Uncertainty of Measurement in the Medical Laboratory

Florent Vanstapel MD PhD
Clinical Chemistry – POCT Coordinator
Laboratory Medicine
UZ Leuven – KULeuven – Belgium

CAT-Series – Leuven April 23 2015

Everything you wanted to know about UM

- **norm-compliance?**
- **clinical relevance?**

But were afraid to ask

- **is it difficult?**
- **do you get payed for it?**

Relevant Standards

- EU IVD Regulation
- ISO 15189
- Technical standards

EU DIRECTIVE 98/79/EC (IVD Directive)

What the industry is required to do

Art. 1. 3. For the purposes of this Directive, calibration and control materials refer to any substance, material or article intended by their **manufacturer** either to **establish measurement relationships** or to verify the **performance characteristics** of a device in conjunction with the **intended use** ...

Annex III. 3. adequate **performance evaluation data** showing the performances claimed ... and **supported by a reference measurement system**, with information on the reference methods, the reference materials, the known reference values, the **accuracy** and measurement units used; such data should originate from **studies in a clinical or other appropriate environment or result from relevant biographical references**,

ISO-15189:2014

What the laboratory is required to do

5.3.1.4 Equipment calibration and metrological traceability
b) **metrological traceability** of the standard and calibration
c) **verifying measurement accuracy** … **at defined intervals**;
traceability shall be to a reference material or procedure

5.5.1.2 Verification of examination procedures
applies to Validated (= established) examination procedures …
independent verification by laboratory … through objective evidence
5.5.1.3 Validation of examination procedures
validate a) **non-standard methods**; b) laboratory developed methods; c) standard methods used outside intended scope;
d) validated methods subsequently modified.
NOTE Performance characteristics include:
trueness, accuracy, precision; uncertainty, analytical specificity, …

ISO-15189:2014

Very narrowly defined

5.5.1.4 Measurement uncertainty of measured quantity values shall … regularly review estimates of measurement uncertainty

NOTE 1 The **relevant components** … actual measurement process, commencing with the presentation of the sample to the measurement procedure and ending with the output of the measured value.

NOTE 2 Measurement uncertainties may be … obtained **by measurement of quality control materials under intermediate conditions** that include e.g. changes of reagent and calibrator batches, different operators, scheduled instrument maintenance.

Upon request, the lab shall make its estimates of measurement uncertainty available to laboratory users.

Internal Quality Control: ANOVA

Technical Standards

How things are done

Eurachem: The Fitness for Purpose of Analytical Methods
Laboratory Guide to Method Validation & Related Topics

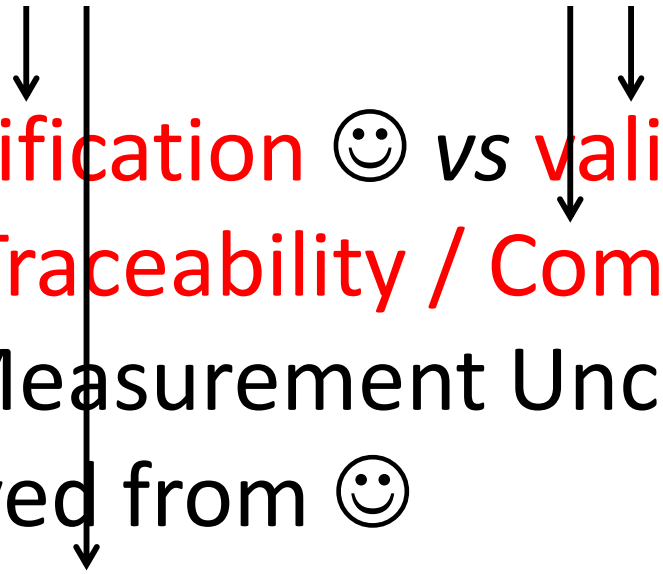
6.5 METHOD DEVELOPMENT
6.6 THE DIFFERENT PERFORMANCE CHARACTERISTICS OF A METHOD
6.6.4 Trueness
6.6.5 Precision
6.6.6 Measurement uncertainty

TrainMic: EC JRC IRMM
Training in Metrology in Chemistry

3. Single Laboratory Validation of Measurement Procedures
4. Uncertainty of Measurement Principles
5. Uncertainty of Measurement: Approaches to Evaluation

