Process Performance Analysis for Roche's Pharmaceutical Manufacturing Network

Yiming Peng, Theo Koulis, Jens Lamerz, and Dan Coleman
Nonclinical Biostatistics
Genentech, A Member of the Roche Group

2017 ASA Biopharmaceutical Section Nonclinical Biostatistics Conference
Rutgers University, New Jersey
Jun 12 - 14, 2017

Contents

Motivation

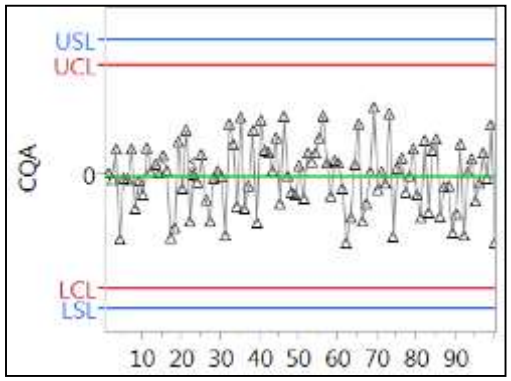
Objective	Characterizing our global manufacturing network and identifying opportunities for continuous process improvement through risk ranking of Critical Quality Attributes **(CQA)**
Client	Senior management and Health Authorities
Scope	• CQA from large and small molecules, drug substance **(DS)** and drug product **(DP)** • ~30 Products • ~2 Sites per Product • ~20 CQAs per Product • ~20-200 batches per CQA per site per year

How to summarize 1,200 time series per year?

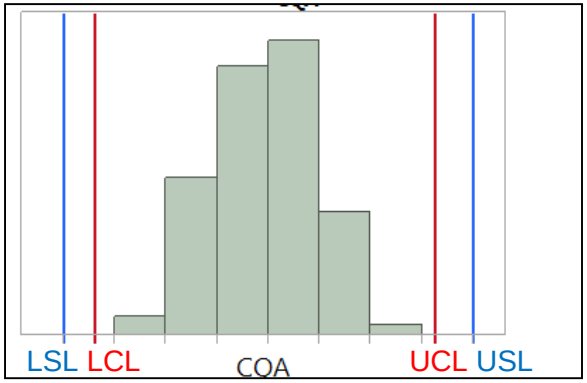
Summarizing Time Series

High Information → **(Very) Low Information**

Time Series



Histogram

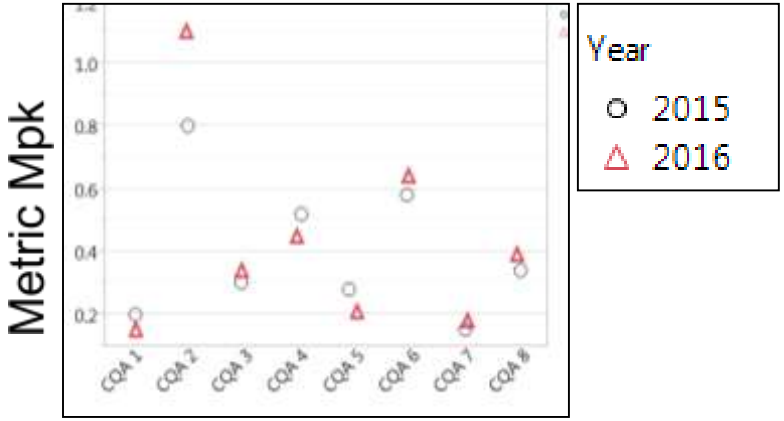


Metric

Mpk = 1.36

Metrics a necessary evil?

Metrics Enable Comparisons of Time Series



Strategy for Characterizing a Network

Idea

- Use metrics to risk rank CQAs
- Drill down to control charts for potentially at risk CQAs

Metric

- Simple and interpretable (2)
- Identify high risk processes (3)
- Enable continuous improvement (3)
- Require minimal/no assumption (4)
- Suitable for automation (5)

2. Roche’s process performance index, Rpk

Simple and Interpretable

Risk index Rpk

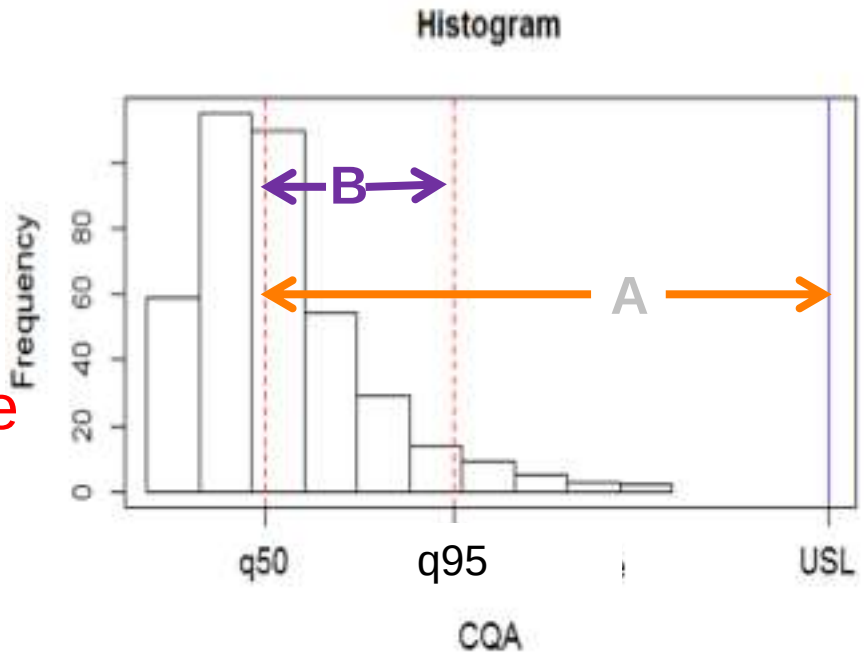
- R_{pk} consistent description of process performance that does not depend on distribution (Clement (1989))

- R_{pk} proportion of the allowable range used by the process
  - R_{pk} is typically between 0 and 1
  - 0 indicates high process performance
  - 1 indicates low process performance
  - Small is Beautiful, Big is Bad!

