Development of a Dose Content Uniformity Test Suitable for Medium and Large Sample Sizes

Yi Tsong,
DB VI, OB, CDER, FDA
Yi.tsong@fda.hhs.gov

Based on collaborative works with
Meiyu Shen, Jinglin Zhong
Xiaoyu Dong and Richard Lostrito,

This presentation represents the presenter's opinions. They do not necessarily represent FDA official position

Outlines

I. Introduction
   - Bias of USP Harmonized Test

II. Alternative Procedures
   - Two Sided and Two One-sided Hypotheses
   - Two-sided Tolerance Interval
   - Two One-sided Tolerance Intervals
   - Tolerance Interval Controlling Both Tails

III. PhRMA Large Sample Test

IV. Hypothesis Development for Large Sample Sizes

V. Proposed Procedure

VI. Discussion

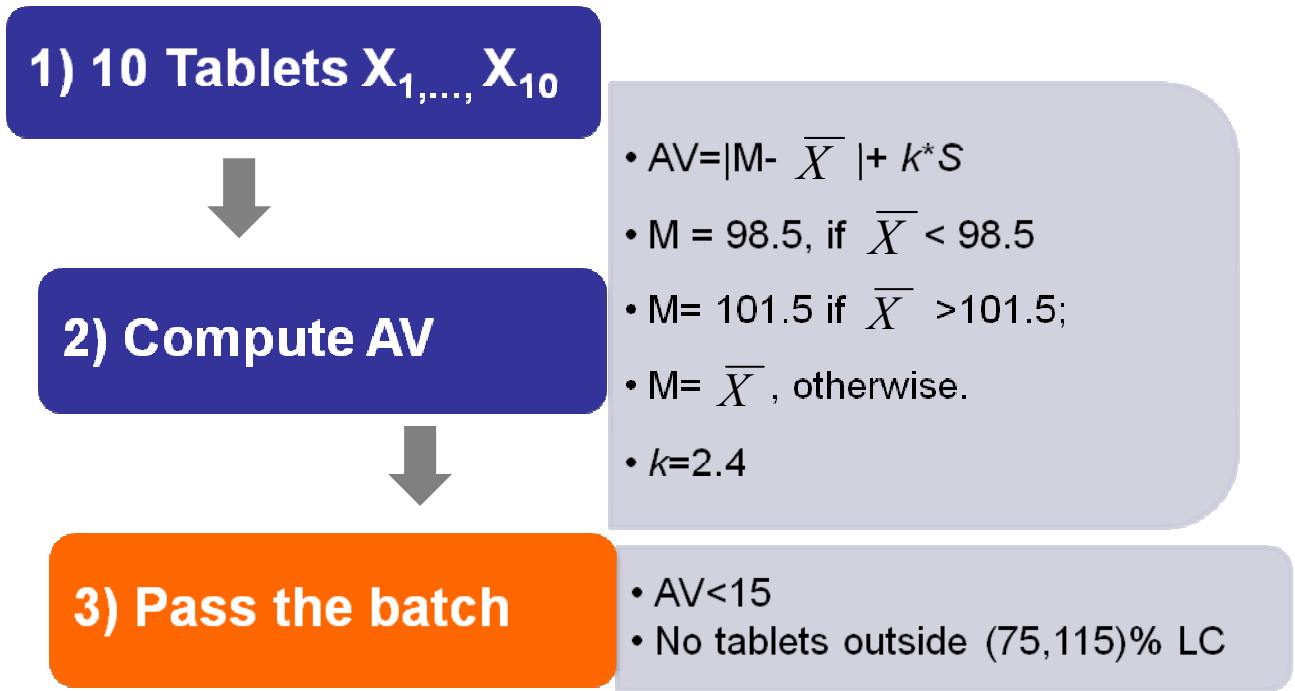
Introduction

Bias of USP Harmonised Test (Shen & Tsong, PF, 2011)

- For a therapeutic product, most of the dose contents should be within (85, 115)%LC to assure the homogeneity of the product.
- DCU compliance can be determined through the sampling acceptance plan.

USP Harmonised Test (USPHT)

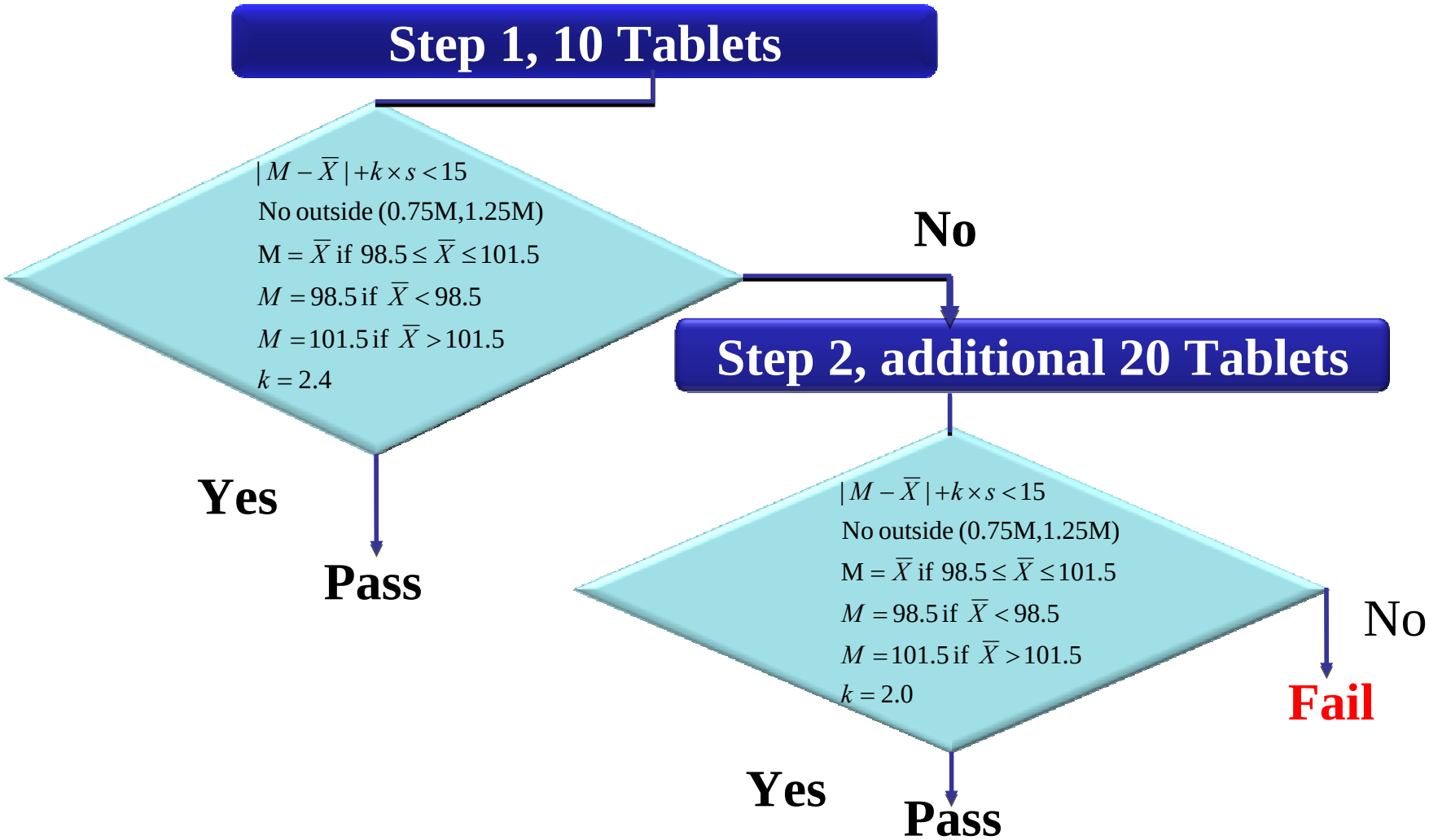
•**Step 1**



•**If Step 1 fails, go to Step 2:**

- sample additional 20 tables
- decision rule is similar as step 1 but based on the total 30 tablets with k=2.0

