Statistical Considerations in Setting Product Specifications

Xiaoyu (Cassie) Dong, Ph.D., M.S.
Joint work with Drs. Yi Tsong, and Meiyu Shen
Division of Biometrics VI, Office of Biostatistics, CDER, FDA

Manufacturing: Regulatory Impact of Statistics
2014 MBSW, Muncie, Indiana

Disclaimer

This Presentation reflects the views of the author and should not be construed to represent FDA's views or policies

Outline

I. Background

II. Statistical Methods to Set Spec.

1. Reference Interval
2. (Min, Max)
3. Tolerance Interval
4. Confidence Interval of Percentiles

III. Comparison at Large Samples

IV. Sample Size Calculation

V. Concluding Remarks

I. Background

- What are specifications?

  Specifications define quality standard/requirements.

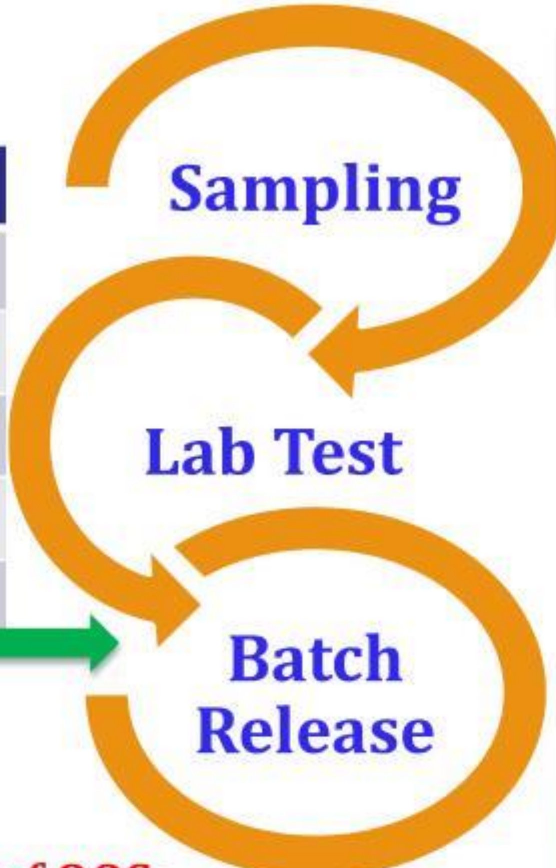
  ICH Q6A/B: a specification is defined as a list of tests, references to analytical procedures, and appropriate acceptance criteria, which are numerical limits, ranges, or other criteria for the tests described.

- It establishes the set of criteria to which a drug substance, drug product or materials at other stages of its manufacture should conform to be considered acceptable for its intended use.
- Specifications are one part of a total control strategy designed to ensure product quality and consistency.

I. Background (2)

- Specifications define quality standard/requirements.

Test	Specification
Assay	90-110%LC
Impurities	≤ 1%
Content Unif.	USP<905>
Dissolution	USP<711>
Microbial	≤ 2 %



Fail

Investigate root cause of OOS:
analytical error, process change, product change

I. Background (3)

- Specifications are important quality standards.
  - Only batches which satisfy specifications can be released to the market;
  - Provide a high degree of assurance that products are of good quality;
  - Assure consistent manufacturing process;
  - Most importantly, directly/indirectly link to product efficacy and safety;
  - Out-of-spec. (OOS) data are informative: analytical error, process change, product change

I. Background (4)

How specifications are determined?

Phase I/II

- May be used as supportive data

Phase III

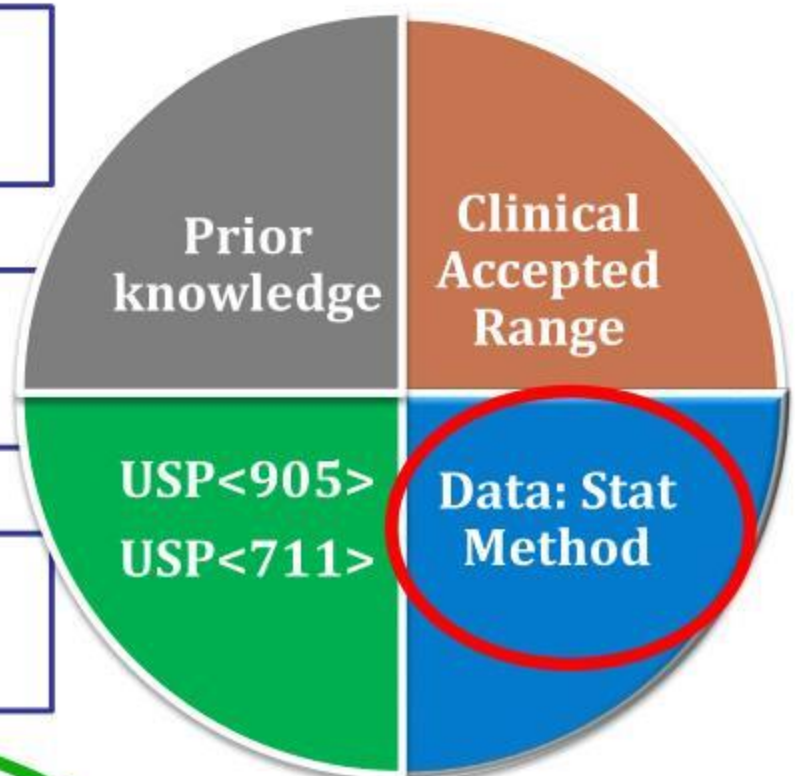
- **Use 3 clinical batches to set spec.**

NDA Submission

- Regulatory Approval

Post-marketing Changes

- Accumulated Data: release/stability



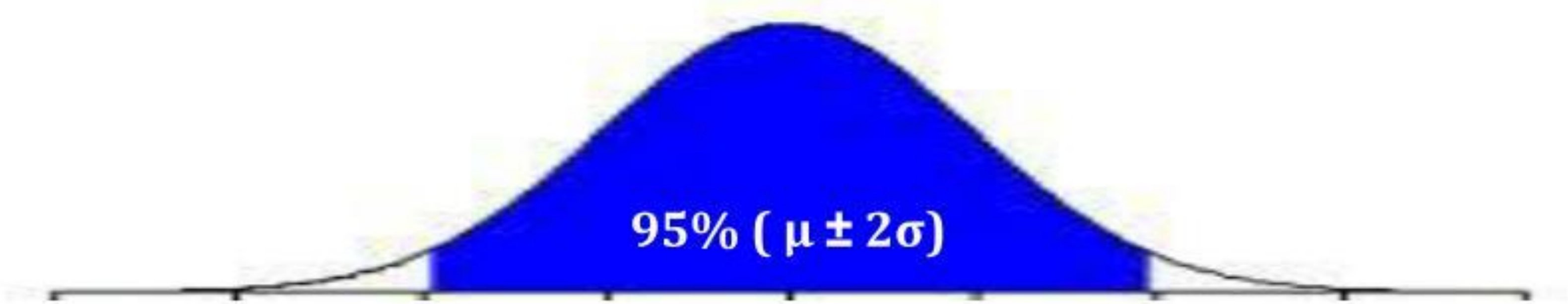
I. Background (5)

What are the impacts of setting inappropriate spec. ?

- **Too wide:**
  - Increase consumer's risk (release poor quality batches)
  - Product recalled or withdrawn from the market
  - Insensitive to detect process drifting/changes
  - Adverse impacts on patients
- **Too narrow:**
  - Increase manufacturer's risk (waste good quality batches)
- Thus, it is important to choose proper stat. method to set **meaningful, reasonable**, and **scientifically justified** specifications.

II. Statistical Methods to Set Spec.

- Assume test data $X \sim \mathrm{N}(\mu, \sigma^2)$
- σ^2:(Analytical + Sampling Plan + Manufacturing) Var
- True Spec. = Interval covering central p% of the population, say 95%.



- Use **limited data** from random samples/stability studies to **estimate** the underlying unknown interval.