Message #1

Standards don't ask for much
neither from us ☺ or from the industry ☹

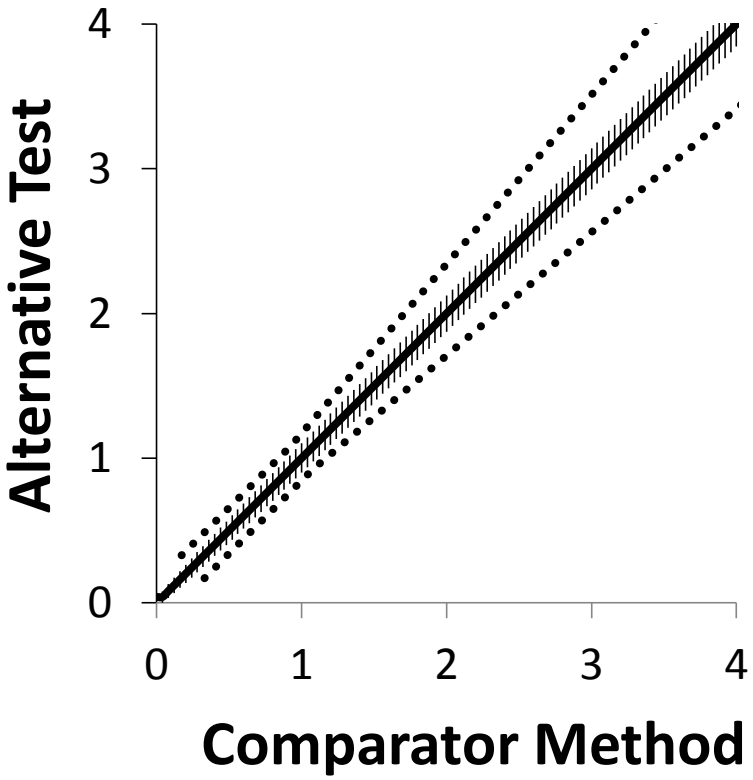


- Method verification ☺ *vs* validation
- Trueness / Traceability / Commutability ☹
- Budget of Measurement Uncertainty
  can be derived from ☺
  internal Quality Control Data ☺ ☹
- …

Casus:
Typical Method Implementation

Validation File:
Fit for the Intended Use?

Method Validation: Traceability to Comparator



Intercept = 0; Slope = 1

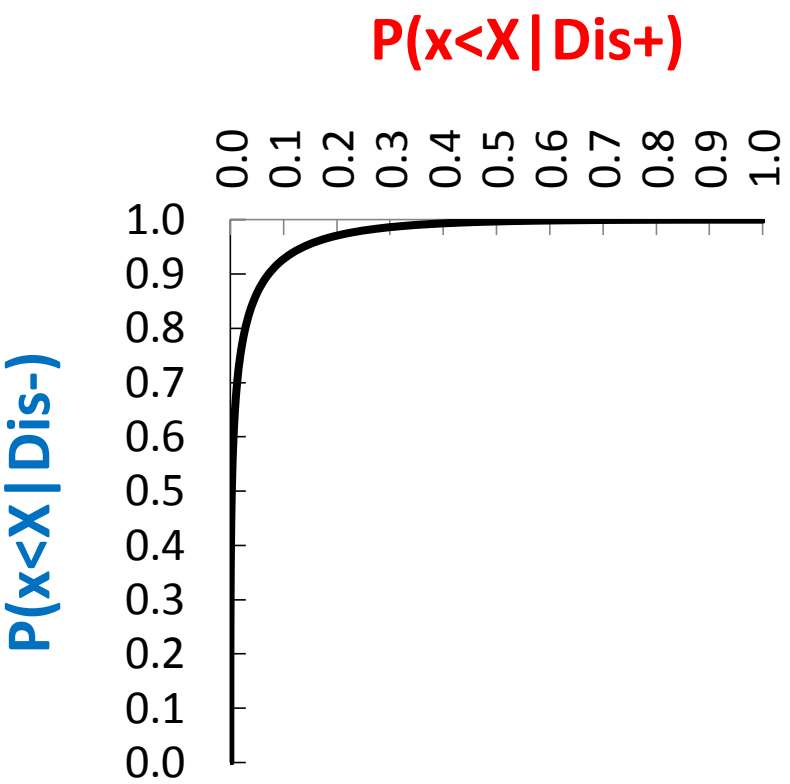
Allowed Total Error (dashed line)
= 31% of Spread in Reference Population
(Tonks criterium)