- Use (k=u, l, 2)
  - R_{pu} for attributes with upper specific. limits, e.g., impurities
  - R_{pl} is for attributes with lower spec. limits, e.g., monomer
  - R_{p2} is for attributes with upper and lower spec. limits, e.g., potency

R_{pu} definition (u = upper specification limit)

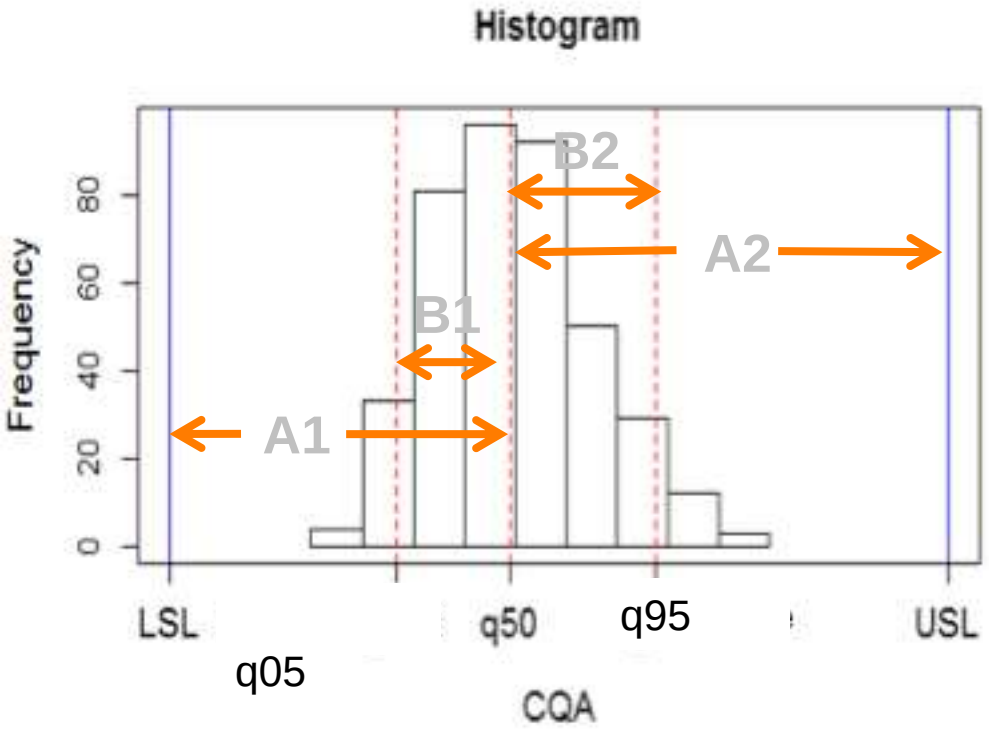
- B = Bulk of data, Percentage of the data between q50 and q95
- A = Allowable Range, interval between q50 and USL
- R_{pu} = the proportion of allowable range used by the bulk of the data
- R_{pu} = B/A



q50 = 50th percentile (median)
q95 = 95th percentile
USL = Upper Specification Limit

R_{p2} definition (2 = two-sided specification)

- R_{p2} = max{B1/A1, B2/A2}
- R_{p2} is the largest proportion of allowable range used by a bulk of the data



LSL = Lower Specification Limit
q05 = 5th percentile
q50 = 50th percentile (median)
q95 = 95th percentile
USL = Upper Specification Limit

3. Pilot study results

Identify high risk processes

Enable continuous improvement

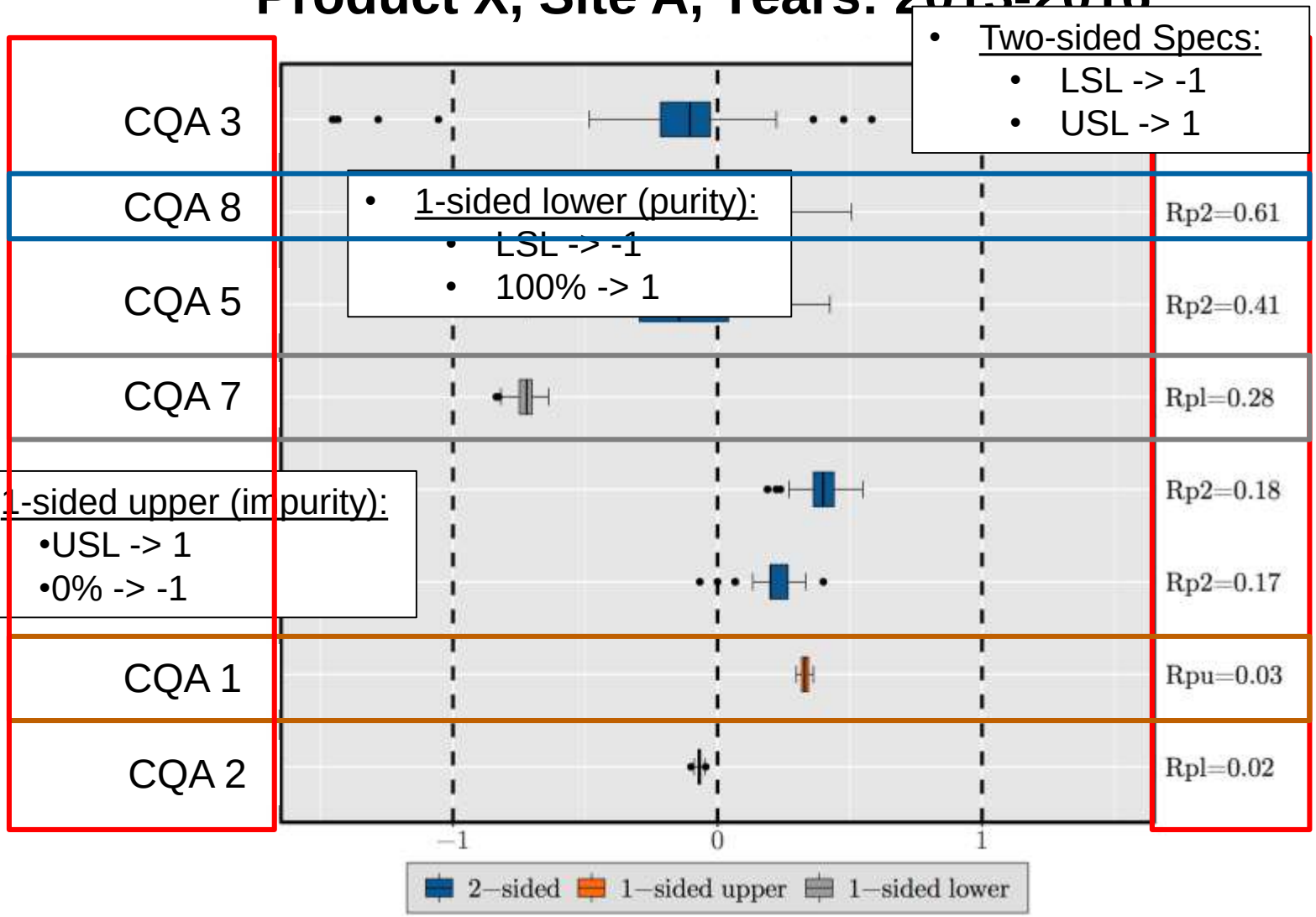
Pilot Study

- Data:
  - 9 products
    - 5 Biologics, 4 Small Molecule
  - DS and DP
  - Multiple years
  - Multiple sites