II. Statistical Methods to Set Spec. (2)

- Commonly used methods in NDA submissions:
  - Reference Interval: $\bar{X} \pm 2SD$
  - (Min, Max)
  - Tolerance Interval: $\bar{X} \pm kSD$, k is (p%, 1-α%) tolerance factor
- Our proposal under study:
  - Confidence limits of Percentiles
- Compare: Coverage and Interval Width

II.1 Reference Interval

- Reference Interval (RI) = $\bar{X} \pm 2SD$
  - Most common method
  - Used in control chart to monitor process changes
- RI is not a reliable estimate for ($\mu \pm 2\sigma$) at small samples
  - Variability
  - Actual Coverage vs. Intended Coverage (95%)

II.1 Reference Interval (2)

- Variability of RI Limits:

$$Var(\bar{X} + Z_{1-p/2} \times S) = \frac{\sigma^2}{n} + Z_{1-p/2}{}^2 \sigma^2 (1 - C^{-2})$$

$$C = \sqrt{\frac{n-1}{2}} \Gamma(\frac{n-1}{2}) / \Gamma(\frac{n}{2})$$

	n = 10	n = 20	n = 50	n = 100
Std. Dev.	0.55	0.39	0.24	0.17
Upper Spec. Range	**(0.90, 3.1)**	(1.22,2.78)	(1.52, 2.48)	**(1.66, 2.34)**
True Upper Spec.	2			

Table 1 – Approx. Ranges of Upper Specification Limit Estimated using Reference Interval Method

II.1 Reference Interval (3)

- Actual vs. Intended Coverage (95%):

Table 2 – Quantiles of Coverage from 10^5 Simulations using Reference Interval Method with Intended Coverage of 95%

Quantiles	n = 10	n = 20	n = 50	n = 100	n = 1,000
Min Cover.	27.2	46.9	75.2	84.3	92.6
25%	86.9	90.6	92.8	93.6	94.6
50%	92.9	94	94.6	94.8	95
75%	96.6	96.4	96.1	95.9	95.3
Max Cover.	100	99.9	99.5	99	96.8

II.2 (Min, Max)

- Specification = (Min, Max) of Obs.
- Not suitable to define spec. :
  - coverage can't be defined.
  - Insensitive to identify OOS obs. as "atypical" or "abnormal" results.
  - With small samples, neither the manufacturer's risk nor the consumer's risk is clear;
  - with large samples, consumer's risk will be greatly inflated due to over-wide spec.

II.2 (Min, Max) (2)

- Spec. = (Min, Max) , say intended coverage = 95%

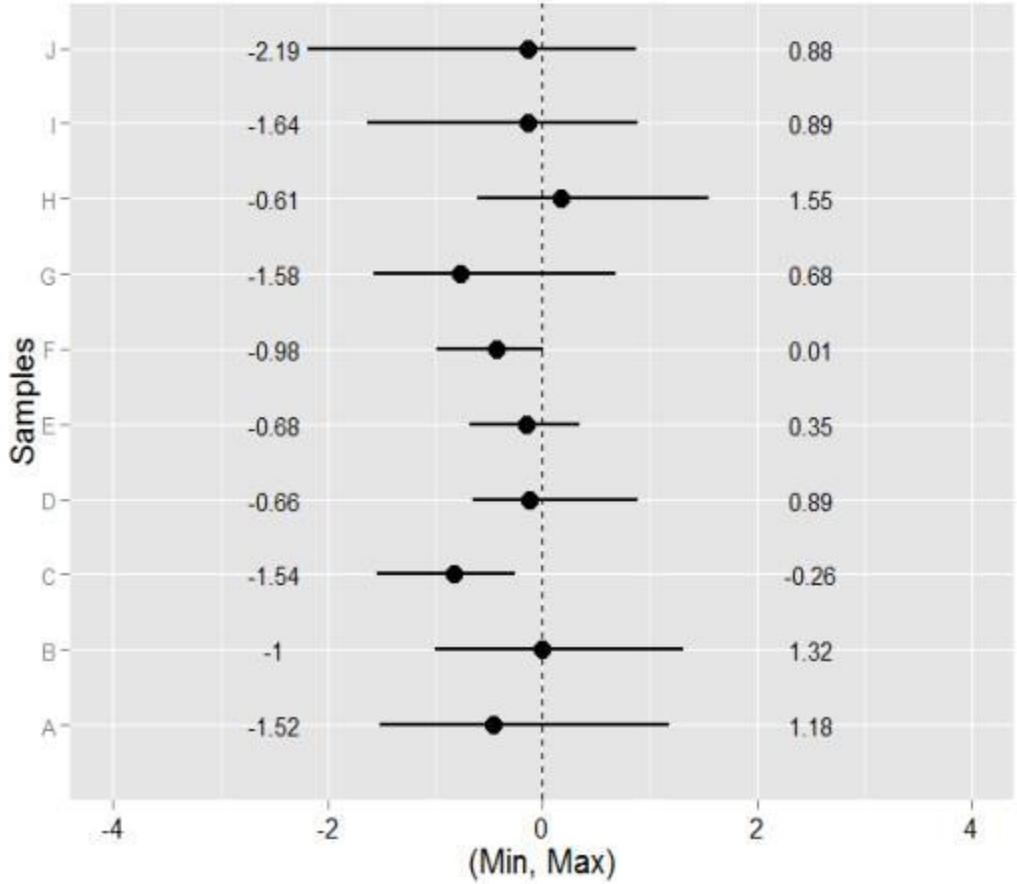


Figure 1 – Plots of (Min, Max) of 10 Simulations with N = 5 from N(0,1)

Coverage
80%
76%
67%
69%
34%
38%
56%
34%
75%
82%

II.2 (Min, Max) (3)

- Spec. = (Min, Max) , say intended coverage = 95%

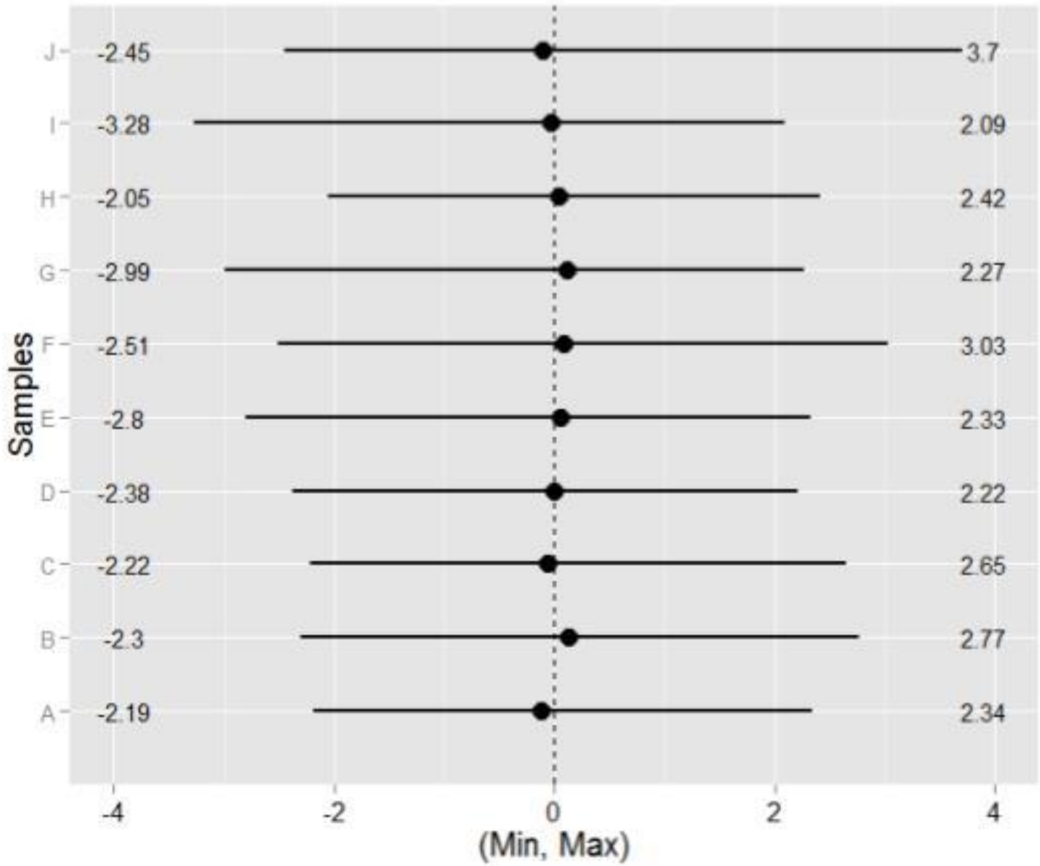
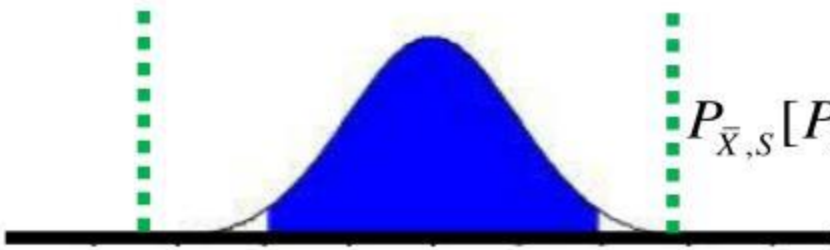