USP Harmonized Test



- It is modified from two-sided tolerance interval (JP XIII).

$$\mathrm{AV} = |M - \bar{X}| - kS < 15$$

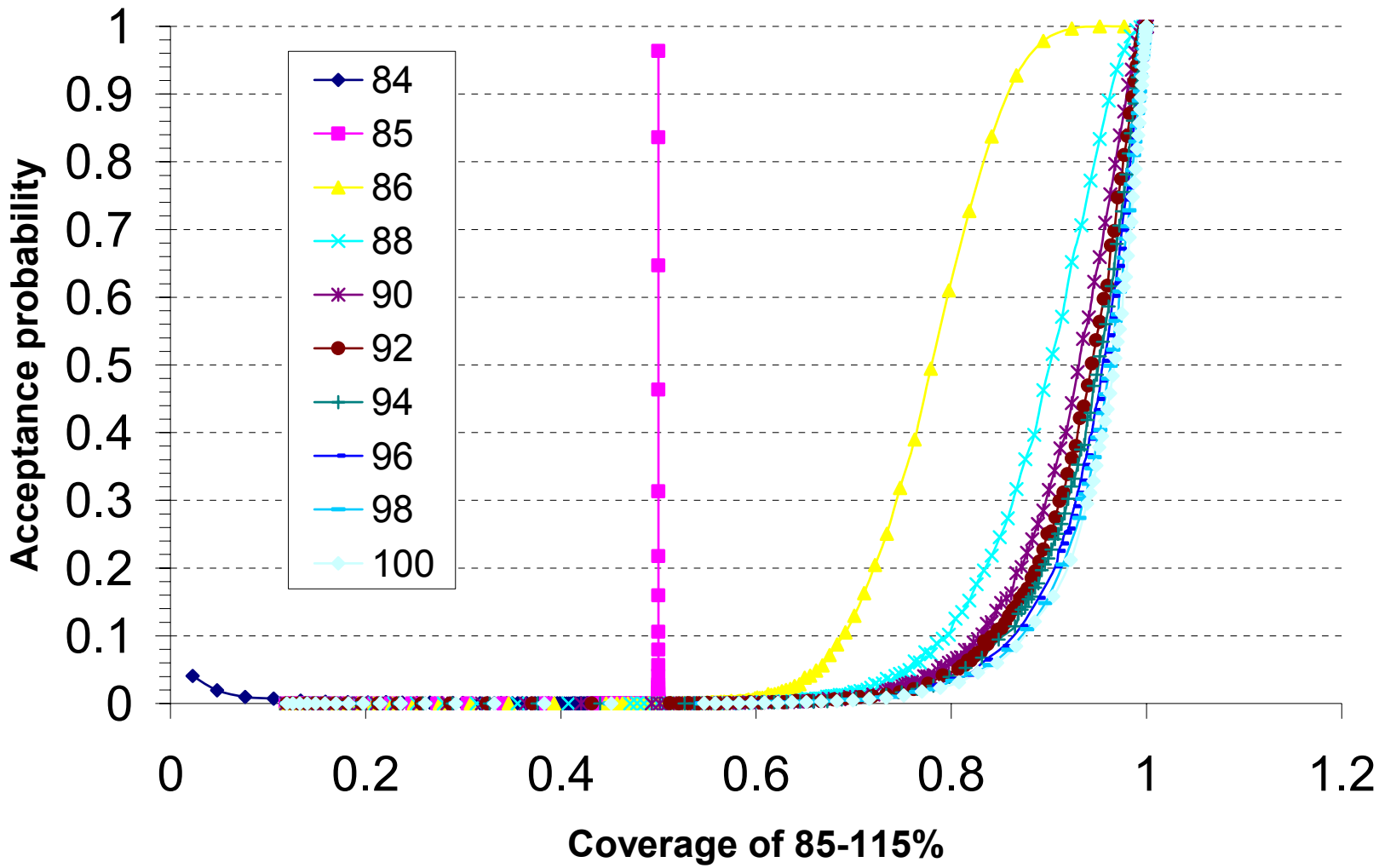
$$\Leftrightarrow \quad \mathrm{M} - 15 < \bar{X} - kS < \bar{X} + kS < M + 15$$

$$M = 98.5 \text{ if } \bar{X} < 98.5$$

$$M = 101.5 \text{ if } \bar{X} > 101.5$$

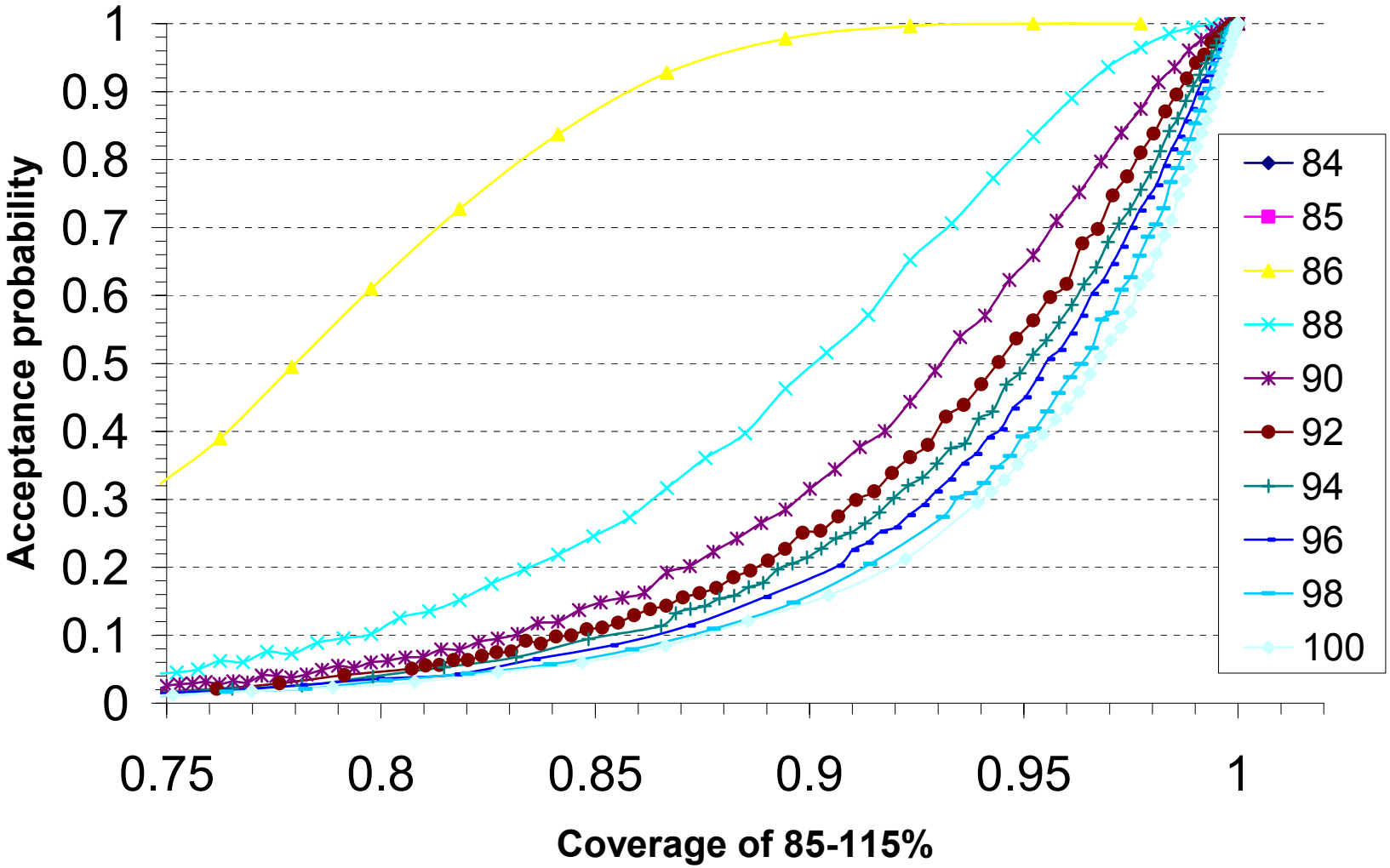
- However, as a consequence, USP harmonised test is biased in favor of off-target products.