- Analysis
  - PPA at site level (single product, single site)
  - PPA at product level (single product, multi-site)
  - PPA at network level (multi-product, multi-site)

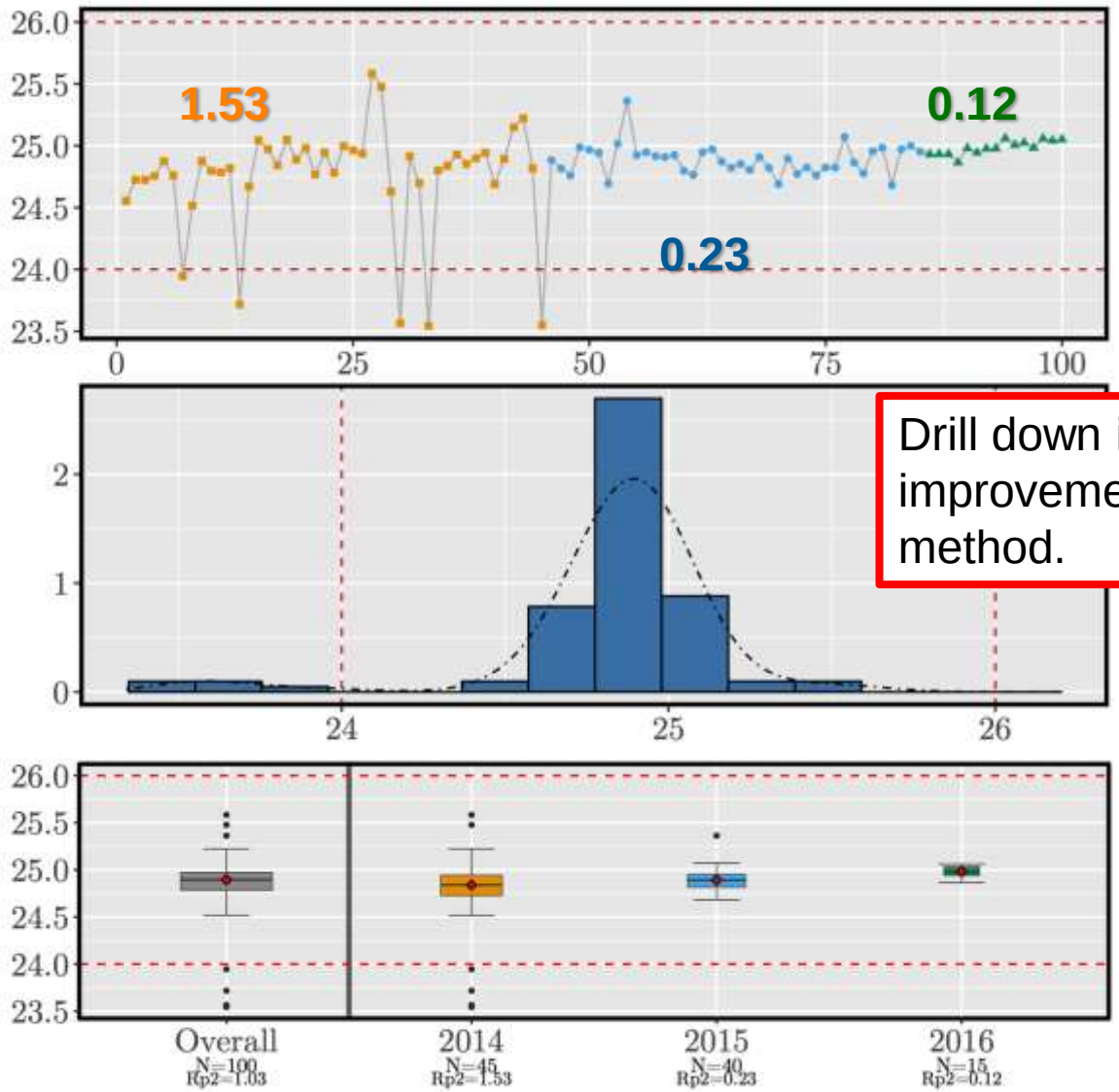
Site level: Risk-Ranking Overview

Product X, Site A, Years: 2013-2016



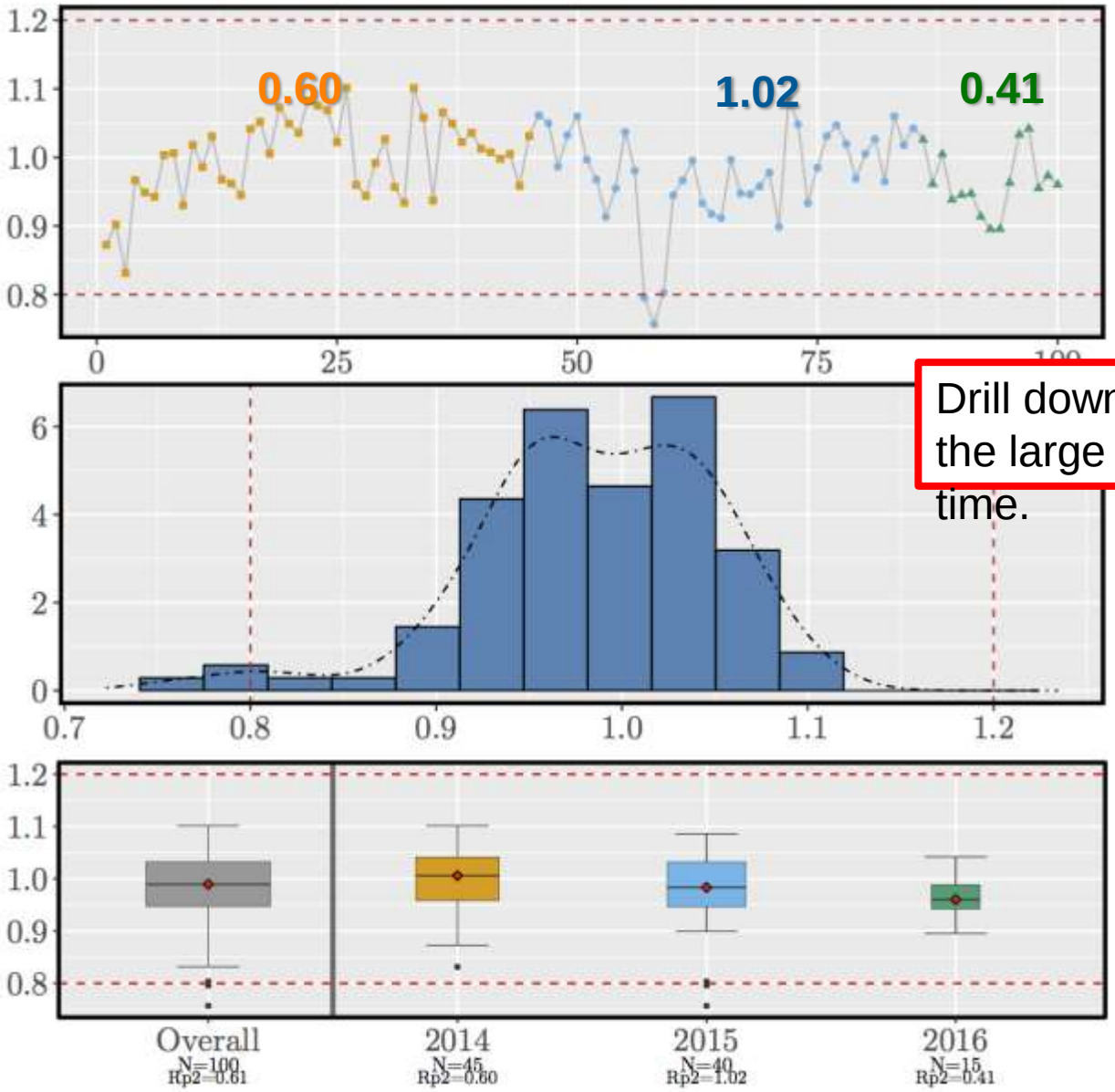
- Two-sided Specs:
  - LSL -> -1
  - USL -> 1
- 1-sided lower (purity):
  - LSL -> -1
  - 100% -> 1
- 1-sided upper (impurity):
  - USL -> 1
  - 0% -> -1

Drill Down: CQA 3 by Year



Drill down is needed to see the improvement in the analytical method.