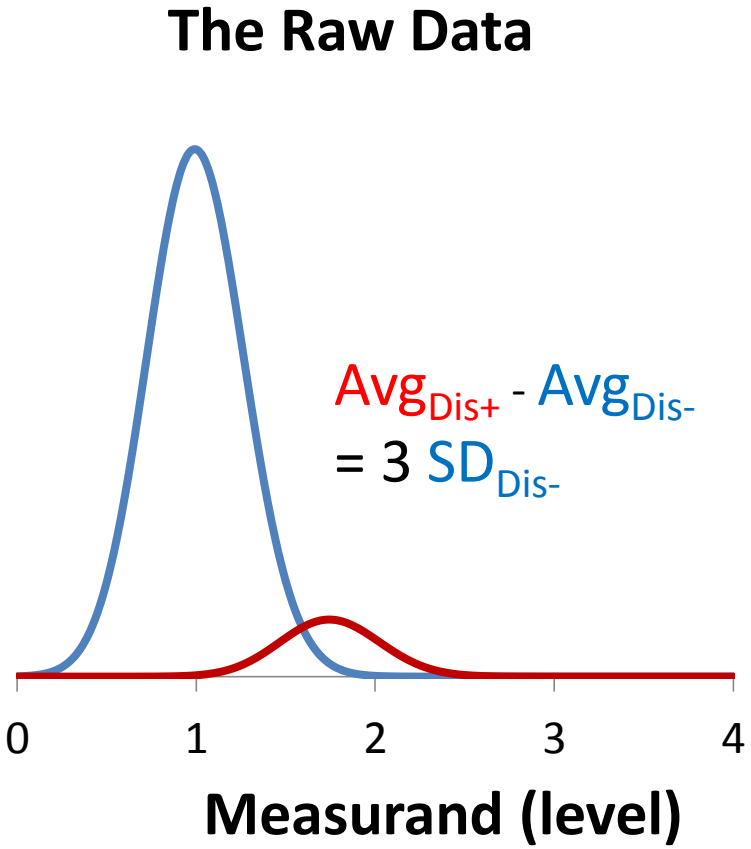
Analytical Imprecision (shaded area)
< 1/3 of Total Error
~ 1/2 of a 6 σ process

Is this a good test?

What was the Question: Does this make it a test
Fit for the Intended Use?

Clinical Validation: ROC-curve



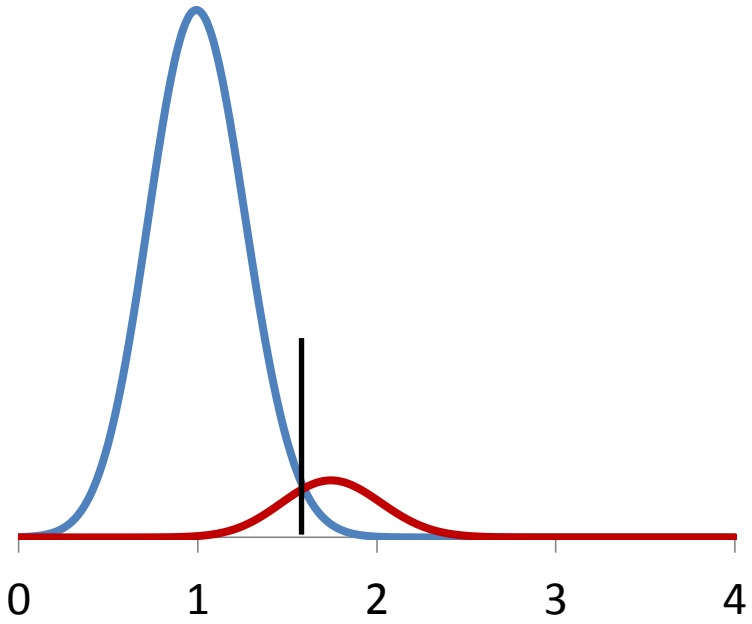


Avg_{Dis+} at 3 SD_{Dis-} = Definition of Limit of Detection

Does this make it a test unfit for purpose?