Figure 2 – Plots of (Min, Max) of 10 Simulations with N = 100 from N(0,1)

Coverage
99.3%
98.1%
97.2%
98.7%
99.3%
98.8%
97.8%
98.3%
98.6%
97.6%

II.3 Tolerance Interval

- Spec. = Tolerance Interval (TI)= Mean ± *k* ×SD
  - Aims to cover **at least** *p*% (say 95%) of the population with conf. level of 1-α.



$$P_{\bar{X},S}[P_X(\bar{X} - kS < X < \bar{X} + kS \mid \bar{X}, S) \geq \beta] = 1 - \alpha$$

  - *k* is (*p*, 1-α) tolerance factor $k = t_{n-1,\gamma}(Z_p\sqrt{n})/\sqrt{n}$

- By definition, TI is almost always **wider than** the targeted interval, especially with small samples.

	n = 5	n = 10	n = 50	n = 100	n = 1000
95%, Z_p = 2.00	4.91	3.40	2.43	2.28	2.05

II.3 Tolerance Interval (2)

- Tolerance interval has issues of over-coverage

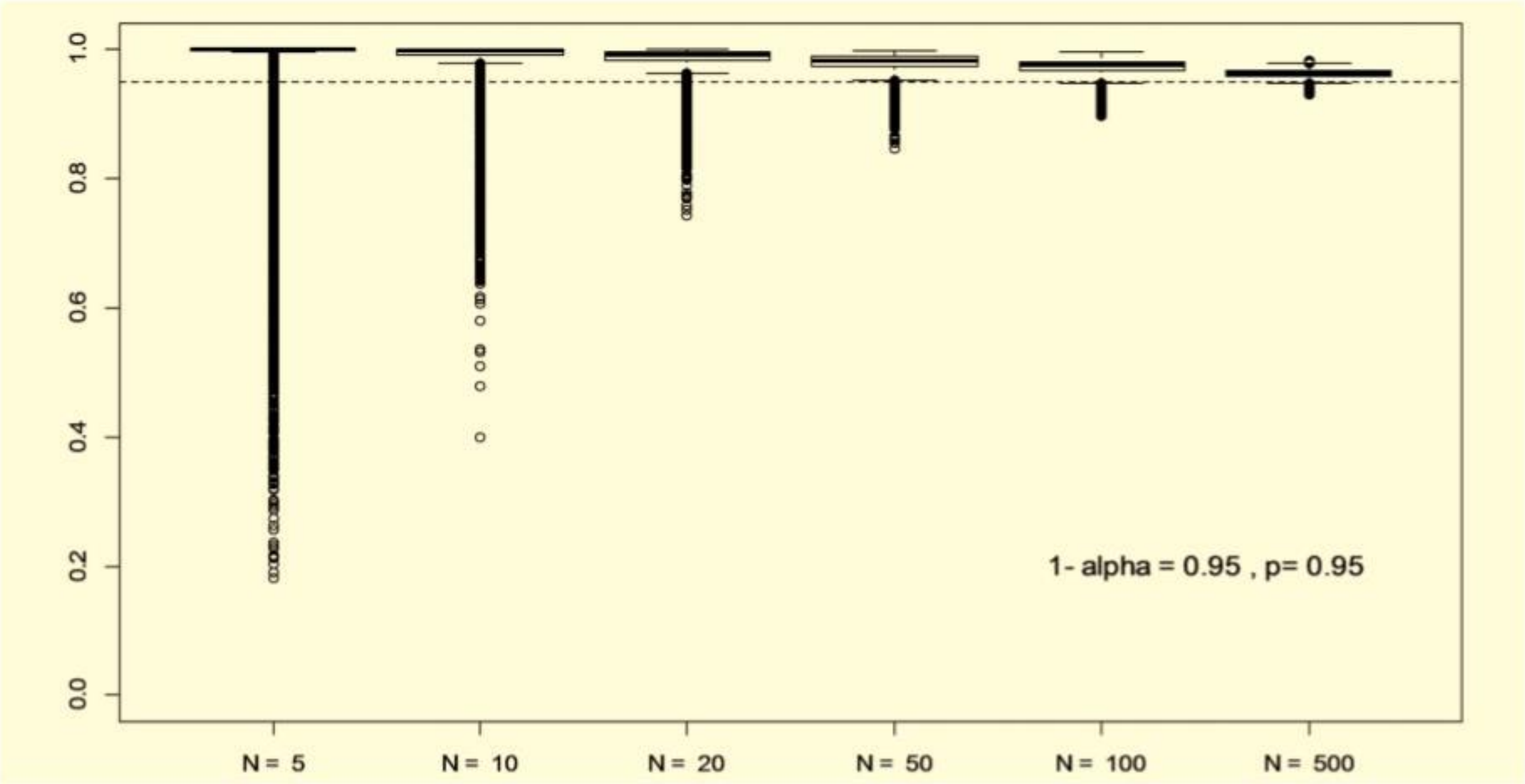


Figure 3 – Box Plots of Coverage Obtained from 10^5 Simulations using Tolerance Interval

II.3 Tolerance Interval (3)