USP harmonized DCU



Shen M, Tsong Y, Bias of USP Harmonised test for dose content uniformity. Pharmacopeia Forum 2010, To appear

USP harmonized DCU



Shen M, Tsong Y, Bias of USP Harmonised test for dose content uniformity. Pharmacopeia Forum 2010, To appear

II. Alternative Procedures

Tolerance Intervals for Controlling Two Tail End Probabilities
(Dong, Shen, Zhong and Tsong, SBR, To submit, 2010)

- Two sided hypotheses

$$H_0 : \Pr(L < X < U) \le P \text{ vs. } H_1 : \Pr(L < X < U) > P$$

- Two one-sided hypotheses (FDA Delivery dose uniformity; Tsong and Shen (JBS, 2007)

$$H_0^L : \Pr(X < L) \ge p_1 \text{ vs. } H_1^L : \Pr(X < L) < p_1$$

$$H_0^U : \Pr(X > U) \ge p_2 \text{ vs. } H_1^U : \Pr(X > U) < p_2$$

Where (L,U) =(85, 115); p_1 and p_2 are the limit proportions of under-filled and over-filled tablet

Two sided hypotheses

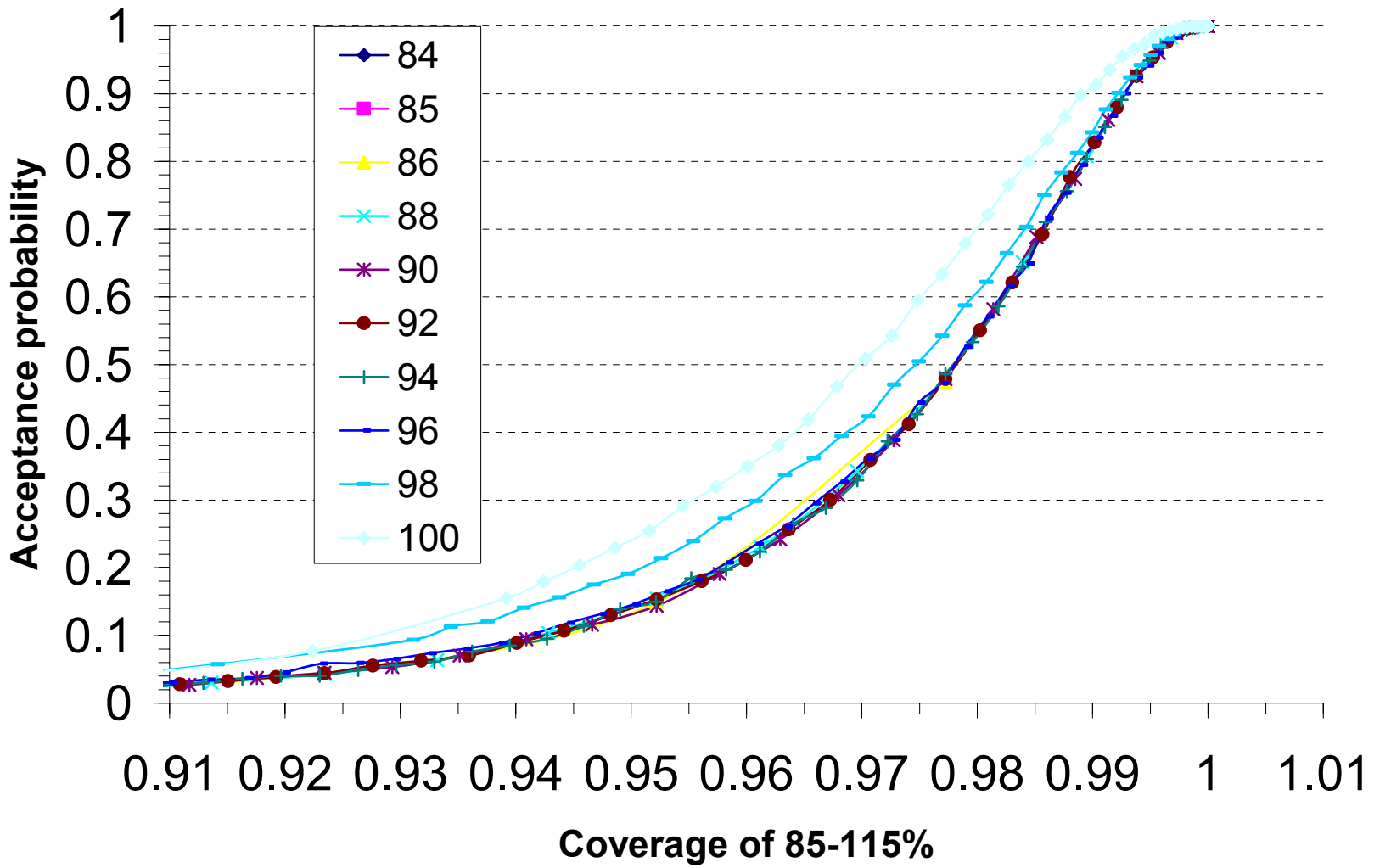
$H_0 : \Pr(L < X < U) \le P$ vs. $H_1 : \Pr(L < X < U) > P$

- Two-sided tolerance interval approach (PTSTI)
  - Calculate 95% confidence P-coverage two-sided tolerance interval ($\bar{x} - Ks,\ \bar{x} + Ks$) with k as the solution of

$$\Pr\{\int_{\bar{x}-Ks}^{\bar{x}+Ks} n(x : \mu, \sigma)dx \ge 87.5\%\} = 0.95$$

  - Reject H_0 if $L < \bar{x} - Ks$ and $\bar{x} + Ks > U$.

