - Both affect content uniformity performance
 - Using USP <905> at L1 (10 tablets) for low dose case
 - With 10% loss of potency an $RSD \leq 2.7\%$ passes
 - With 5% loss of potency an $RSD \leq 4.8\%$ passes
 - For reference, at 100% LC, an $RSD \leq 6.25\%$ passes
 - USP <905> has a zone of indifference of $1.5\% \pm LC$
 - Meeting USP <905> content uniformity with $\leq 5\%$ potency loss may provide reasonable assurance that strengths do not lose order of potency on stability.

Case Study 2: Split Tablets

- Economic driving forces encourage tablet splitting
 - Typically, the 2X strength product is not doubly expensive
 - Patients, prescribers, and insurers recognize this
- Current practices for immediate release tablets
 - Splitting of scored and non scored tablets
 - Hand splitting
 - Kitchen knives and other household implements
 - Tablet splitters (not regulated)

Case Study 2: Split Tablets

- What are some potential problems?
 - If scored, is the score functional?
 - Ease to split (splitability)
 - Even splitting by weight
 - Loss of mass (crumbs) typically 1-2%
 - Performance of the major pieces (i.e., halves) regarding:
 - Friability
 - Dissolution rate
 - Physical and chemical stability
 - Content uniformity (esp low dose cases)

Case Study 2: Split Tablets

- Some Points to Consider for CU of the pieces
 - Which tablets are appropriate for splitting (IR vs. ER)?
 - Is a functional score necessary?
 - Valid assessment of “splitability”?
 - Stability assessment (90 days post dispensing?)
 - New active surface exposed
 - New sharp edges and corners exposed
 - Dissolution assessment (meet same specifications?)
 - Loss of mass (what’s a maximum allowable loss?)
 - Content Uniformity
 - Should each piece Meet USP <905> ?

Potential Alternatives to USP <905>

- Large N
 - For RTRT approaches, FDA has approved Large N approaches using WV and NIR
 - May be appropriate when large numbers of tablets can be adequately measured; usually non destructively
 - The Agency is working with a goal to develop harmonized criteria with other regions for Large N approaches
- Parametric Tolerance Interval Testing (PTIT)
 - Similar to some Large N approaches, but typically smaller sample sizes
 - Discussed in FDA Draft Guidance (add reference)

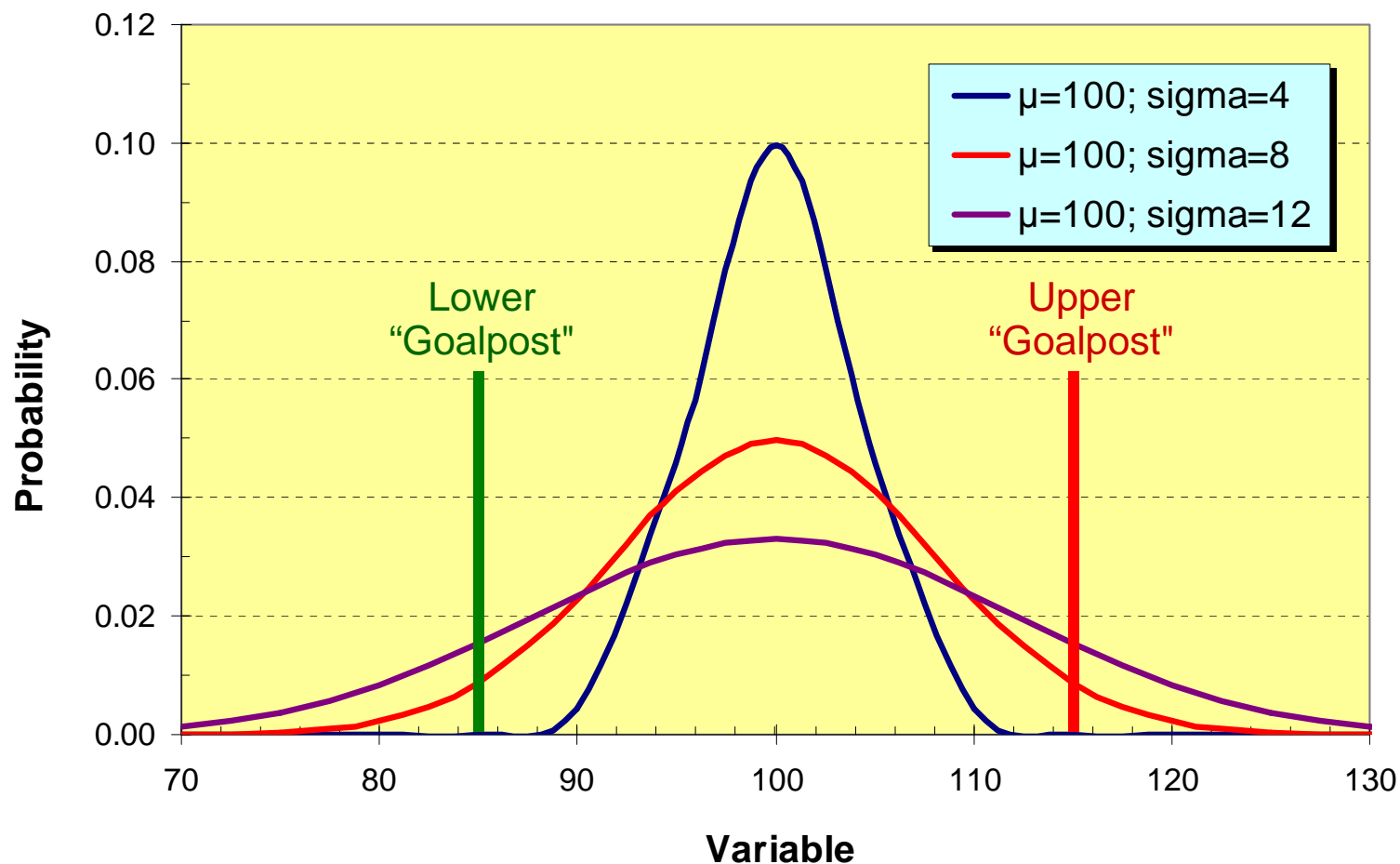
Parametric Tolerance Interval Testing (PTIT)