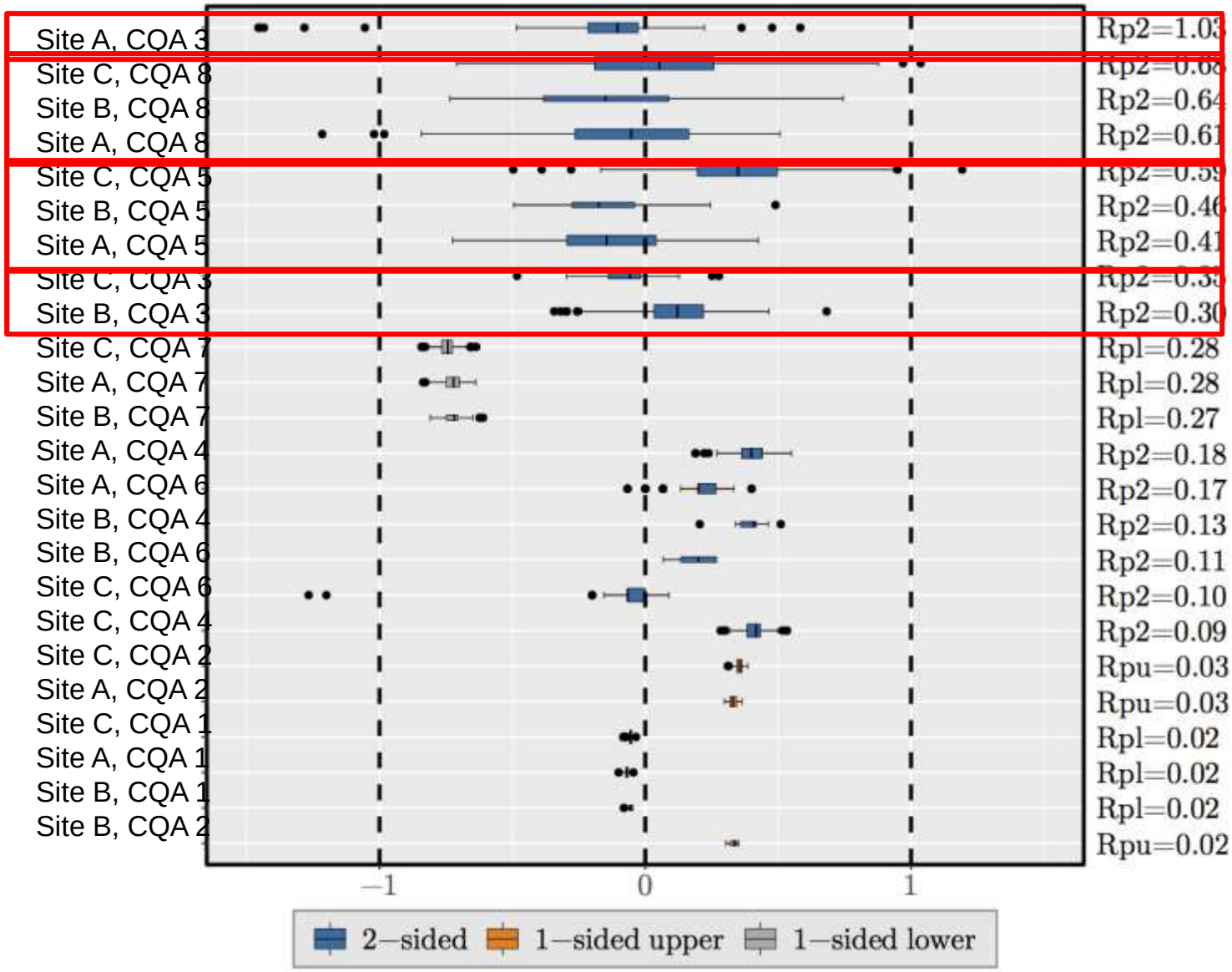
Drill Down: CQA 8 by Year



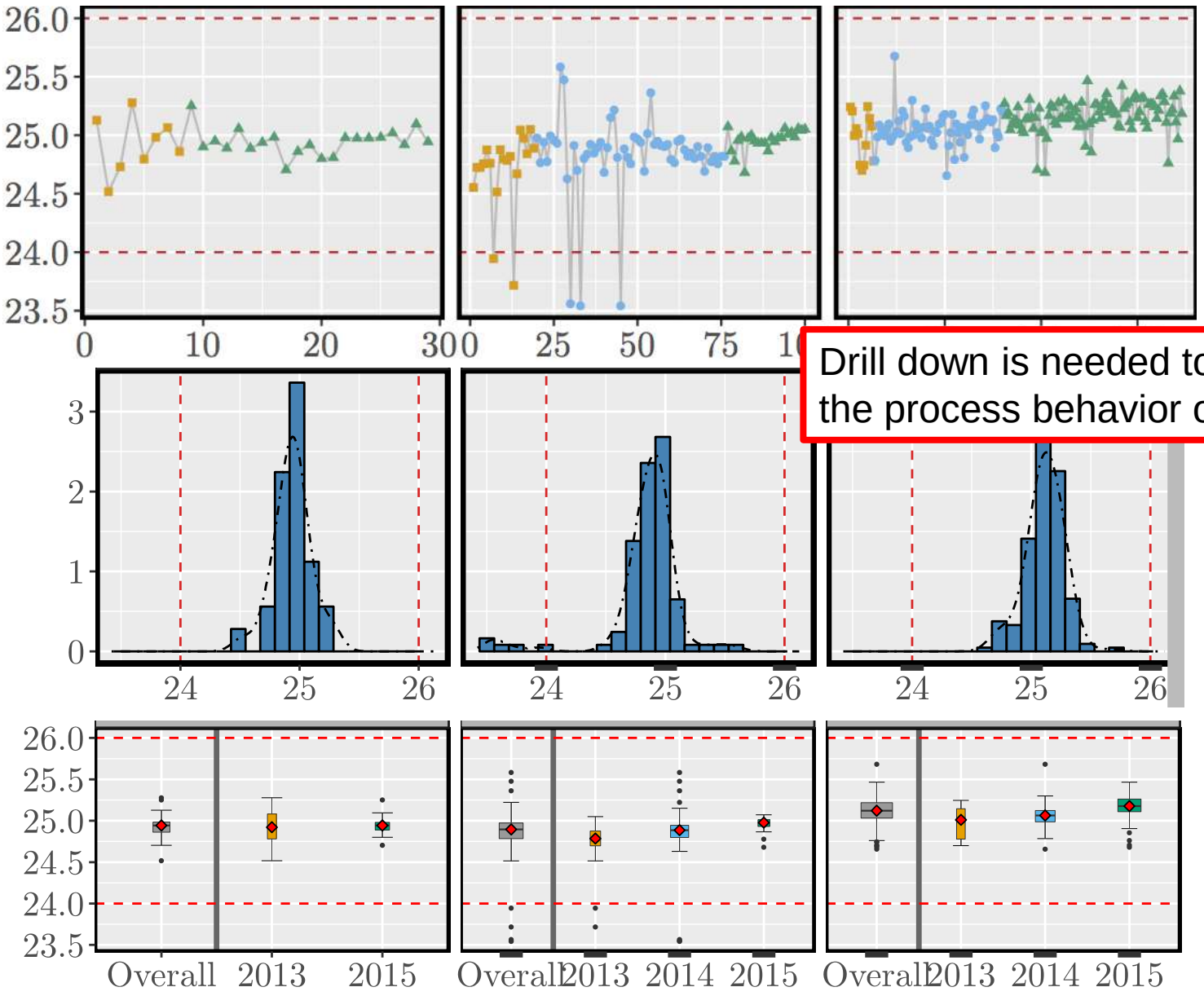
Drill down is needed to understand the large process variation over time.