- Tolerance interval is too wide

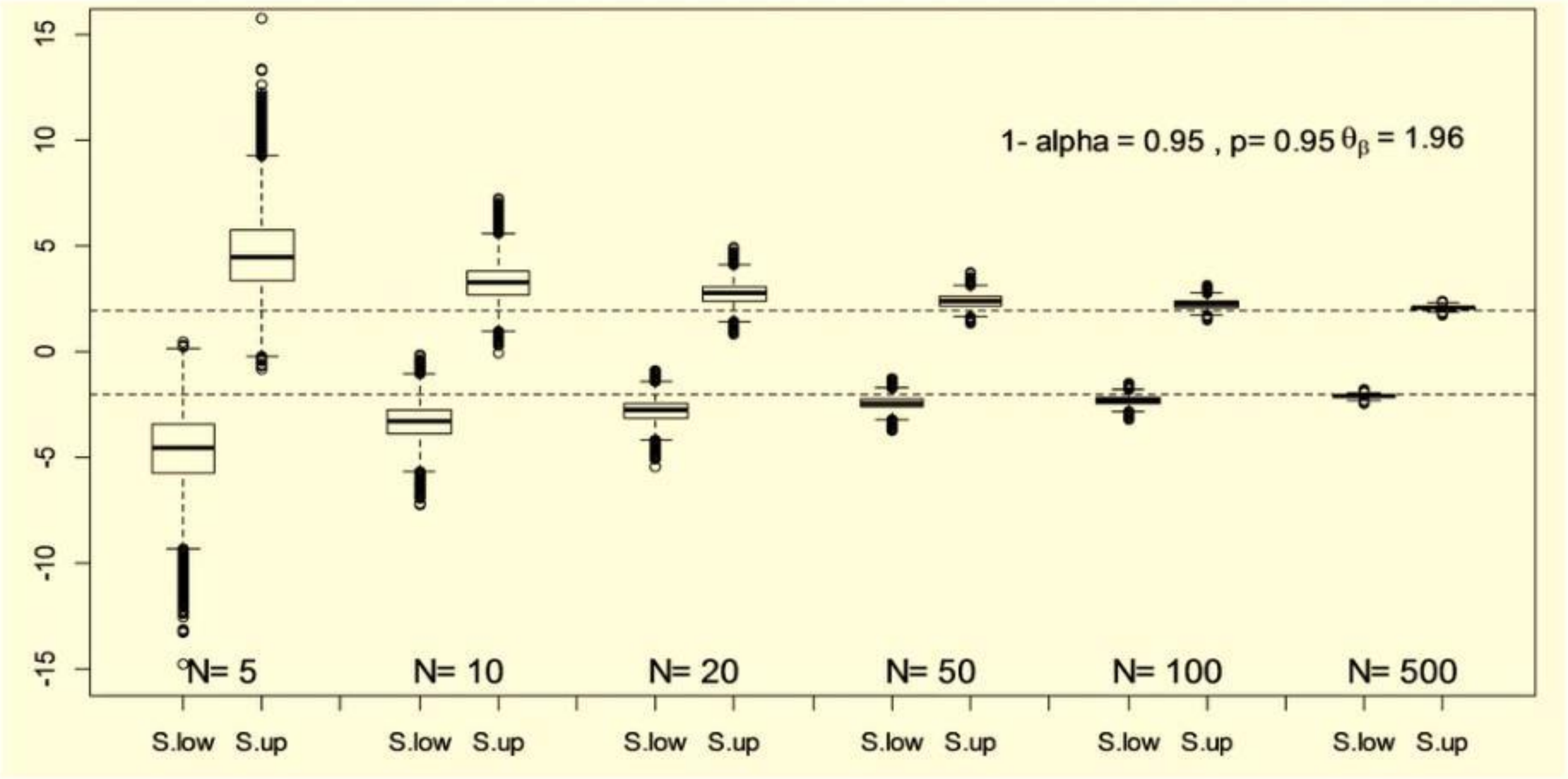


Figure 4 – Box Plots of Lower and Upper Bounds Obtained from 10^5 Simulations using Tolerance Interval

II.4 Confidence Limits of Percentiles

- Basic Idea

Tolerance Interval

Lower CL of LB, Upper CL of UP

$$\mu - Z_p\sigma\,, \qquad \mu + Z_p\sigma$$

(UCL of LB, LCL of UB)

Conf. Limits of Percentiles

II.4 Confidence Limits of Percentiles (2)

- Let the true interval be

$$\left(\theta_{low} = \mu - Z_p\sigma, \theta_{up} = \mu + Z_p\sigma\right)$$

- 1-α upper CL of θ_{low}:

$$\theta_{lowCL} = \hat{\theta}_{low} + Z_{1-\alpha}\sqrt{\hat{\text{var}}(\hat{\theta}_{low})} = (\bar{X} - CZ_pS) + Z_{1-\alpha} \times \frac{S}{\sqrt{n}}\sqrt{1 + nZ_p^2(C^2 - 1)}$$

- 1-α Lower CL of θ_{up}:

$$\theta_{upCL} = \hat{\theta}_{up} - Z_{1-\alpha}\sqrt{\hat{\text{var}}(\hat{\theta}_{up})} = (\bar{X} + CZ_pS) - Z_{1-\alpha} \times \frac{S}{\sqrt{n}}\sqrt{1 + nZ_p^2(C^2 - 1)}$$

II.4 Confidence Limits of Percentiles (3)

- Intended Coverage = 95%

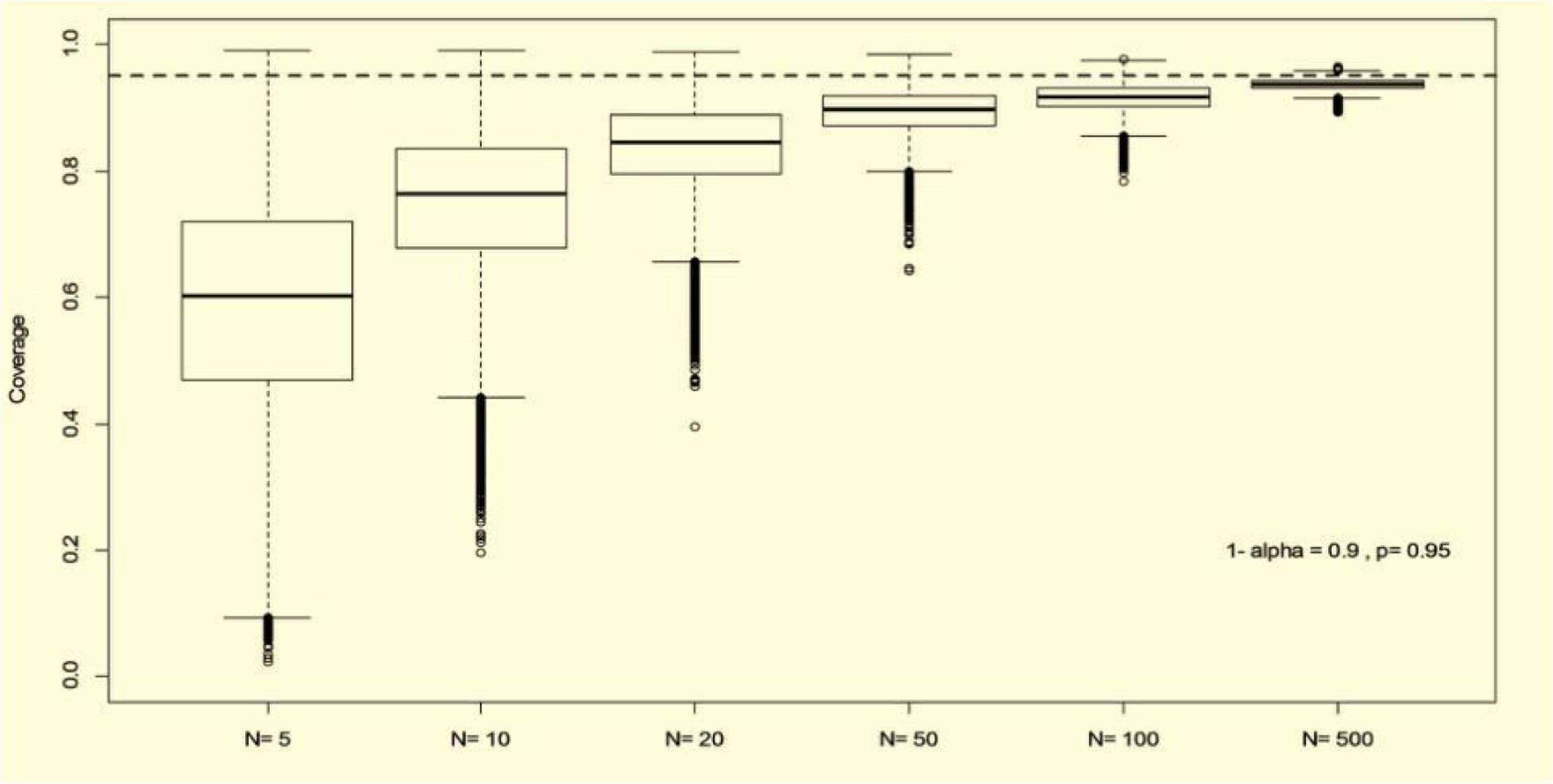


Figure 5 – Box Plots of Coverage from 10^5 Simulations using Conf. Limits of Percentiles

II.4 Confidence Limits of Percentiles (4)

- Intended interval= (-1.96, 1.96)