Dichotomic Analysis

Black – Gamble – White



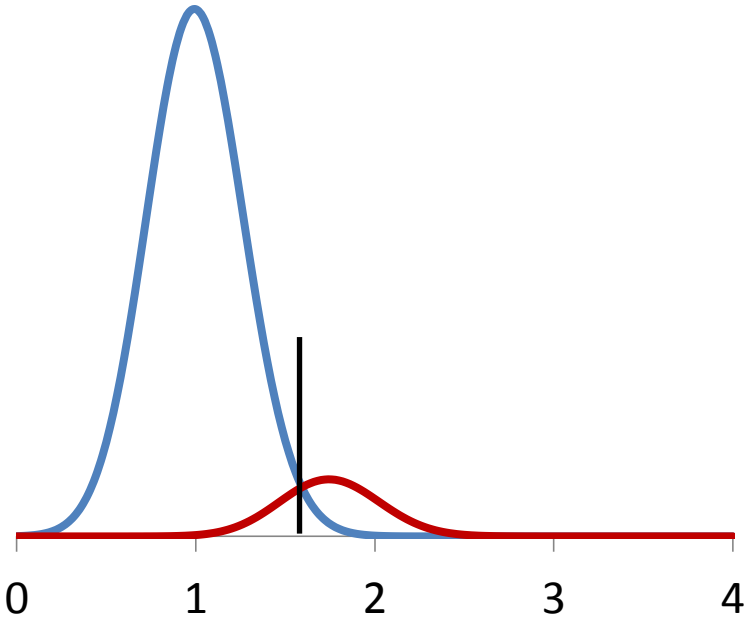
Loss of Truth due to misclassification

	D-	D+
T-	1	-1
T+	-1	1

Kappa-weights

Dichotomic Analysis

Black – Gamble – White



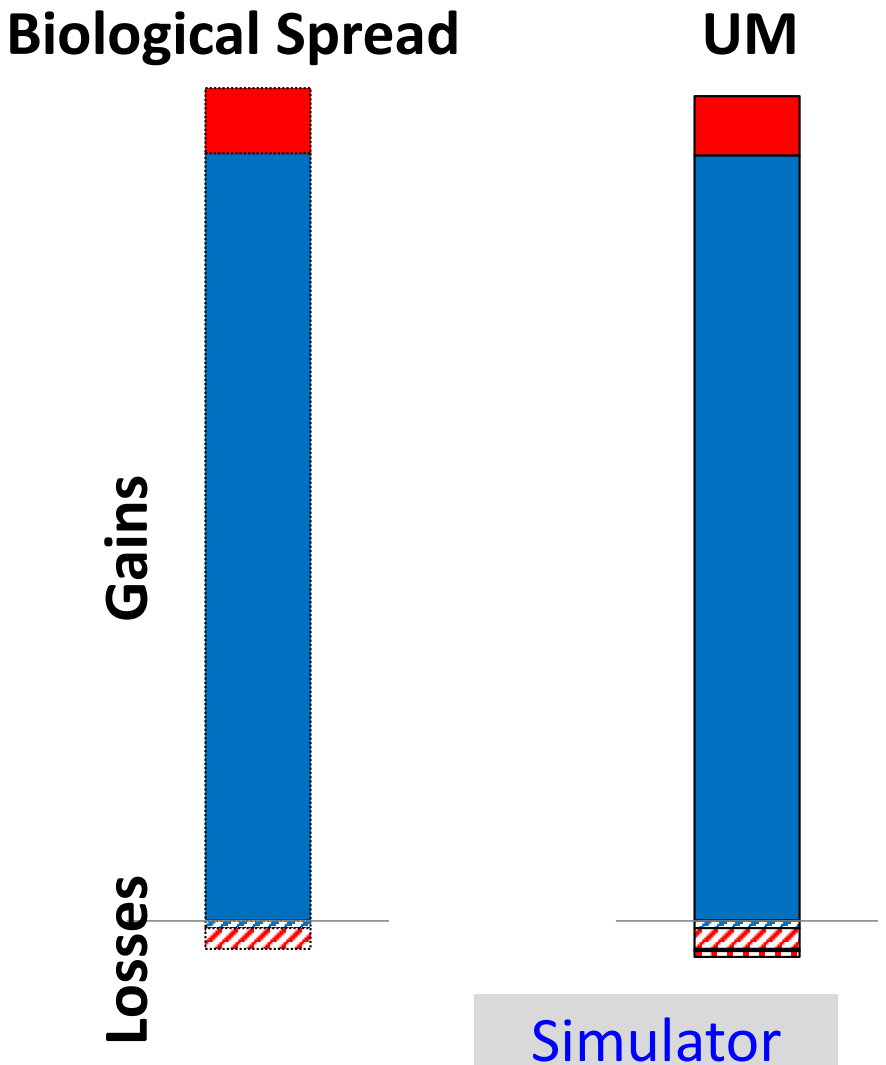
Uncertainty Budget

Total

Analytical

Biological

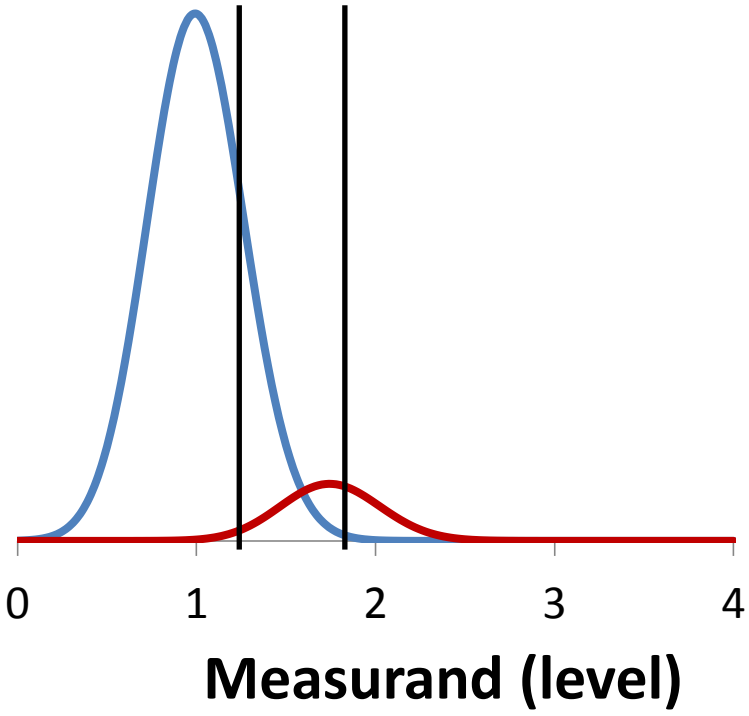
Loss of Certainty due to



Simulator

How a Test Result is Interpreted

Black – Grey – White



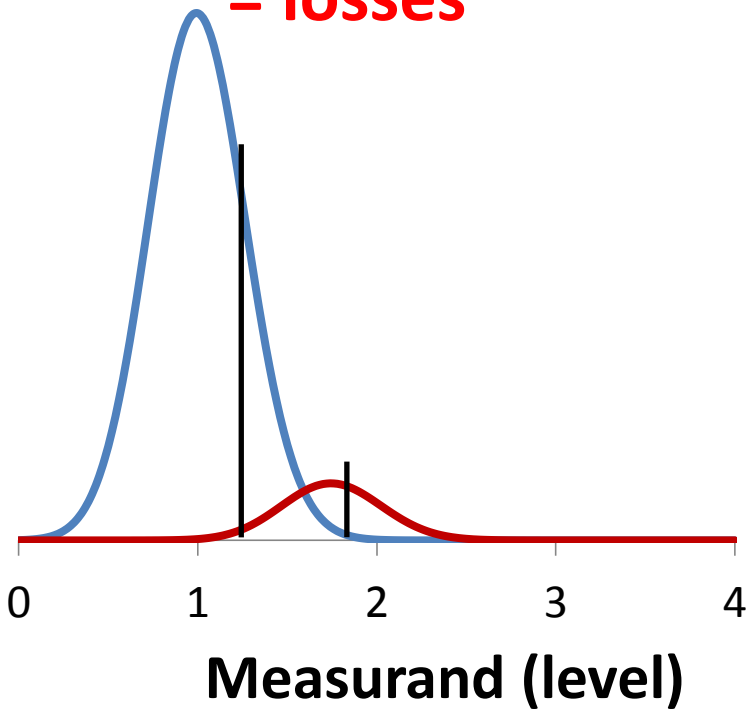
Differential Diagnosis

- Don't miss what
  needs urgent intervention
- Workup
  list of common causes
- But also exclude rare causes