- What is PTIT?
 - A statistically relevant way of estimating CU by measuring it in a sample
- The test has two sample tiers Tier-1 and Tier-2
 - A different concept than L1 and L2 of USP<905>
 - Applicant chooses and specifies the tier sizes based on confidence in their process and product performance.
- Let's explore further through an example
 - Case 3

Parametric Tolerance Interval Testing (PTIT)

- What are the advantages of PTIT over USP<905>
 - Sample size choices are more flexible
 - Sample sizes (2 tiers); typically larger than USP but smaller than “Large N”
 - Passing at Tier-1 rewards high quality
 - No zone of indifference
 - No restriction on individual values
 - No zero tolerance

PTIT: BATCH Distributions, Goalposts, and the Concept of COVERAGE.



Parametric Tolerance Interval Testing (PTIT)

- If the mean is $>100\%$ LC $MSD = [120 - SM] / K$
- If the mean is $< 100\%$ LC $MSD = [SM - 80] / K$
- Where
 - SM = sample mean
 - MSD = the maximum sample standard deviation allowable to PASS at this Tier (1 or 2)
 - K is a PTI model derived constant and depends upon
 - Sample size
 - Tier-I or Tier-II
 - Relationship to Tier-I to Tier-II
 - The goalposts
 - The desired coverage between the goalposts

Parametric Tolerance Interval Testing (PTIT)

- Note that:
 - The smaller K is, the larger the allowable MSD is
 - The closer Sample Mean (SM) is to 100% of target, the larger is the allowable MSD
 - The smaller the sample SD is, the more off-target SM may be and still pass the test
 - SM must be within the goalposts (obvious)
 - There is no model constraint on individual values
- Here are some values of K ...

PTIT K values

N = sample size						
% coverage	N=10	N=30	N=20	N=60	N=30	N=90
	Tier-I	Tier-II	Tier-I	Tier-II	Tier-I	Tier-II
	K10	K30	K20	K60	K30	K90
82.5	2.82	1.94	2.20	1.74	2.00	1.66
85.0	2.96	2.04	2.32	1.83	2.11	1.75
87.5	3.12	2.16	2.45	1.94	2.23	1.86
90.0	3.31	2.30	2.60	2.07	2.37	1.98

Case 3: Simulated Batch for PTIT

- Design a sim batch
 - Let batch mean = 100 %LC
 - Let batch Sd = 8.0 %
- Let Goalposts be 80% to 120% of LC
- Set coverage to 90% within goalposts
- Try the smallest possible sampling plan
 - Tier-I = 10
 - $MSD = [SM-80]/K10$ or $MSD = [120-SM]/K10$
 - Tier-II = 30
 - $MSD = [SM-80]/K30$ or $MSD = [120-SM]/K30$
- Select appropriate K values from the Table

Case 3: Simulated Batch for PTIT

N = sample size						
% coverage	N=10	N=30	N=20	N=60	N=30	N=90
	Tier-I	Tier-II	Tier-I	Tier-II	Tier-I	Tier-II
	K10	K30	K20	K60	K30	K90
82.5	2.82	1.94	2.20	1.74	2.00	1.66
85.0	2.96	2.04	2.32	1.83	2.11	1.75
87.5	3.12	2.16	2.45	1.94	2.23	1.86
90.0	3.31	2.30	2.60	2.07	2.37	1.98

Case 3: Simulated Batch for PTIT

- Tier-I = 10
 - $\text{MSD} = [\text{SM}-80]/3.31$ or $\text{MSD} = [120-\text{SM}]/3.31$
- Tier-II = 30
 - $\text{MSD} = [\text{SM}-80]/2.30$ or $\text{MSD} = [120-\text{SM}]/2.30$
- Pass if sample standard deviation < MSD
- Do that 30 times
- Here's the results of the first 10 runs