Beta=87.5%, PTSTI



How about use it for testing the two-sided hypotheses?

Power (×100) of PTSTI factor k(=2.000) with n=30, p=0.875, α=5% for the two-sided hypothesis test setting.

High End: p_2=P(X>115)100% \ Low End: p_1=P(X<85)100%	12	11	10	9	8	7	6	5	4	3
0.5	0.45	0.74	1.22	1.99	3.27	5.35	8.77	14.35	23.35	37.51
1.5	0.43	0.70	1.13	1.84	2.99	4.87	7.88	12.71	20.3	31.79
2.5	0.36	0.58	0.94	1.51	2.42	3.89	6.20	9.81	15.31	23.28
3.5	0.28	0.44	0.71	1.13	1.80	2.86	4.49	6.99	10.68	15.85
4.5	0.20	0.32	0.51	0.81	1.28	2.00	3.11	4.76	7.15	10.39
5.5	0.14	0.22	0.35	0.56	0.87	1.36	2.08	3.16	4.67	6.67
6.5	0.10	0.15	0.24	0.37	0.58	0.90	1.37	2.06	3.01	4.23
7.5	0.06	0.10	0.16	0.25	0.38	0.59	0.89	1.32	1.92	2.66
8.5	0.04	0.07	0.10	0.16	0.25	0.38	0.57	0.85	1.21	1.67
9.5	0.03	0.04	0.07	0.10	0.16	0.24	0.37	0.54	0.76	1.04

Pr(85<X<115)=0.875

Pr(85<X<115)>0.875

How about use it for testing the two one-sided hypotheses?

Power(×100) of PTSTI factor k(=2.000) with n=30, p=0.875, α=5% for TOST.

High End: $b=\mu+Z_p\sigma$ -115 \ Low End: $a=\mu-Z_p\sigma$-85	-14	-10	-2	-1	0	1	2	10	14
14	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0
2	0	0	0.16	0.26	0.42	0.62	0.85	0.97	0.54
1	0	0	0.26	0.47	0.78	1.22	1.75	2.88	2.21
0	0	0	0.42	0.78	1.37	2.24	3.39	7.79	7.79
-1	0	0	0.62	1.22	2.24	3.85	6.08	18.46	22.36
-2	0	0	0.85	1.75	3.39	6.08	10.01	37.19	49.19
-10	0	0	0.97	2.88	7.79	18.46	37.19	100	100
-14	0	0	0.54	2.21	7.79	22.36	49.19	100	100

Pr(X<85) =0.0625, or Pr(X>115) = 0.0625

Pr(X<85) < 0.0625, and Pr(X>115) < 0.0625

- X~ $N(\mu,\sigma^2)$, The two one-sided hypotheses are equivalent to test for:

$$H_0^L:\mu - Z_p\,\sigma \le \mathrm{L} \quad \text{vs.} \quad H_1^L:\mu - Z_p\,\sigma > \mathrm{L}$$

and

$$H_0^U:\mu + Z_p\,\sigma \ge \mathrm{U} \quad \text{vs.} \quad H_1^L:\mu + Z_p\,\sigma < \mathrm{U}$$

P=1-p_1-p_2

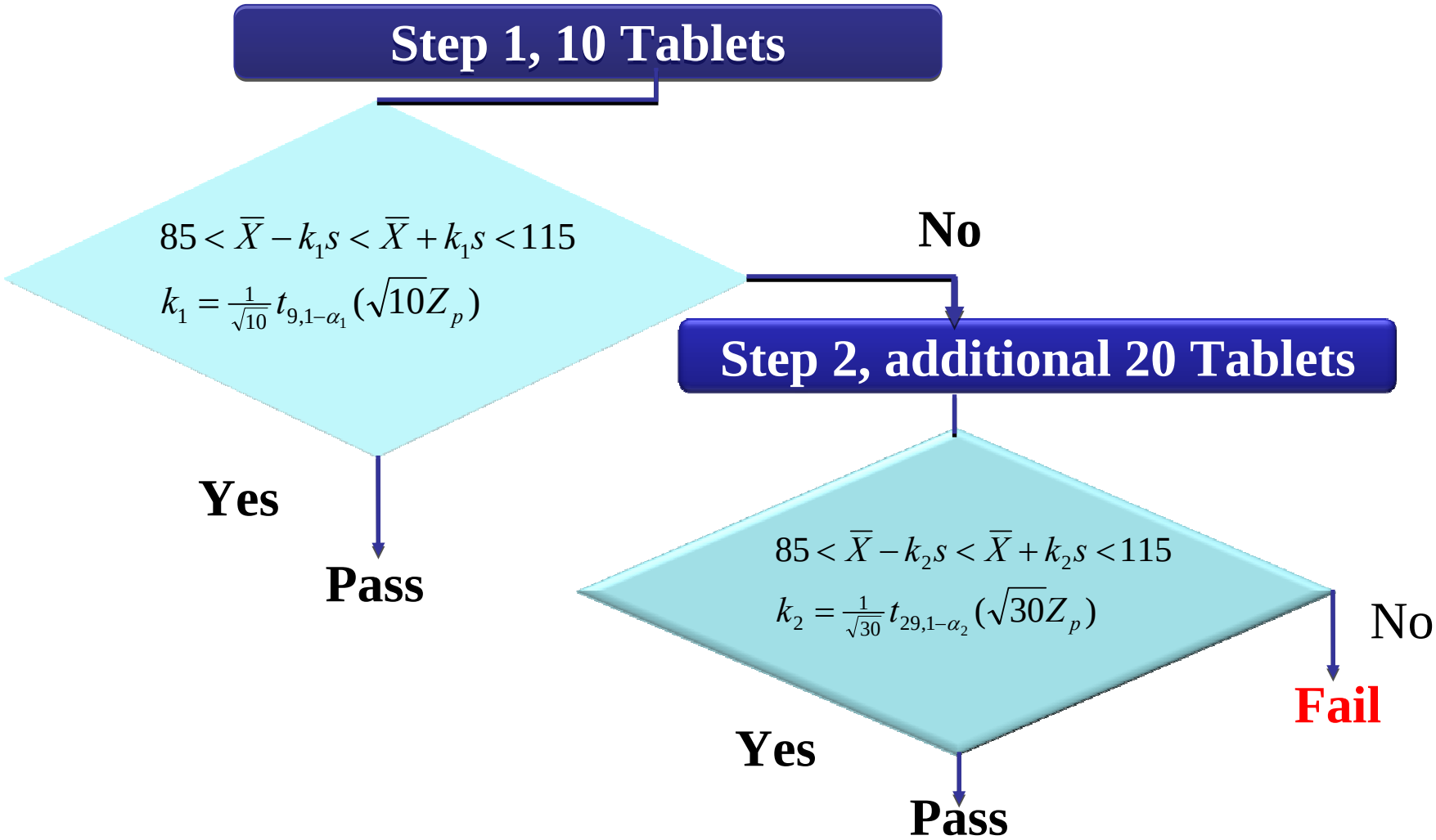
- Each hypothesis is tested at α; pass the batch if