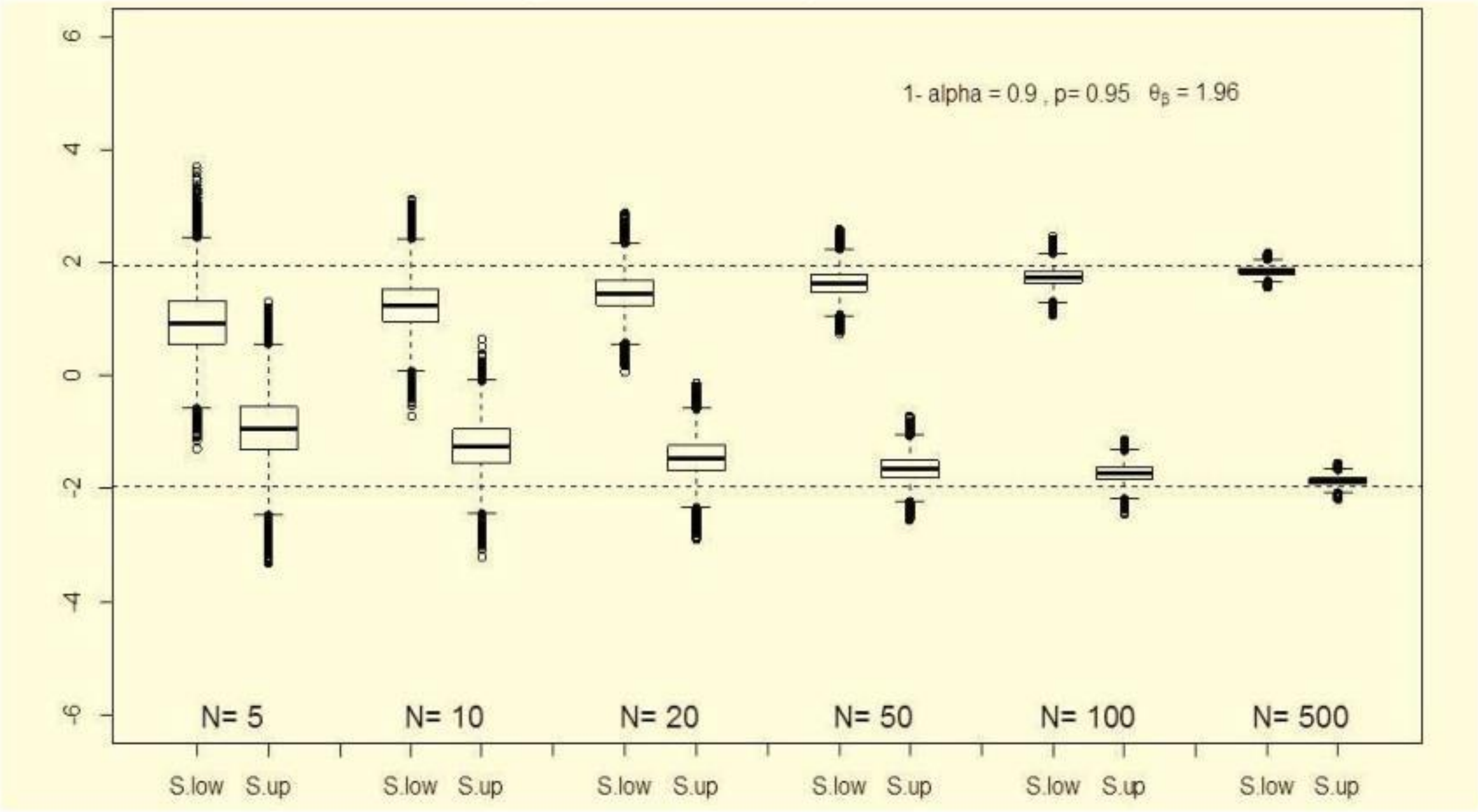


Figure 6 – Box Plots of Lower and Upper Bounds from 10^5 Simulations using Conf. Limits of Percentiles

III. Comparisons at Large Samples

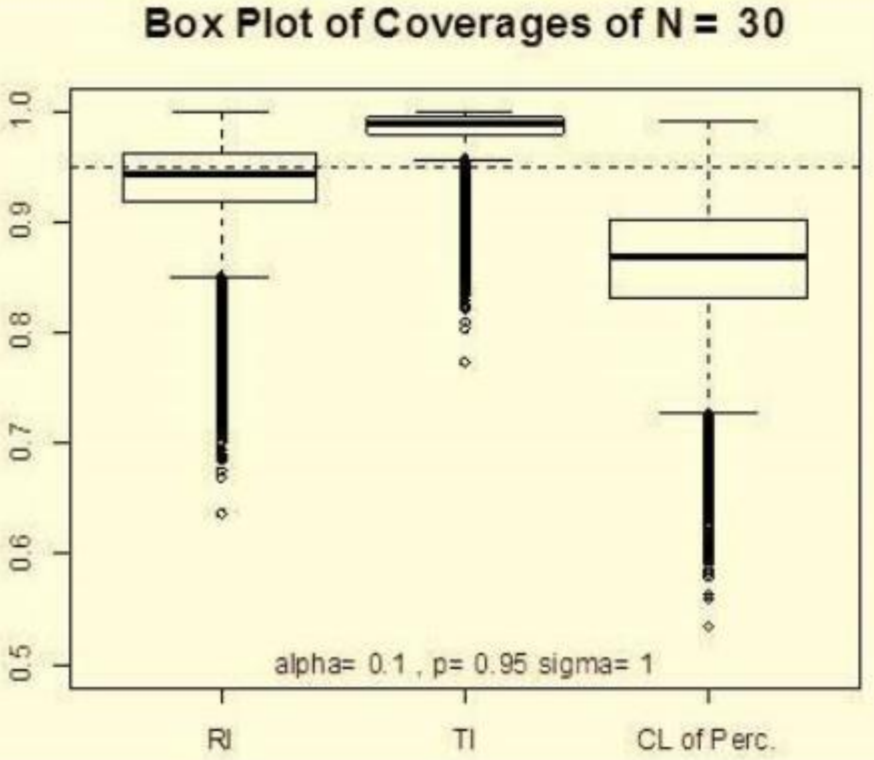
Small Samples	Reference Interval	(Min, Max)	Tolerance Interval	Conf. Limits of Percentiles
Coverage	Not assured	Not assured	$\geq$ p%	< p%
Interval Width	Large Var.	Large. Var.	Too wide	Too narrow
Our RECOM.	✗	✗	✗	✗

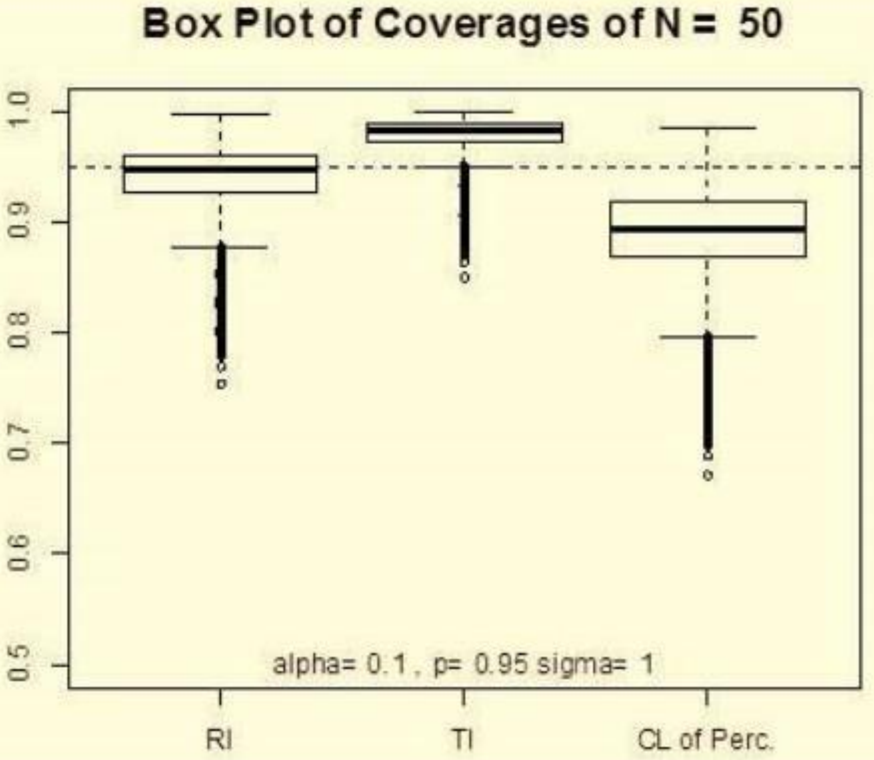
It is not suitable to set specification when small sample sizes are small, especially when the data variability is large.