Case 3. SM=100% and SD=8.0%

90% coverage using smallest sampling plan: **10 / 30**

Run #	N=10		N=30		Tier-I	Tier-II
	SM	SD	SM	SD	1=pass	1=pass
1	100.5	10.7	99.5	8.6	0	0
2	98.9	7.1	98.0	7.8	0	1
3	98.3	9.5	100.4	8.8	0	0
4	100.6	5.9	101.9	6.3	0	1
5	101.4	5.7	99.4	8.7	0	0
6	98.9	8.8	100.2	8.6	0	1
7	99.0	5.8	100.3	6.7	0	1
8	98.6	8.8	100.5	7.6	0	1
9	98.6	8.5	99.5	6.1	0	1
10	101.7	8.6	101.0	8.5	0	0

Case 3: Simulated Batch for PTIT

- Undesirable outcome
 - The batch passed only 21/30 times
 - There were no Tier-I passes
- Explaining this outcome
 - Goalposts too far apart?
 - Coverage too high?
 - Increase coverage -> larger K
 - Sample sizes too small ?
 - Decrease sample size -> larger K
- **Let's try larger sample tier sizes**
 - Larger samples are more likely to represent the batch
 - The requirements for larger samples are less stringent

Case 3: Simulated Batch for PTIT

N = sample size						
% coverage	N=10	N=30	N=20	N=60	N=30	N=90
	Tier-I	Tier-II	Tier-I	Tier-II	Tier-I	Tier-II
	K10	K30	K20	K60	K30	K90
82.5	2.82	1.94	2.20	1.74	2.00	1.66
85.0	2.96	2.04	2.32	1.83	2.11	1.75
87.5	3.12	2.16	2.45	1.94	2.23	1.86
90.0	3.31	2.30	2.60	2.07	2.37	1.98

Case 3. $S_m=100\%$ and $SD=8.0\%$

90% coverage using LARGEST sampling plan: 30 / 90

- $K_{30} = 2.37$
- $K_{90} = 1.98$
- Run sampling and testing 30 times
- Results
 - Pass Tier-I about half the time
 - Pass Tier-II 29/30 times

Case 3. $Sm=100\%$ and $SD=8.0\%$

90% coverage using LARGEST sampling plan: 30 / 90

- This may be an acceptable situation
 - Reasonable opportunity to pass at Tier-I
 - (good)
 - High probability to pass at Tier-II
 - (very good)
 - No limit on individuals
 - (excellent)
 - The applicant has the option to *select and specify* the sampling plan to suit their risk and resources
 - May be resource intensive at Tier-II
 - Typically, coverage is set at 90% between goalposts

Parametric Tolerance Interval Testing (PTIT)

- What are the issues with PTIT?
 - Technically assumes a normal distribution of the measured property in the batch. However, provided the distribution is symmetrical, it should be ok.
 - If CU variability is not random, look for, evaluate, and control for operative non-random causes
 - There are multiple PTIT approaches
 - Tier sizes and ratios
 - Various ways to weight the power of tiers
 - Various goalposts and coverage within goalposts
 - There is no universal agreement on best approach
 - However...

Parametric Tolerance Interval Testing (PTIT)

- The Agency has identified some PTIT approaches that may be acceptable for CU testing. For inhalation products these include:
 - Goalposts 80% to 100% of label claim.
 - Coverage within goalposts, 90% of batch population
 - Two sample size tiers in a 1:3 ratio
 - For example, 10:30, 20:60, 30:90, etc.
 - Power (as 1-alpha) equally distributed between tiers
 - The applicant chooses and specifies the tier sizes
 - Cases vary; **recommend discussion** with the Agency before implementing

Case Study 4: Drug Eluting Cardiac Stents

- Very low drug levels in or on the stent
- Potent drugs used (e.g., paclitaxel)
- Extended release design, in place a long time
- Local treatment of vulnerable target tissue
- Local tissue levels likely in therapeutic range.
 - Dose response relationships exist
- The vast majority of these products comply with USP <905> for tablets at L1 (n=10)

Summary and Conclusions

- In some cases USP <905> works fine for CU testing, but it has flaws:
 - Small fixed sample sizes for L1 and L2
 - Zone of indifference of 1.5%
 - Limits on individual values and they may vary according to the sample mean
- PTIT testing offers greater flexibility and appears to have greater statistical relevance provided the batch distribution is near-normal or symmetrical

Summary and Conclusions

- Loss of potency on stability for a narrow therapeutic drug may put more demands on what is acceptable CU performance.
- Split tablet performance is a current and pervasive issue that merits careful consideration in terms of CU
- Pharmacokinetic and pharmacological input can further help to determine the criticality of CU performance to safety and efficacy

Acknowledgements

- Yi Tsong, Office of Biostatistics, CDER
- Meiyu Shen, Office of Biostatistics, CDER
- Christine Moore, Office of New Drug Quality Assessment, CDER

Thank you

Address Questions or Comment to:
NewDrugCMC@fda.hhs.gov