$$L < \bar{X} - kS < \bar{X} + kS < U$$

With $k = \frac{1}{\sqrt{n}} t_{n-1,1-\alpha}(\sqrt{n}Z_p)$,

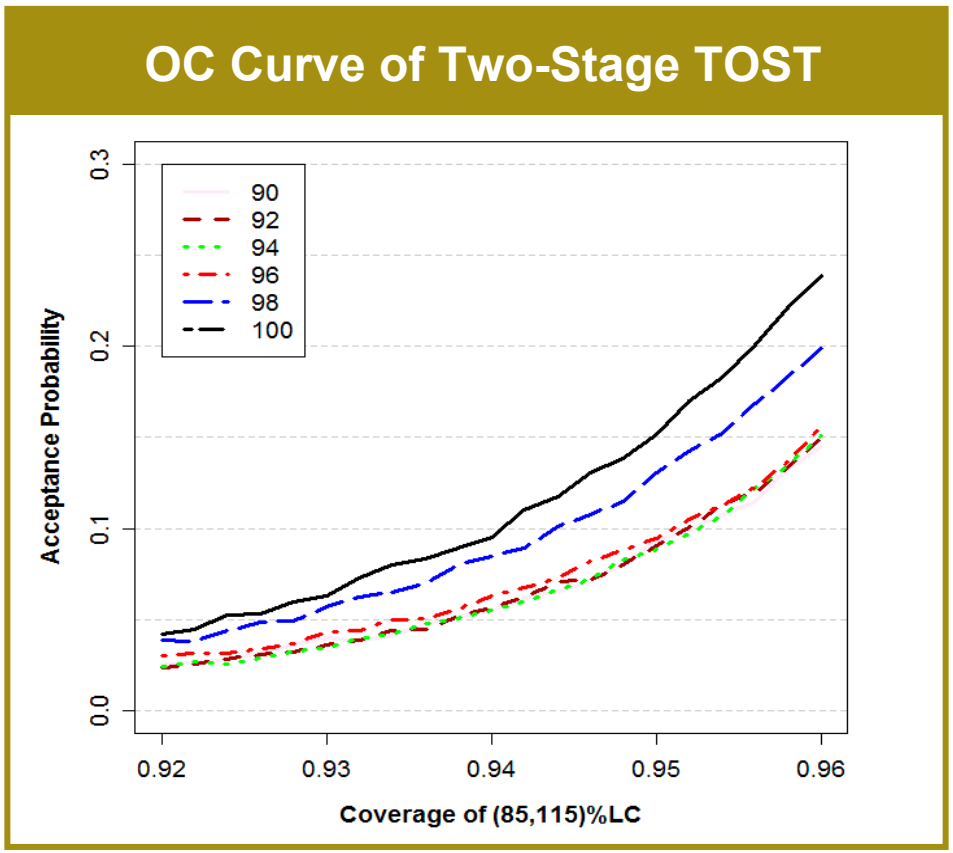
i.e. intersection of two one-sided tolerance intervals.

Two-Stage TOST

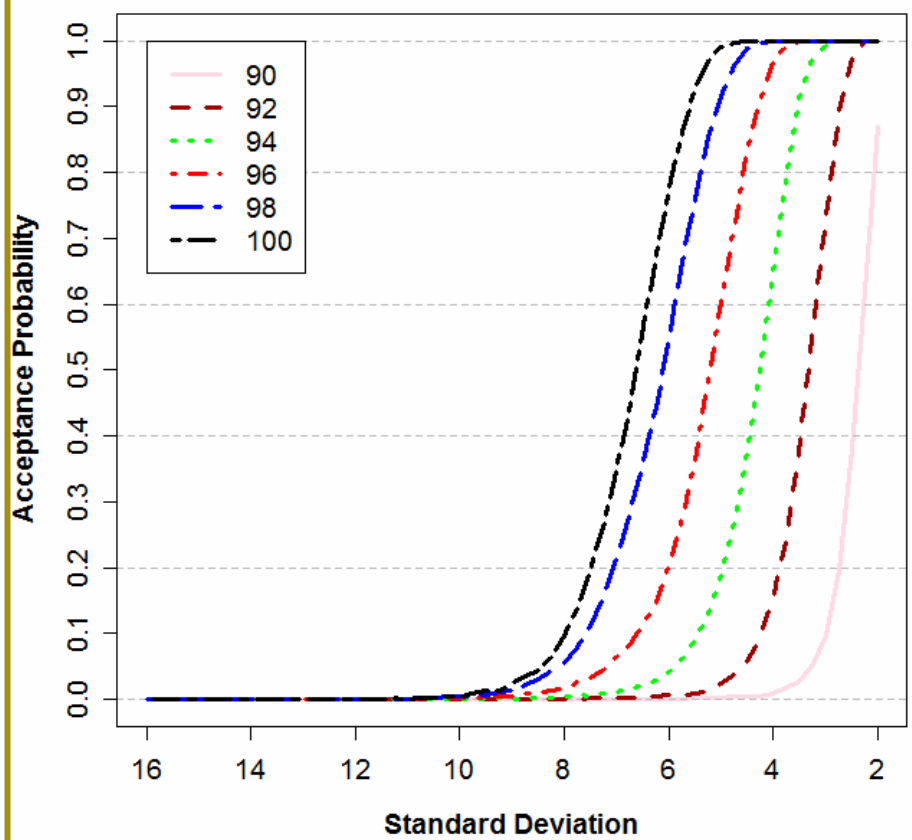


Properties of TOST

- Two-Stage TOST is **unbiased** for off-target products



OC Curve of Two-Stage TOST



- Intersection-Union Test

Power (×100) of TOST factor k(=2.084) with n=30, p=0.875, α=5% for TOST.

High End: $b=\mu+Z_p\sigma$ -115 \ Low End: $a=\mu-Z_p\sigma$-85	-14	-10	-2	-1	0	1	2	10	14
14	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0
2	0	0	0.07	0.13	0.21	0.32	0.44	0.54	0.3
1	0	0	0.13	0.23	0.4	0.64	0.95	1.73	1.3
0	0	0	0.21	0.4	**0.72**	1.22	1.91	**5**	**5**
-1	0	0	0.32	0.64	1.22	**2.18**	**3.58**	12.85	15.9
-2	0	0	0.44	0.95	1.91	3.58	6.17	28.29	39.2
-10	0	0	0.54	1.73	5	12.85	28.29	100	100
-14	0	0	0.30	1.30	5	15.90	43.21	100	100

Size α

Biased

μ=100, δ↓ Monotone

Pr(X<85) =0.0625, or Pr(X>115) =0.0625

Pr(X<85) <0.0625, and Pr(X>115) <0.0625

Equal-tailed Two-Sided Tolerance Interval

- Owen (1964) proposes equal-tailed two-sided tolerance interval to infer the both tail proportions simultaneously.

$$P(\overline{X} - k^* S < \mu - Z_P \sigma < \mu + Z_P \sigma < \overline{X} + k^* S) = 1 - \alpha$$

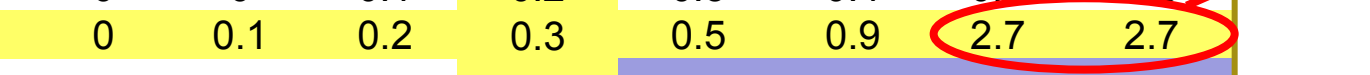
Accept the batch if

$$L < \overline{X} - k^* S < \overline{X} + k^* S < U$$

- We also consider this method for batch quality assessment.

Power (×100) of Equal-tailed tolerance factor k(=2.195) with n=30, p=0.875, α=0.05 for TOST.