You need not to confirm doubt
but to exclude Dis- or Dis+

Dichotomic Analysis

Black – Undecided – White
= losses

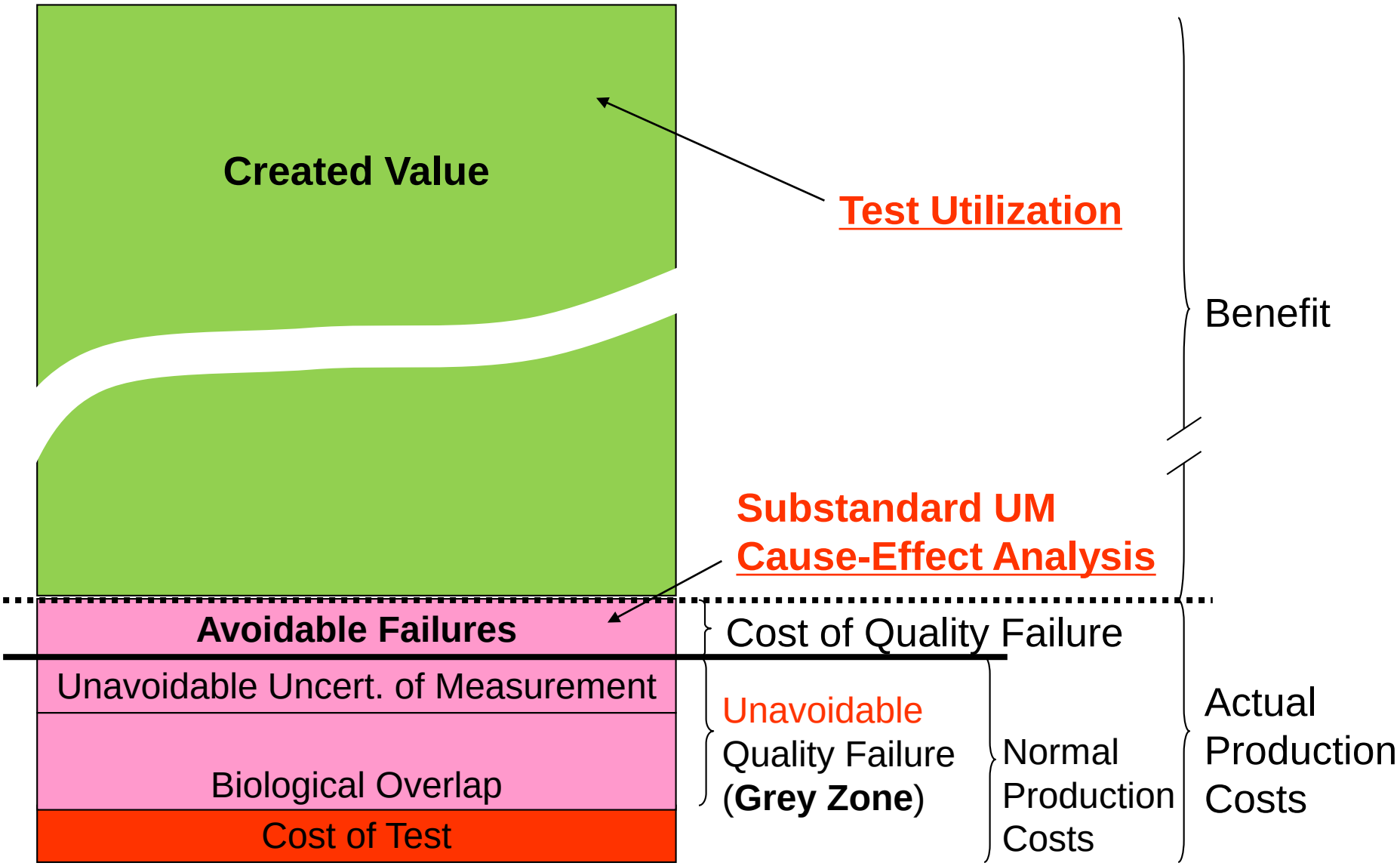


If you refuse to Gamble

predictive value increases
trade-in: sensitivity
& specificity

CAVE: grey zone
no decision = also a decision

Where can the validation exercise create value ?



Message #2

Diagnostic Value depends on whether you ordered a test with Power of Exclusion

Typically, Analytical Uncertainty has little effect on Diagnostic Power ☺

UM doesn't ruin a good test ☺
& can't rescue a bad test ☹

Everything you wanted to know about UM

- norm-compliance?
- clinical relevance?