Product level: risk ranking overview

Product X, Site A B C, Years: 2013-2016

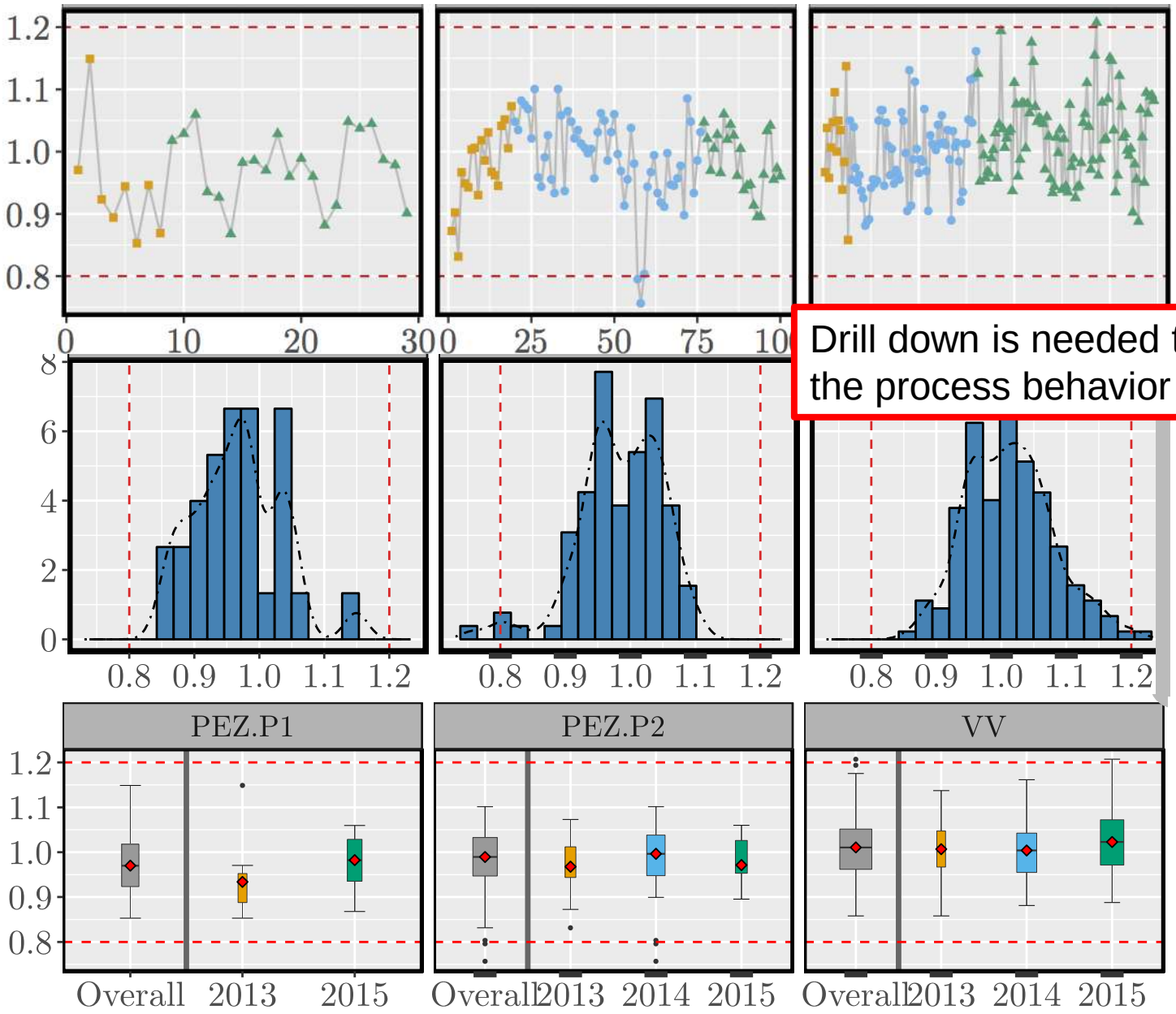


Drill Down: CQA 3 by Year, Multi-site



Drill down is needed to understand the process behavior over time.

Drill Down: CQA 8 by Year, Multi-site

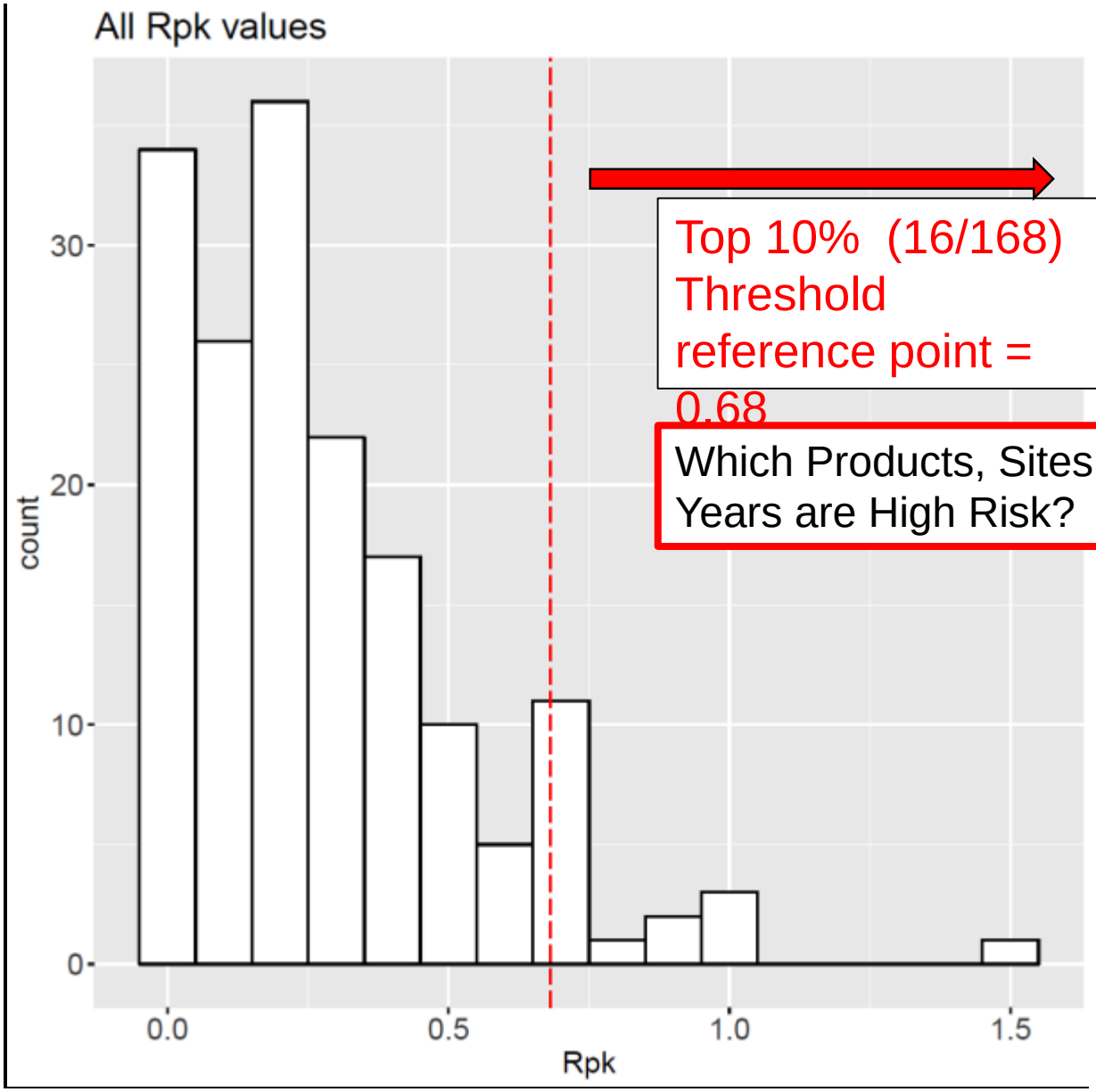


Drill down is needed to understand the process behavior over time.