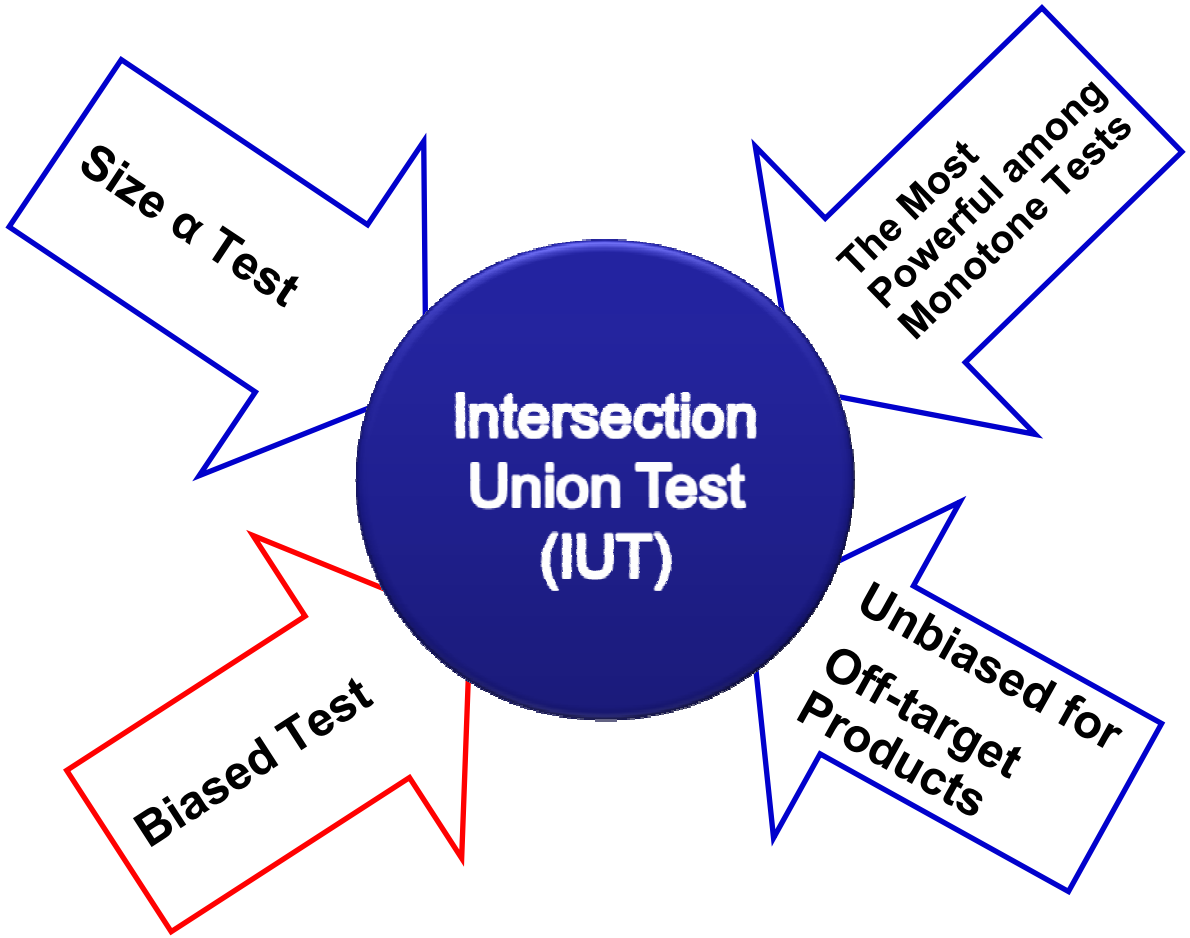
High End: $b=\mu+Z_p\sigma$ -115 \ Low End: $a=\mu-Z_p\sigma$-85	-14	-10	-2	-1	0	1	2	10	14
14	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0.1	0.1	0.2	0.3	0.1
1	0	0	0	0.1	0.2	0.3	0.4	0.9	0.6
0	0	0	0.1	0.2	0.3	0.5	0.9	2.7	2.7
-1	0	0	0.1	0.3	0.5	1.0	1.7	7.7	9.7
-2	0	0	0.2	0.4	0.9	1.7	3.1	18.9	27.5
-10	0	0	0.3	0.9	2.7	7.7	18.9	100	100
-14	0	0	0.1	0.6	2.7	9.7	27.5	100	100



Pr(X<85) =0.0625, or Pr(X>115) =0.0625

Pr(X<85) <0.0625, and Pr(X>115) <0.0625

- **TOST Properties In summary**



Proof see Appendix

Issues for Larger Sample Testing

- Developing DCU test for larger sample size (n>30) because
  - Simultaneous multi-tablet content measurement is becoming feasible with near infrared spectroscopy (NIRS).
  - Higher power to separate on-target and off-target batch.
  - More insights for the manufacturing process.
  - Trend for future quality control.
- Alternative approaches
    - PhRMA Counting Method
    - Two One-sided Tests (TOST)
    - Equal-tailed Tolerance Interval
- What to choose for P, p_1, p_2 for n > 30 „

III. PhRMA Large Sample Test (Counting Method)

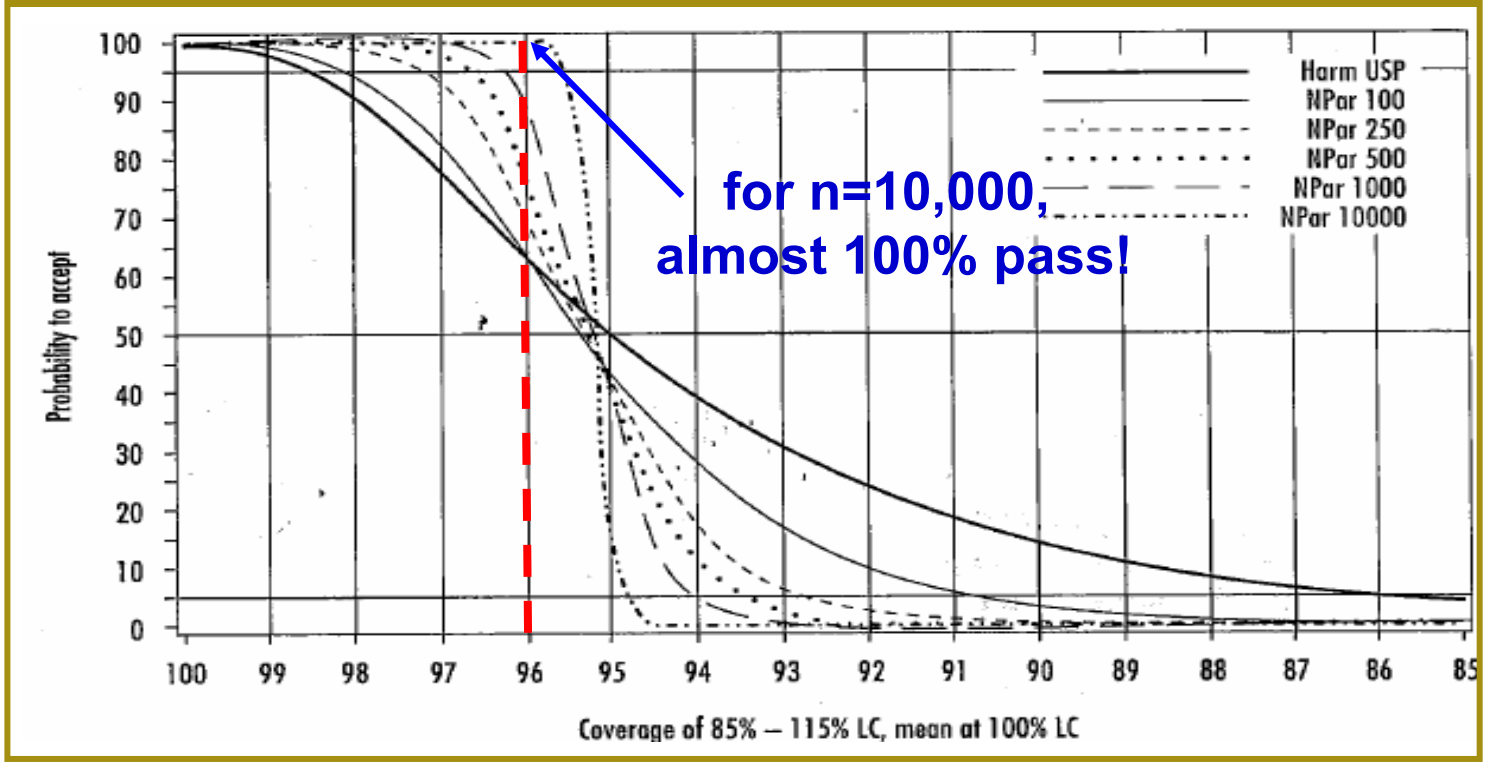
- PhRMA (2006, DIJ) proposes an alternative CUT for large sample sizes.
- The batch is rejected if number of dosage units outside (85, 115) is too high (>c).