But were afraid to ask

- **is it difficult?**
- **do you get payed for it?**

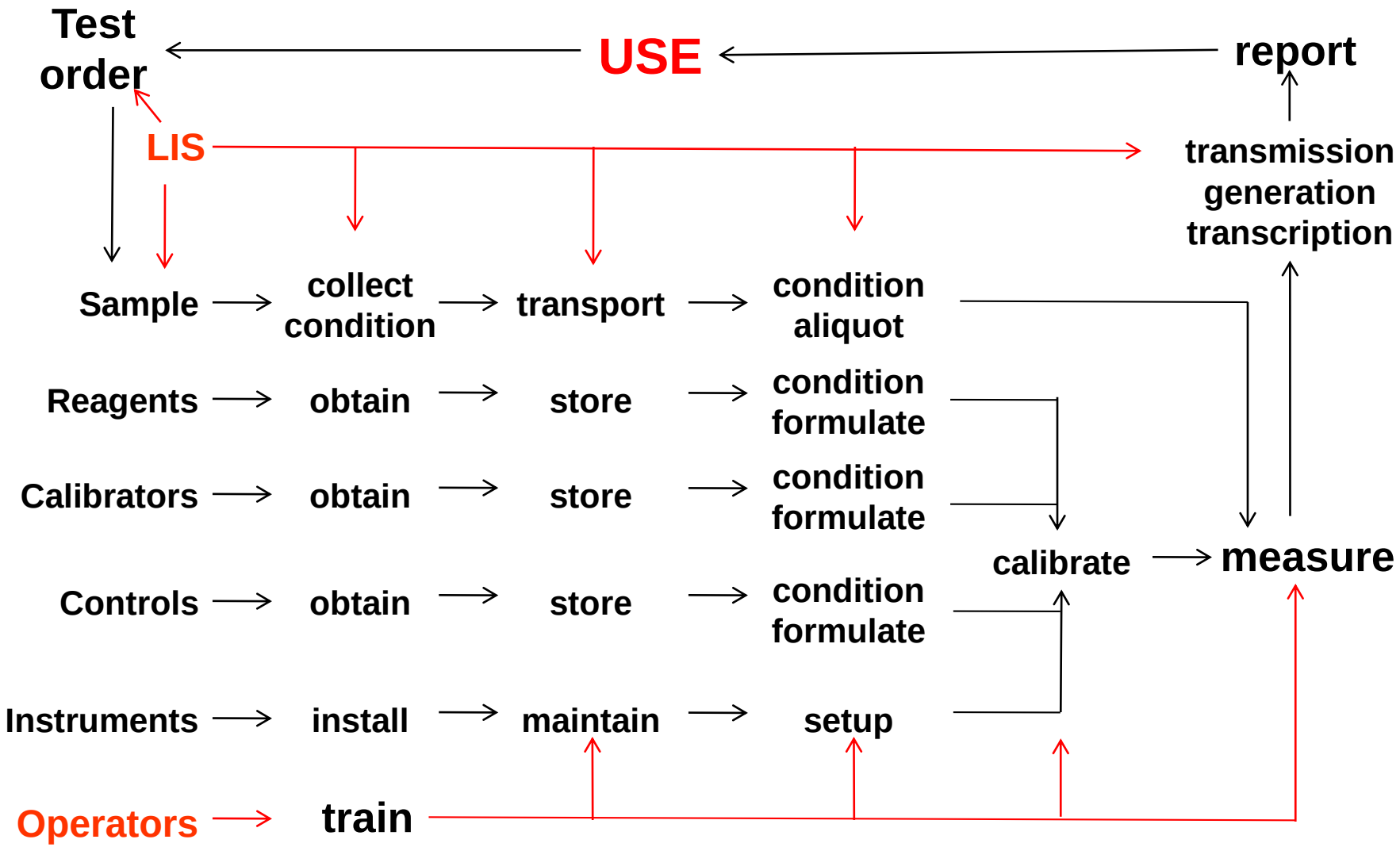
Casus:
UM properly

The Nuts and Bolts of UM Evaluation

UM Evaluation starts with

Inventory of Sources of Error

Risk Analysis Generic Laboratory Process



After ISO/TS 22367:2008 **= Workshop Approach**

= Fishbone analysis

Risk Analysis: Failure-Mode analysis: What can cause relevant failures?

Test Ordering

Right Test at the Right Time … for the Right Patient …

Pre-analytical

Biological Variation
Pulsitility, Diurnal & Seasonal Variation
Physiological (Starvation, Exercise, …) Variation, TimeStamp

Specimen Collection
Posture
Stasis, Hemolysis, Filling of the tube
Right patient, Correct labelling, sample, recipient, …

Interpretation

Adequacy, sampling details , specimen quality, …, …
Adequacy cut-offs, reference ranges, …