Network level:

- Review network manufacturing process performance
- Design threshold to drive continuous process improvement

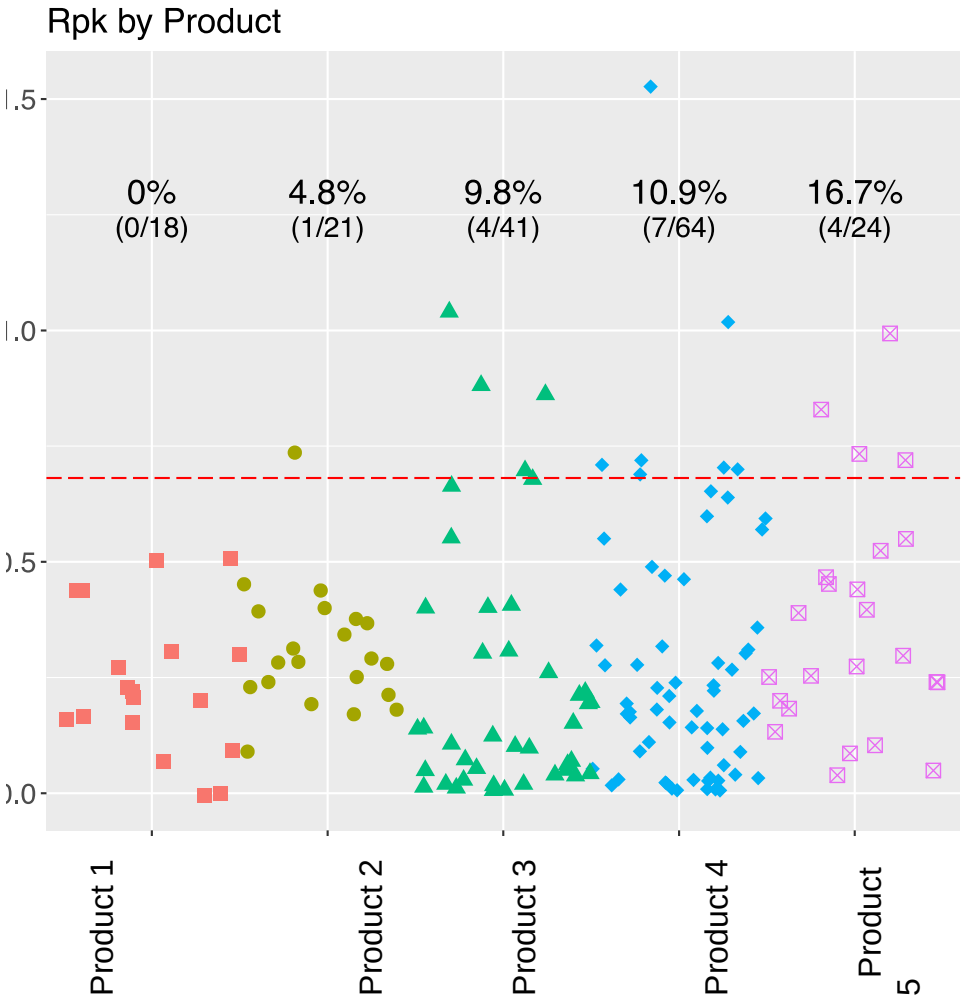
Annual Rpk for Biologics (N≥20 batches)



Top 10% (16/168)
Threshold reference point = 0.68

Which Products, Sites, CQAs and Years are High Risk?

Characterization of Biologics Network
Rpk by sites and product



Empirically derived Thresholds

Request:

Manufacturing sites want thresholds to classify Rpk as low or high risk

Proposal:

Use 10% percentile from network results, from previous year to set threshold for the next year.

Possibly separate thresholds for biologics DS, biologics DP, small molecule DS, and small molecule DP.

Properties:

If process performance does not change, expect 10% high risk and 90% low risk

If process performance improves, expect fewer high and more low risk

4. FAQ

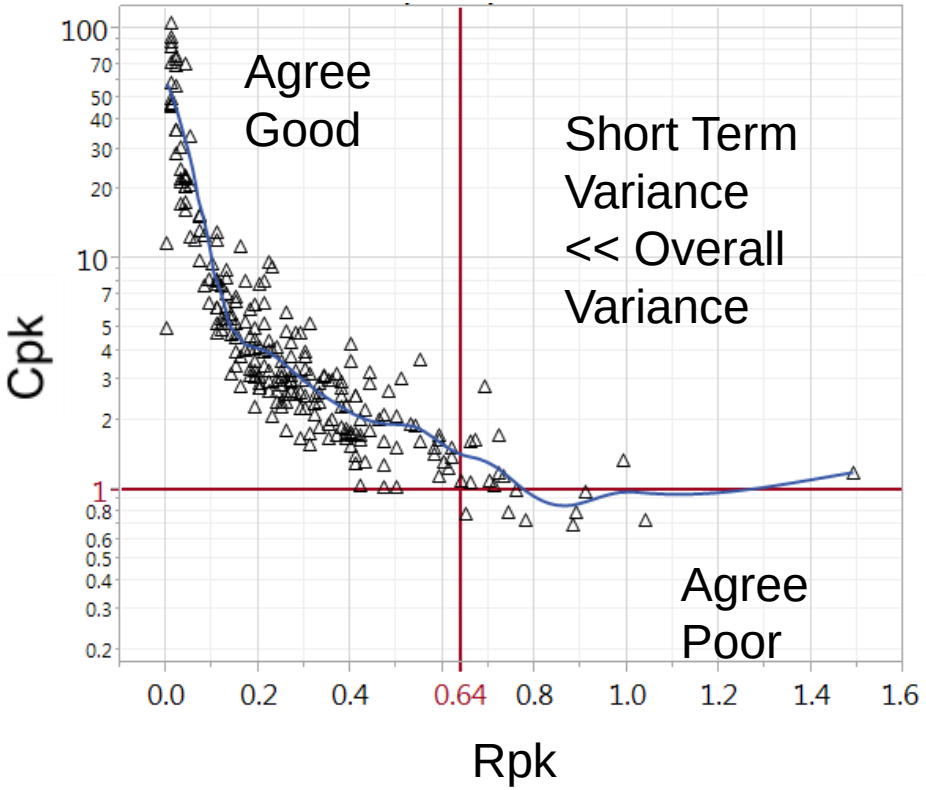
Rpk requires minimal / no assumption

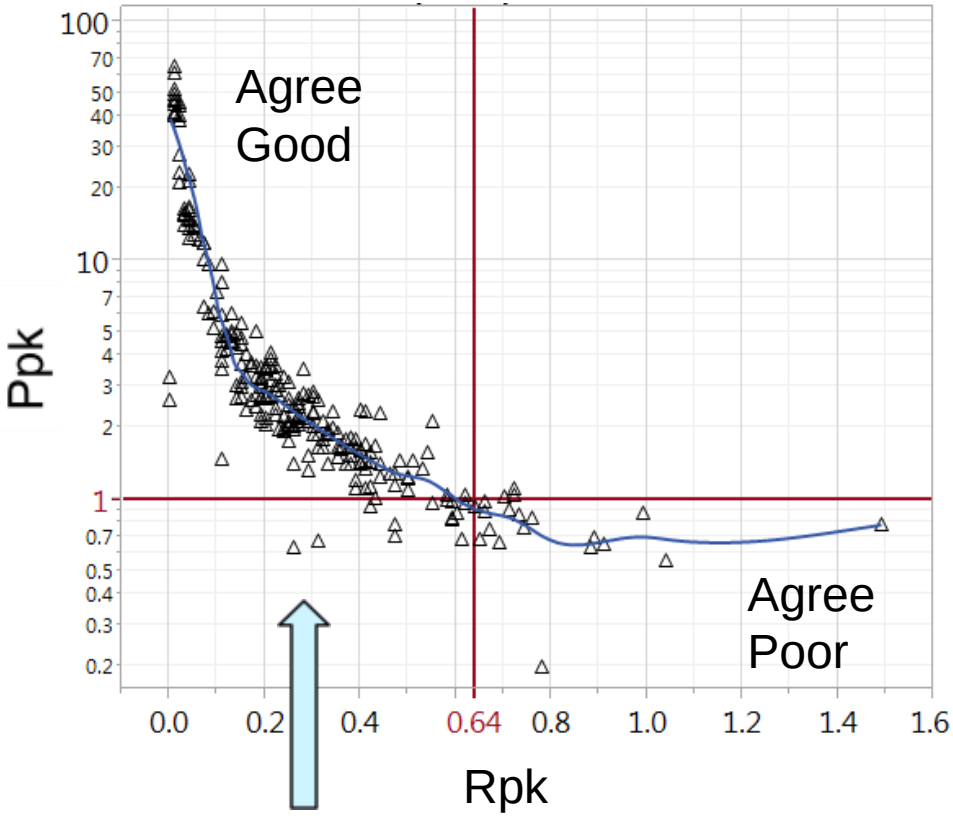
Q1: Why not Cpk or Ppk?

- Cpk uses the **short term variation**, which excludes variation due to trends, shifts, and other sorts of process wandering
- These components of the overall variation need to be included for risk assessment -- Cpk is not suitable for risk ranking

- Ppk uses the overall variation, but similar as Cpk, the calculation assumes process data follows Normal distribution -- mean ± 3 sd describes the data when the distribution is Normal (symmetric)
- Many Pharmaceutical data **may not follow Normal distribution** – Ppk and Cpk may be misleading

Comparison of annual Rpk to Ppk and Cpk on pilot study (biologics & small molecules, N≥20)





Short tail distributions or special cause outliers inflating variance

Q2: Why use the .05 and .95 quantiles, *q.05* and *q.95* ? -- Why not the range, i.e., minimum and maximum, or the .00135 and .99865 quantiles?

- The range is not acceptable because interpreting it correctly and consistently depends on the sample size.
  - E.g. the range of 10 CQA values is likely to be a lot smaller than the range of 100 CQA values -- **Quantiles are better**

- *q.05* and *q.95* is a **practical compromise** between making the "bulk" as large as possible and having sufficient data to reliably compute the sample quantiles. For datasets with at least **20 = 1/.05 batches**, our experience is that Rpk has been reliable for risk ranking

- *q..00135* and *q.99865* make the "bulk" of the data very large but require large sample sizes 741=1/0.00135 to estimate. For many pharmaceuticals, **741 batches** are far more than would be made in a year

Q3: Why is Rpk “the smaller the better” instead of “the larger the better”?

- Rpk behaves linearly with the process range, i.e., reduce the process range by half, the Rpk is reduced by half; double the process range, the Rpk is doubled – **linearity makes it more interpretable**
- When the process range is within specifications, Rpk is **between 0 and 1**, while the Cpk and Ppk are between 1 and infinity.

Q4: Why do we not require processes to be in "statistical control" for Rpk to be valid?