Acceptance Limit of Proposed Test for a Selection of Sample Sizes										
n	**100**	**250**	**500**	**750**	**1000**	**2000**	**3000**	**4000**	**5000**	**10000**
c	**4**	**11**	**23**	**35**	**47**	**95**	**143**	**191**	**239**	**479**

Source: Sandell D, et al. DIJ, 40 337-344, 2006

- **FDA reviewers don't agree with PhRMA counting procedure.**

OC Curve of the PhRMA Counting Procedure



- PhRMA method can not assure good quality of a batch.

Source: Sandell D, et al. DIJ, 40 337-344, 2006

Tolerance Interval Approaches for the Two One-sided Hypotheses

PhRMa

- Based PTSTI
- Does not control tail end probabilities

TOST

- Correct the bias of USP Harmonized Test
- Control for Type I Error Rate
- Assure both efficacy and safety
- Easy to apply.

n>30

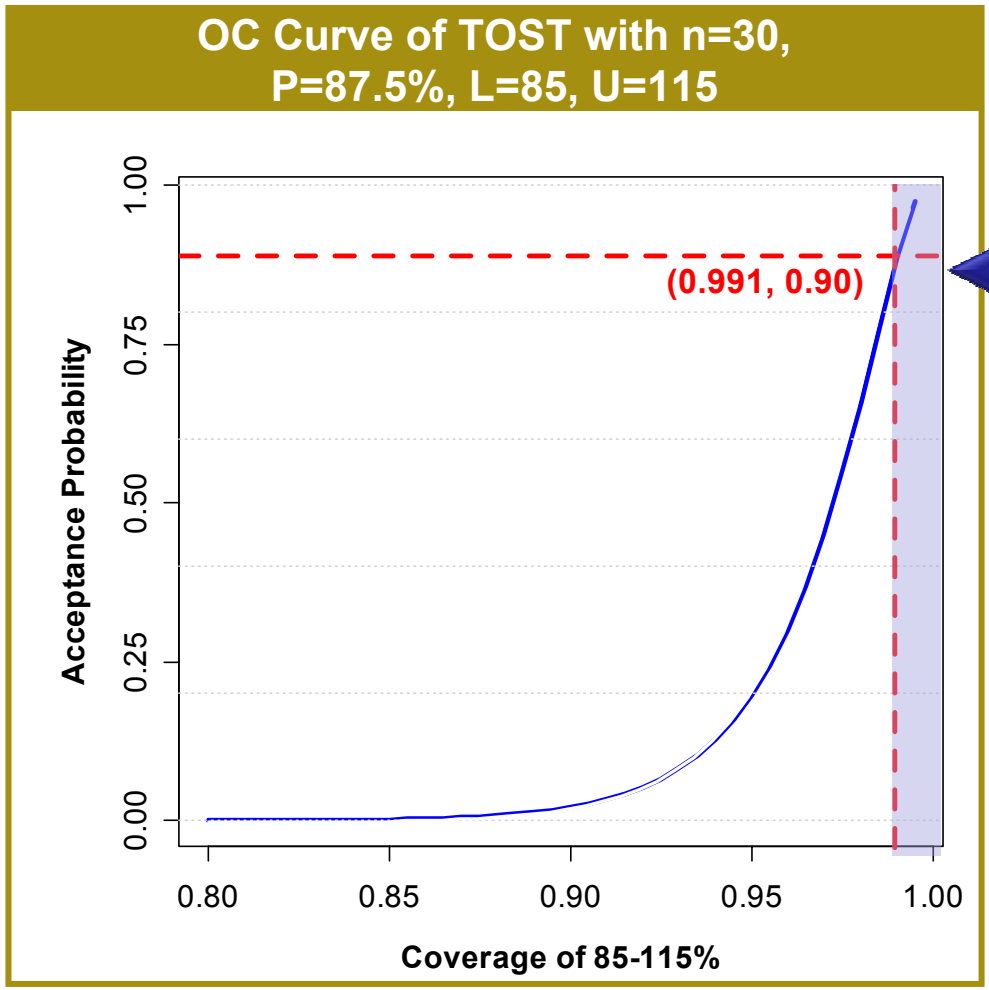
Tolerance Interval

- Too conservative
- Lower Power than TOST
- No close form

Changes: PhRMA – Based on PTSTI; Does not control tail end probabilities

IV. Hypothesis Development for DCU Test with Large Sample Sizes

(Dong, Tsong, Shen and Zhong, JBS, 2011, To submit)