Adequacy interpretation support

Process Approach of Diagnostic Cycle

Post-analytical

Reporting for the right patient
Transcription errors , …, …
Data transfer errors, ….

Analytical = Measurement Process

Sample reception & processing
Right identification of primary and secondary sample
Completeness of coagulation, …
Micro cloths: obstruction needle, light scattering, …
Membrane ghosts and fragments, …

Uncertainty of Measurement
Bias, Specificity, …
Calibration, Linearity, …
Inter-batch random error, …
Other sources of (random) imprecision
Equipment faults (aspiration, carry-over, sporadic faults, ..)

Clasical Focus of Method Validation
Is this where most relevant failures originate?

Message #3

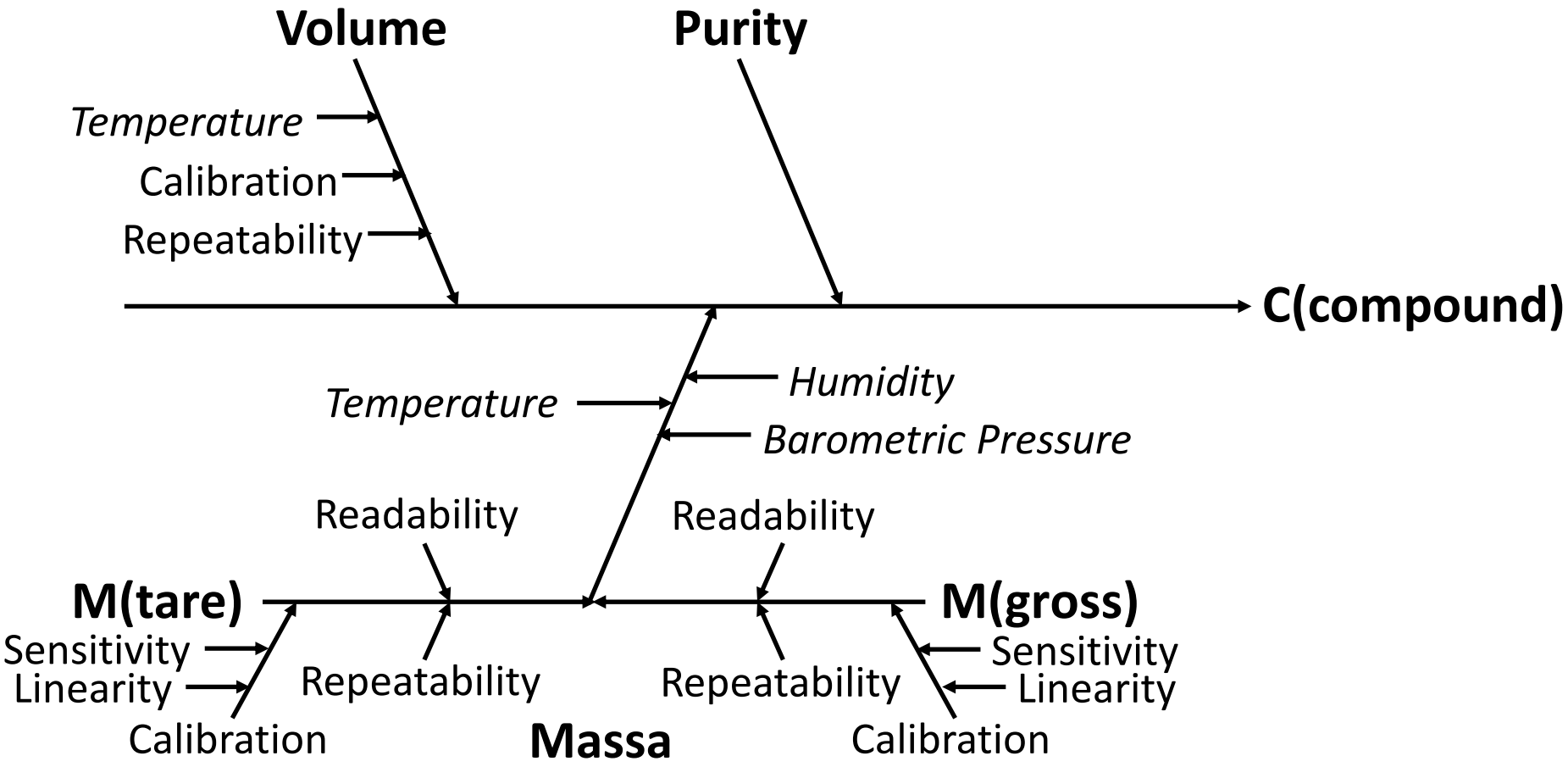
The Core Analytical Process
and the associated UM
should not be your only focus,
maybe not even your primary focus?