What about large samples?

III. Comparisons at Large Samples (2)

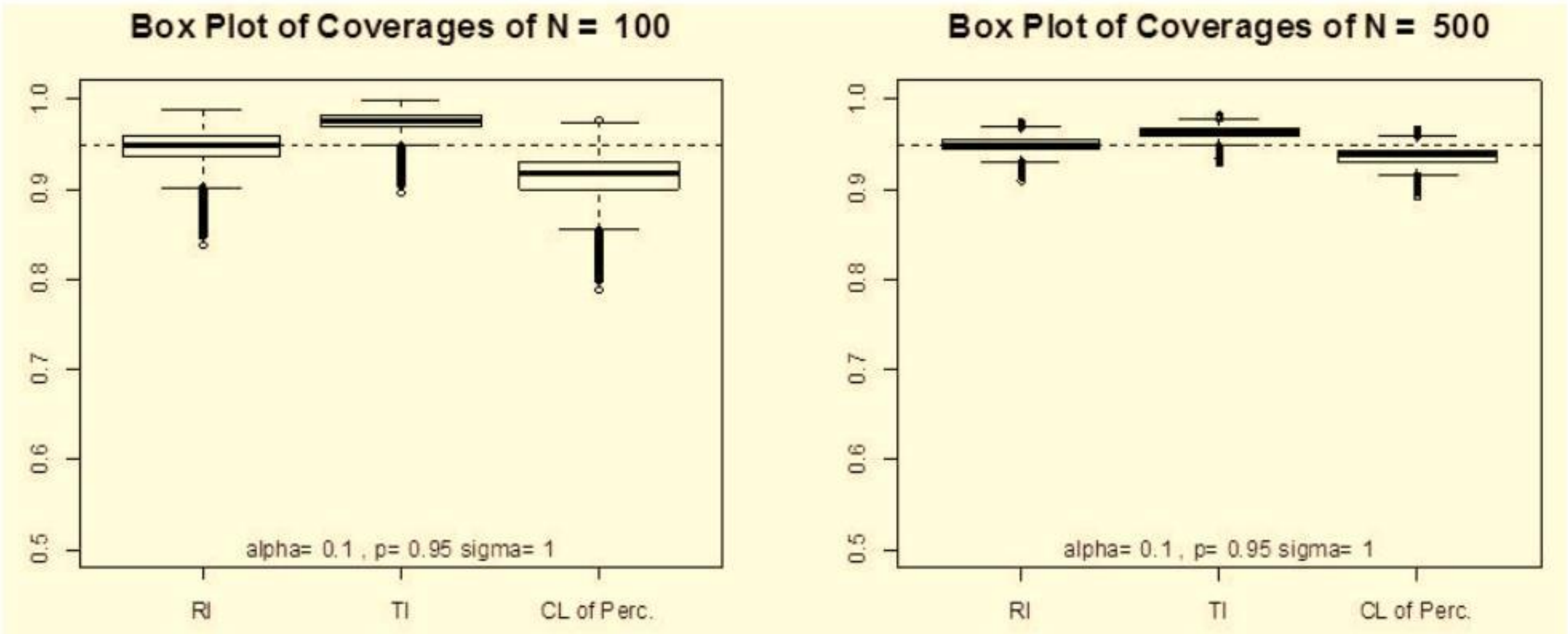
- Intended Coverage (95%)





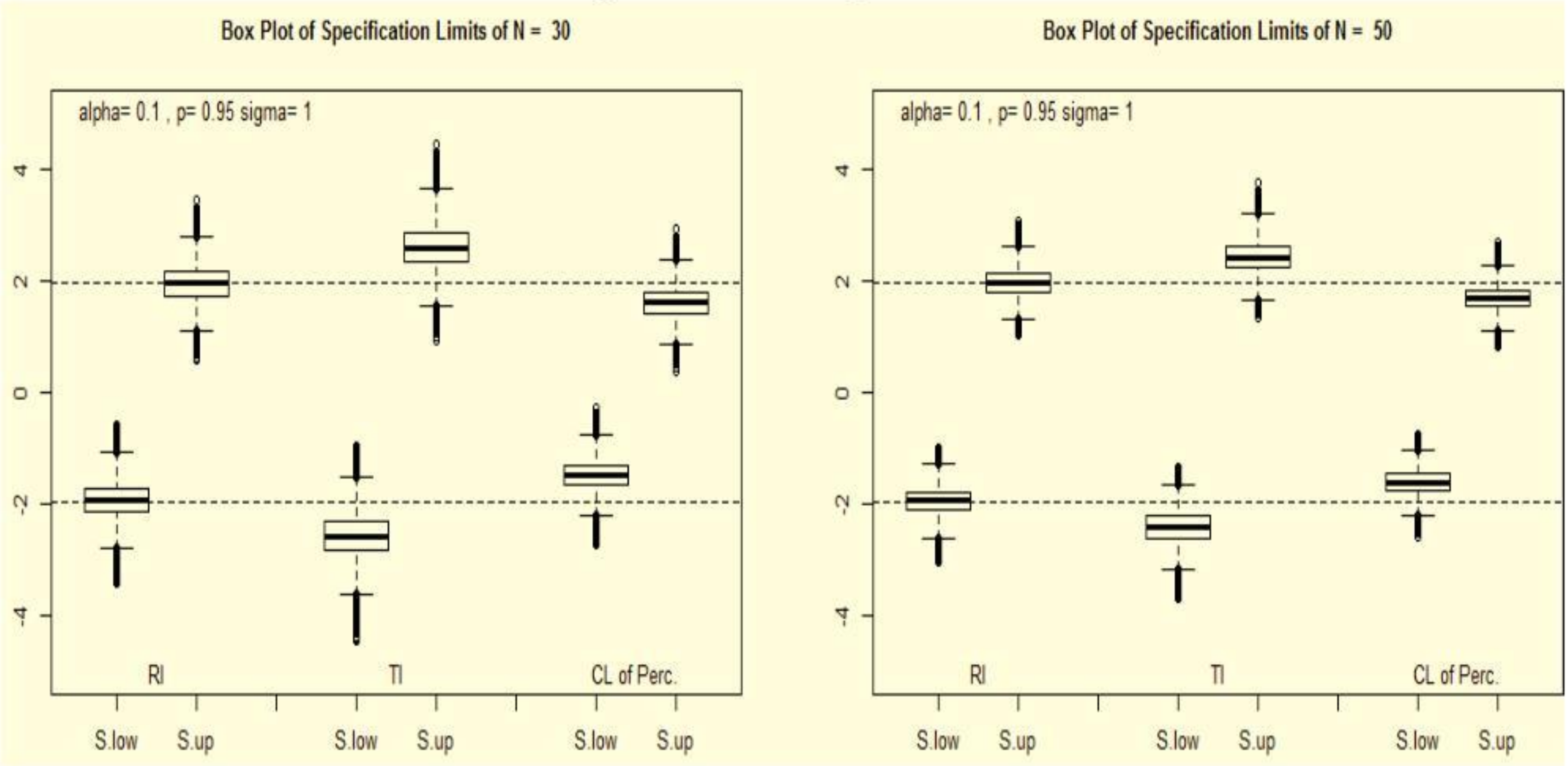
III. Comparisons at Large Samples (3)

- Intended Coverage (95%)