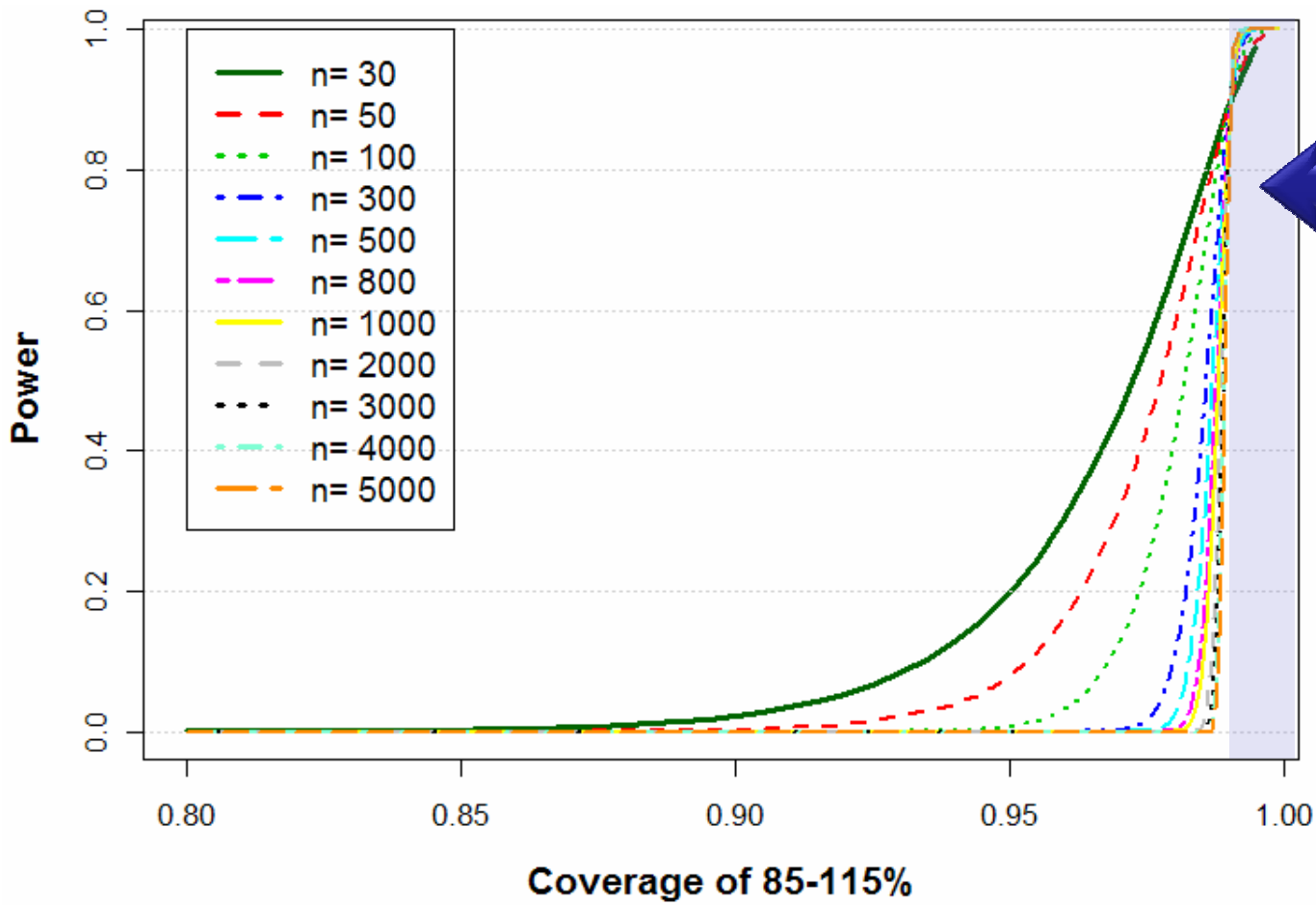


Example

(FDA Chemists) Good Quality: cp>99.1%

control producer's risk

• **Project Goal:** For n>30, batch with 99.1% coverage for (85,115) or higher should have **higher passing rate** while **controlling type I error**



Good Quality

Hypothesis Development for Large Sample Sizes

- Define ***p(n)*** in H_0 adjusted for sample size *n*.

$$H_0^L: \mu - Z_{p(n)}\,\sigma \le \mathrm{L} \quad \text{vs.} \quad H_1^L: \mu - Z_{p(n)}\,\sigma > \mathrm{L}$$

and

$$H_0^U: \mu + Z_{p(n)}\,\sigma \ge \mathrm{U} \quad \text{vs.} \quad H_1^L: \mu + Z_{p(n)}\,\sigma < \mathrm{U} \qquad (8)$$

- All power curves (n>30) intersect at (cp=99.1%, power=90%).

- Power of TOST: Reject both H_0^L and H_0^U.

$$
\begin{aligned}
\gamma &= \Pr(R_{H_0^L} \cap R_{H_0^U} \mid H_1) \\
&= \Pr(L < \bar{X} - kS < \bar{X} + kS < U \mid H_1) \\
&= \Pr\left(\tfrac{L-\mu}{\sigma} < \tfrac{\bar{X}-\mu}{\sigma} - k\tfrac{S}{\sigma} < \tfrac{\bar{X}+\mu}{\sigma} + k\tfrac{S}{\sigma} < \tfrac{U-\mu}{\sigma} \mid H_1\right) \\
&= E_W\{\Phi(\sqrt{n}(\tfrac{U-\mu}{\sigma} - k\sqrt{\tfrac{W}{n-1}})) - \Phi(\sqrt{n}(\tfrac{L-\mu}{\sigma} + k\sqrt{\tfrac{W}{n-1}})) \mid H_1\}
\end{aligned}
\qquad (9)
$$

where $W = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$ and Φ is the cdf for N(0,1)

- Based on the power function, we develop a two-step method to determine *p(n)*:

Step 1: Solve for k subject to

$$\gamma(n, k, \mu = 100, p_1 = p_2 = 0.45\%) = 0.90$$

k

Step 2: Solve for $p(n)$ from

$$k = \frac{1}{\sqrt{n}} t_{n-1,1-\alpha}(\sqrt{n} Z_{p(n)})$$

$p(n)$

V. Proposed Procedure

Quality standard $p(n)$ determined by the proposed two-step method for μ=100, L=85, U=115, α=0.05 and various sample sizes.

n	50	100	300	500	800	1000	2000	3000	4000	5000
k	2.185	2.296	2.415	2.454	2.482	2.493	2.522	2.536	2.543	2.548
$p(n)$	0.920	0.952	0.974	0.979	0.982	0.983	0.986	0.987	0.987	0.988

- For n=500, $p(n)$=0.979. If $p_1 \geq 1.05\%$ or $p_2 \geq 1.05\%$, the batch is rejected; it is equivalent to test

$$H_0^L: \mu - Z_{0.979}\sigma \leq 85 \text{ vs. } H_1^L: \mu - Z_{0.979}\sigma > 85$$

and

$$H_0^U: \mu + Z_{0.979}\sigma \geq 115 \text{ vs. } H_1^L: \mu + Z_{0.979}\sigma < 115$$

VI. Discussion

- Examined the relationship between intersection-union test and tolerance interval controlling tail end probabilities
- Studied power of three tolerance interval procedures for testing two one-sided hypotheses
- Specification of the null hypothesis for content uniformity changes with sample size.
- Such a specification is derived based on controlling the power.
- Further work assessing the impact of non-normality on the parametric TOST and on the Binomial approach are also compeleting.

Reference

- U.S. Pharmacopoeia XXIV. (2005). Easton, Penn: Mack Printing Company.
- Williams R.L., et al. (2002). Content uniformity and dose uniformity: current approaches, statistical analyses, and presentation of an alternative approach. With special reference to oral inhalation and nasal drug products. *Pharmaceutical Research.* 19: 359-366.
- Tsong, Y., Shen, M., (2007). Parametric two-stage sequential quality assurance test of does content uniformity. *J Biopharm. Stat.* 17: 143-157.
- Sandell D., et al. (2006). Development of a content uniformity test suitable for large sample sizes. *Drug Information J.* 40: 337-344.
- D. B. Owen (1964). Control of percentages in both tails of the normal distribution. *Technometrics*. 6: 377-387.
- Berger R.L., Hsu J.C. (1996). Bioequivalence trials, intersection-union tests and equivalence confidence sets. *Statistical Science*. 11: 283-319.

Thanks!