- Conventional use of Cpk, e.g. in consumer products industries, demands that **processes be in "statistical control"** [Montgomery, 2009] so that the indices describe "stable distributions".
- This enables the data to meet strong statistical assumptions needed to predict future product results. Such demands are generally **unrealistic for the pharmaceutical industry**
- Rpk is designed for **risk ranking** by upper level managers: the aim is to **provide an informative, accurate overview of network performance** without making inappropriate statistical assumptions and oversimplifications.
- Rpk is meant to be **descriptive**, not predictive.

Q5: What distributional assumptions must be met for Rpk to give results suitable for risk ranking?

- **None**

- Several “rules of thumb” to help assure the meaningful characterization
  - There should be **at least 20 data points** per CQA.
  - There should be **at least 10 distinct values** in the data

5. PPA Roll out

Automation

Suitable for automation

- No outlier detection, no data transformation, etc.
- R Shiny app (excel data in, word document report out)

Conclusions

- Rpk
  - Rpk is a reasonable metric for risk ranking CQAs
  - appropriate for all data, including low volume, non-normal data
- Business drivers:
  - Rpk provides a common metric for characterizing a pharmaceutical manufacturing network
  - Goal is to make Rpk a component in management's decision making, e.g., prioritizing process improvement through risk ranking CQAs
- Challenge:
  - integration into the quality system

Acknowledgements

- Bert Gunter, Pia Krieger, David Cate, etc.

References

- “The Use and Abuse of Cpk,” - B.H. Gunter, Quality Progress, Parts 1-4 (1989-1991)
- Process Capability Indices – Samuel Kotz and Norman L. Johnson
- Beyond Capability Confusion 2nd Ed. – Donald J. Wheeler
- Advanced Topics in Statistical Process Control 2nd Ed. – Donald J. Wheeler
- Statistical Quality Control 7th Ed. - Douglas C. Montgomery
- “Statistical Quantiles in Statistical Packages” – Rob J. Hyndman and Yanan Fan, The American Statistician (1996)
- “Process Capability Computations for Non-Normal Distributions” – John A. Clement, Quality Progress (1989)
- “Process Capability Indices for One-sided specification intervals and skewed distributions – Kerstin Vannman and Malin Albing, Quality and Reliability Engineering International (2007)

Doing now what patients need next