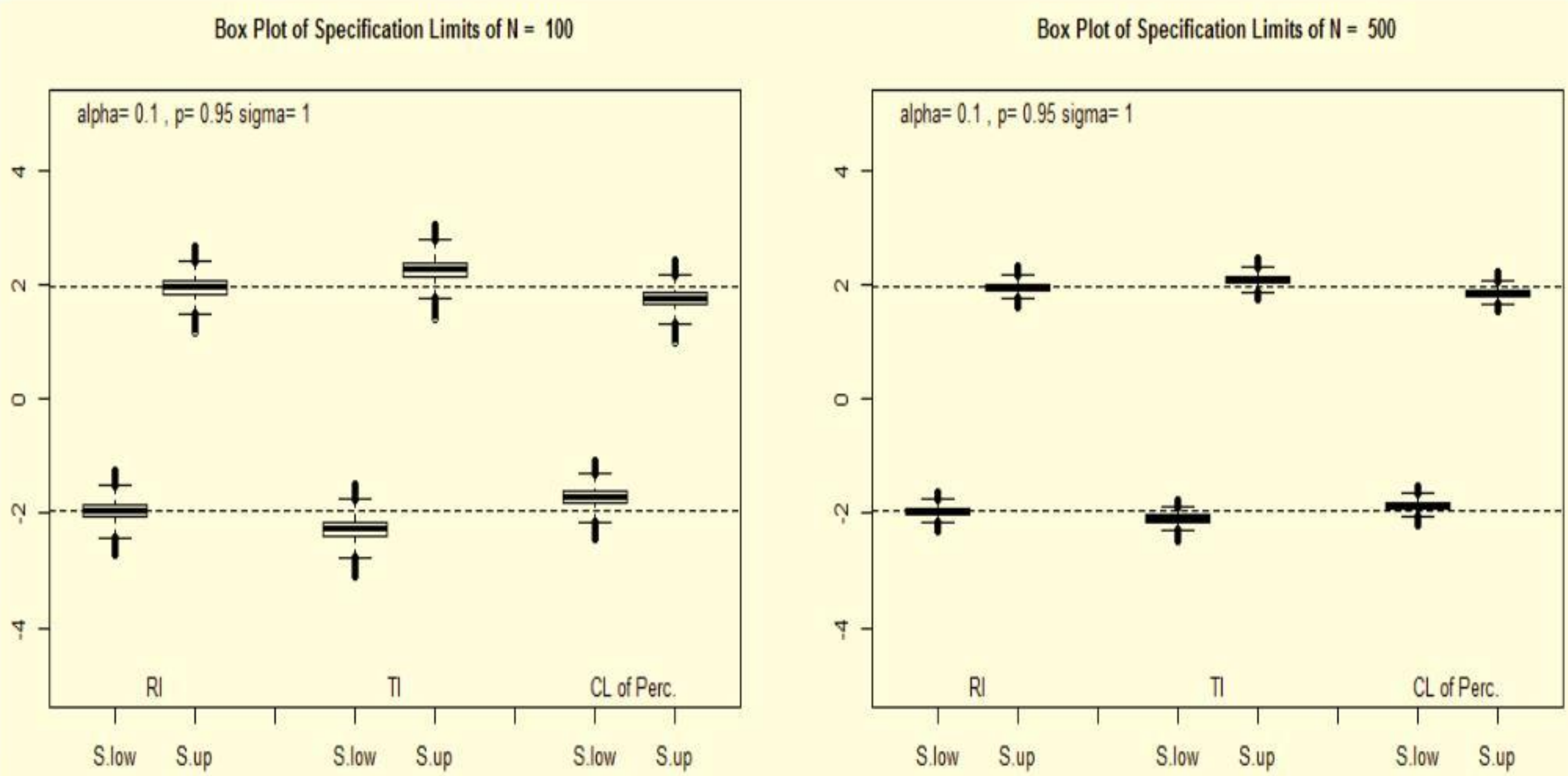
III. Comparisons at Large Samples (4)

- Intended Interval = (-1.96, 1.96)



III. Comparisons at Large Samples (5)

- Intended Interval = (-1.96, 1.96)



III. Comparisons at Large Samples (6)

- Inflation of Consumer's Risk: too wide

outside (μ ± 2σ)	Within Spec.	Outside Spec.
Poor Quality Product	**Pass**	Fail

$$P_{consumr\ \text{inflate}} = I(\hat{\theta}_{low} < \theta_{low})\Pr(\hat{\theta}_{low} < X < \theta_{low} \mid \hat{\theta}_{low}) + I(\hat{\theta}_{up} > \theta_{up})\Pr(\theta_{up} < X < \hat{\theta}_{up} \mid \hat{\theta}_{up})$$

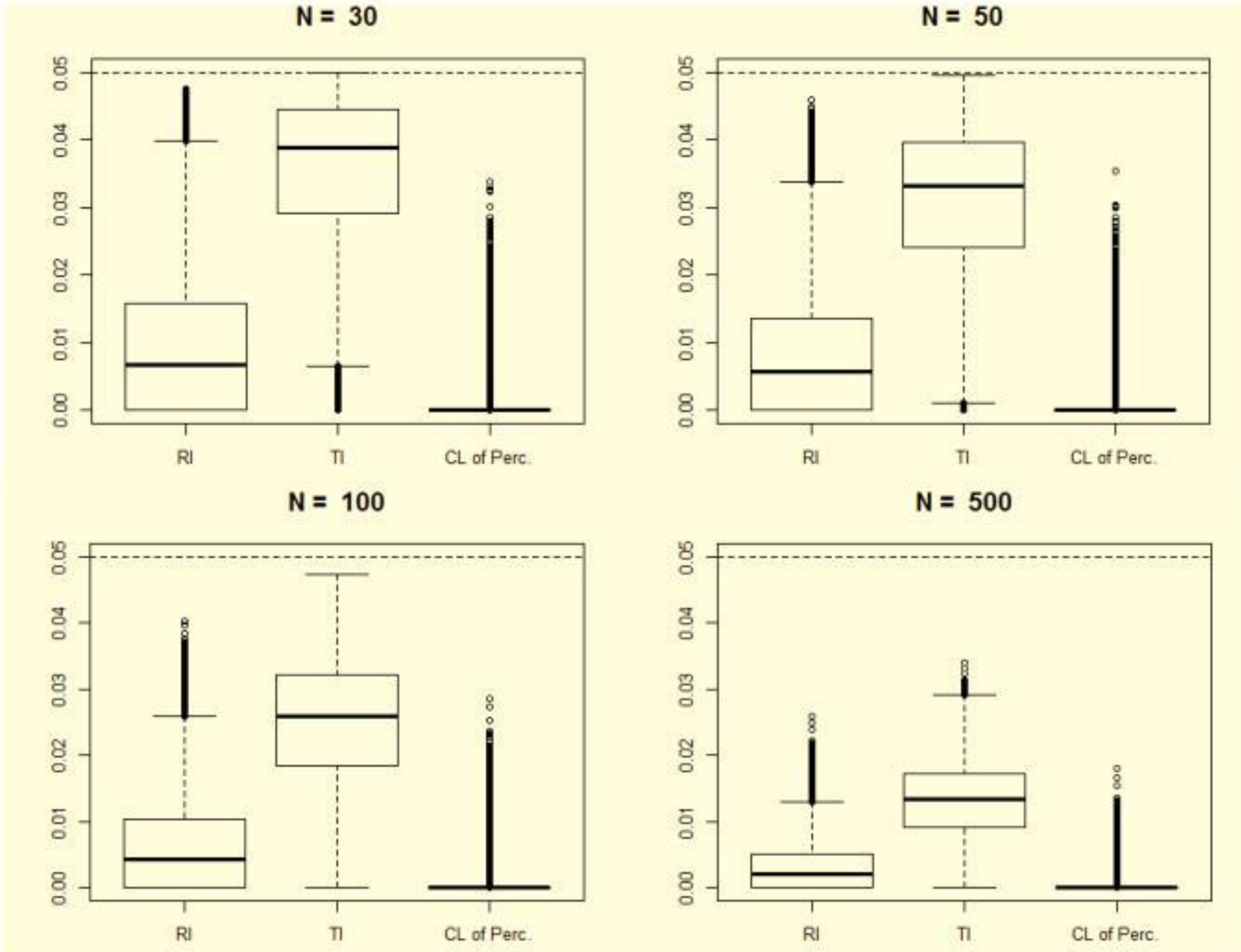
- Inflation of Manufacturer's Risk: too narrow

within (μ ± 2σ)	Within Spec.	Outside Spec.
Good Quality Product	Pass	**Fail**

$$P_{\text{manufacture inflate}} = I(\hat{\theta}_{low} > \theta_{low})\Pr(\theta_{low} < X < \hat{\theta}_{low} \mid \hat{\theta}_{low}) + I(\hat{\theta}_{up} < \theta_{up})\Pr(\hat{\theta}_{up} < X < \theta_{up} \mid \hat{\theta}_{up})$$

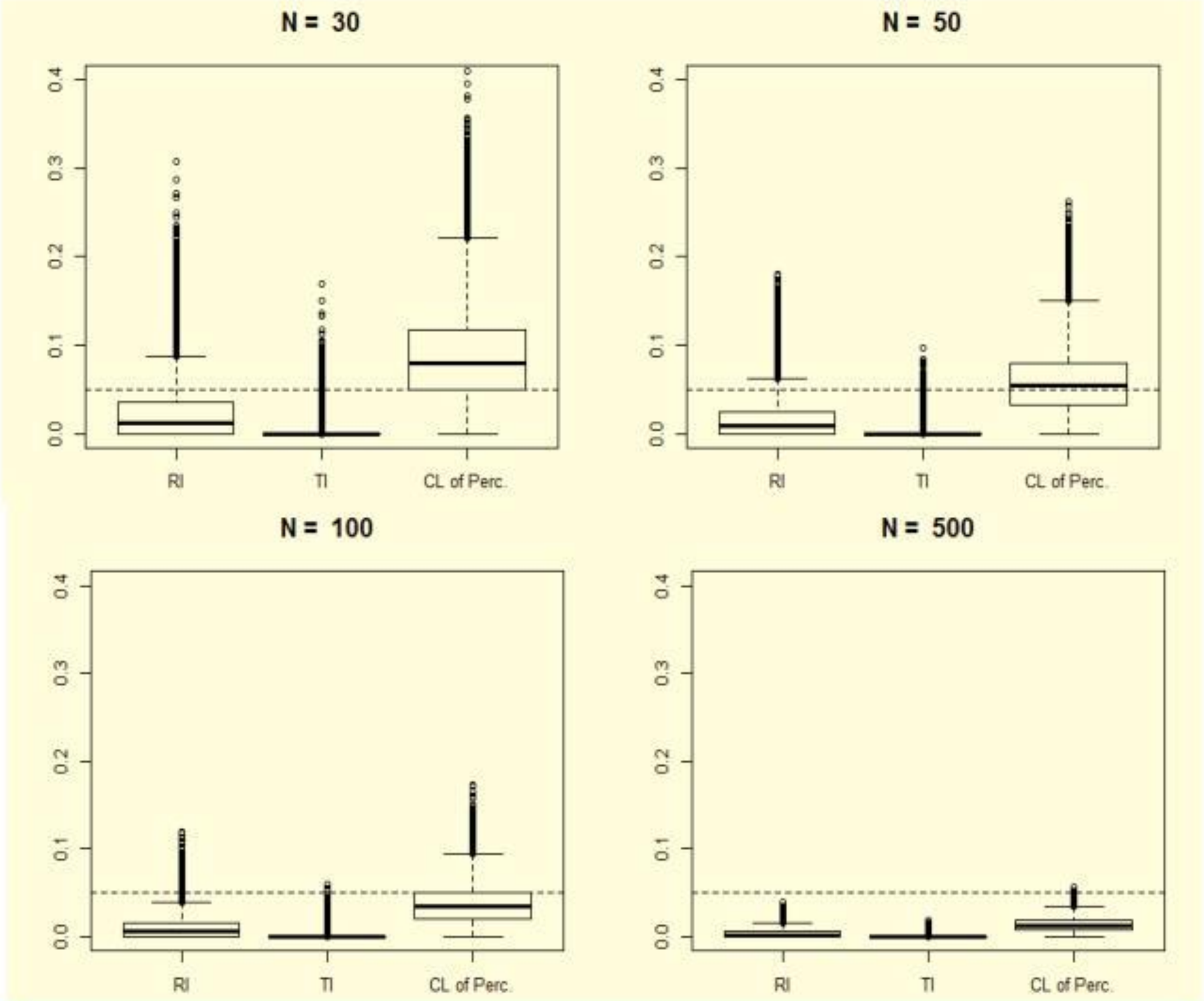
III. Comparisons at Large Samples (7)

- Inflation of Consumer's Risk: release the poor quality product



III. Comparisons at Large Samples (8)

- Inflation of Manufacturer's Risk: waste the good quality product



IV. Sample Size Calculation

- It would be helpful if we can plan the sample size of setting spec. in advance.
- Similar concept of SS calculation used in TI methods;
- Compute sample size so that

$$P_{\bar{X},S}\left[\, p-\delta \le P_X(\bar{X}-kS < X < \bar{X}+kS \mid \bar{X},S) \le p+\delta \,\right] \ge \gamma$$

Take p = 95%, δ = 3%, γ = 90% for example: Determine the sample size so that 90% (γ) of time, the absolute distance between the actual coverage and the targeted value of 95% (p) is less than 3% (δ).

IV. Concluding Remarks

Small Samples	Reference Interval	(Min, Max)	Tolerance Interval	Conf. Limits of Percentiles
Coverage	Not assured	Not assured	≥ p%	< p%
Interval Width	Large Var.	Large. Var.	Too wide	Too narrow
Our RECOM.	✗	✗	✗	✗

Large Samples	Reference Interval	(Min, Max)	Tolerance Interval	Conf. Limits of Percentiles
Coverage	Close to p%	~100%	≥ p%	< p%
Interval Width	Close to Target	Too Wide	Close to Target (Wider)	Close to Target (Narrower)
Our RECOM.	✓	✗	✓	✓

IV. Concluding Remarks (2)

- Specifications are a critical element of a total control strategy;
- Statistical considerations are important to set reasonable specifications in order to ensure quality, efficacy and safety of products at release and during the shelf life;
- When setting specifications, consumer's risk should be well controlled.
- Large sample size can't fix the issues caused by the underlying statistical concept of each method.
- Keep in mind, specifications estimated by statistical methods are subject to scientific or clinical justification.

Acknowledgment

- Dr. Yi Tsong
- Dr. Meiyu Shen
- Dr. Youngsook Lee
- Chemists and Biologists I have worked with.

References

- Chakraborti, S. and Li, J. (2007): Confidence Interval Estimation of a Normal Percentile, *The American Statistician*, 61:4, 331-336.
- ICH Guideline Q6A. (1999). “Specifications: test procedures and acceptance criteria for new drug substances and new drug procedures - chemical substances”.
- ICH Guideline Q6B. (1999). “Specifications: test procedures and acceptance criteria for biotechnological/biological products”.
- EBE Concept Paper (2013), “Considerations in Setting Specifications”, European Biopharmaceutical Enterprises.
- EMA (2011), “Report on the expert workshop on setting specifications for biotech products”. European Medicines Agency, London, 9 September 2011.
- Thomas, J. DiFeo. (2003). Drug product development: a technical review of chemistry, manufacturing, and controls information for the support of pharmaceutical compound licensing activities. *Drug Development and Industrial Pharmacy*. 29: 939-958.
- Kumar, R. and Palmieri Jr, M. J. (2009). Points to consider when establishing drug product specifications for parenteral microspheres. *The AAPS Journal*. 12: 27-32.

References

- Wessels, P., Holz, M., Erni, F., Krummen, K., and Ogorka, J. (1997). Statistical evaluation of stability data of pharmaceutical products for specification setting. *Drug Development and Industrial Pharmacy*. 23: 427-439.
- Yang, H. (2013). Setting specifications of correlated quality attributes through multivariate statistical modelling. *Journal of Pharmaceutical Science and Technology*. 67: 533-543.
- Shen, M., Dong, X., Lee, Y. and Tsong, Y. (2014). Statistical evaluation of several methods for cut point determination of immunogenicity screening assay. (Under peer review)
- Faulkenberry, G.D. and Daly, J.C. (1970). Sample size for tolerance limits on a normal distribution. *Technometrics*, 12, 813-821.
- Faulkenberry, G.D. and Weeks, D. L. (1968). Sample size determination for tolerance limits. *Technometrics*, 43, 147-155.
- Guenther, W.C. (1972). Tolerance intervals for univariate distributions. *Naval Research Logistic Quarterly*, 19, 309-333.
- Krishnamoorthy, K. and Mathew, T. (2009). Statistical tolerance regions – Theory, Applications and computations. *Wiley series in probability and statistics*.

Thank you!