UM Evaluation
(in the narrow sense)

Approach #1
Inventory of Sources of Analytical Error

Fishbone Analysis of an Analytical Assay

e.g. Calibrator formulation



Message #4

Fishbone Analysis identifies

- Input Variables

= Model Equation (see next section)

- *& Influence Variables*

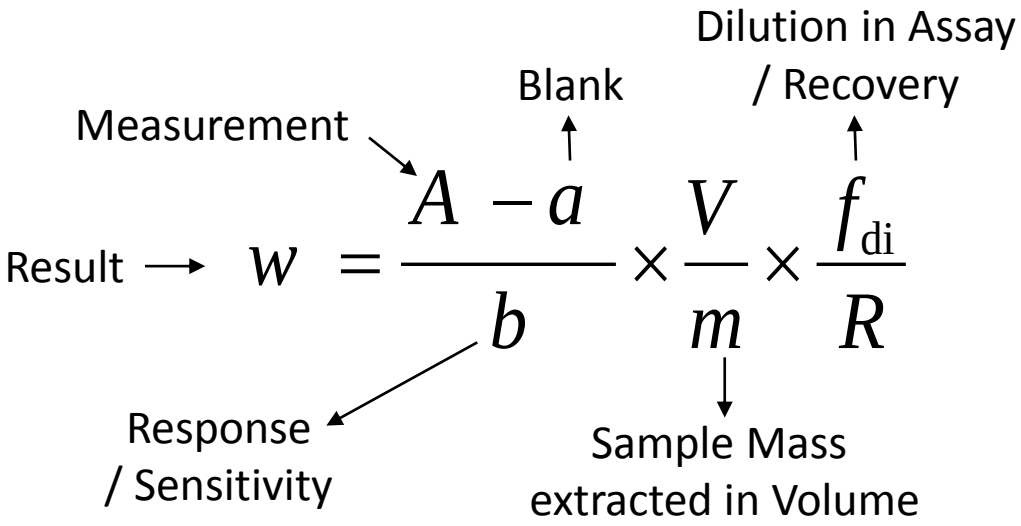
typically not represented in the model

did your supplier develop a rugged test?

Approach #2

Model Equation -> Error Propagation: Sensitivity or Ruggedness Analysis

Example: Extraction & Calibration



What are the critical issues ?

m/V: Unfavourable?
f_{di}: Unfavourable dilution in the assay?
R: Incomplete and haphazard extraction?
(A-a)/b: Unfavourable dynamic Range A *vs* a; Low response factor A/b?
a: Constructed Blank or Extrapolated from Calibration Line ?
b: Historical Colour Yield or In Run Calibration ?

Message #5

Writing down the Model Equation & critically inspecting it tells you a lot ☺

First order Taylor Series approximation of the process

The combined standard uncertainty =
reasonable dispersion of measurement results =
function of the uncertainty of the component quantities

Pythagoras : destructive accumulation of noise ☺

$$u^2(y(x_1,x_2,\ldots)) = \sum_{i=1}^{N} \delta y/\delta x_i \; u_i^2 + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} (\delta y/\delta x_i)\,(\delta y/\delta x_j)\, c_{i,j}$$

☺ Sensitivity Analysis

Covariance: $u_i u_j r_{ij}$

Concentrate on what matters = ruggedness

Partial derivatives ? Don't panick :
mostly additions/subtractions & multiplications/divisions ☺

For independent parameters, the second term is absent ☺

Kragten (or spreadsheat) Method:

The simplest of cases

$$X = A*B/C$$

X: concentration resulting from

- **A: dilution of the sample**
- **B: measurement**
- **C: calibration factor**

Kragten Method: Propagation of error

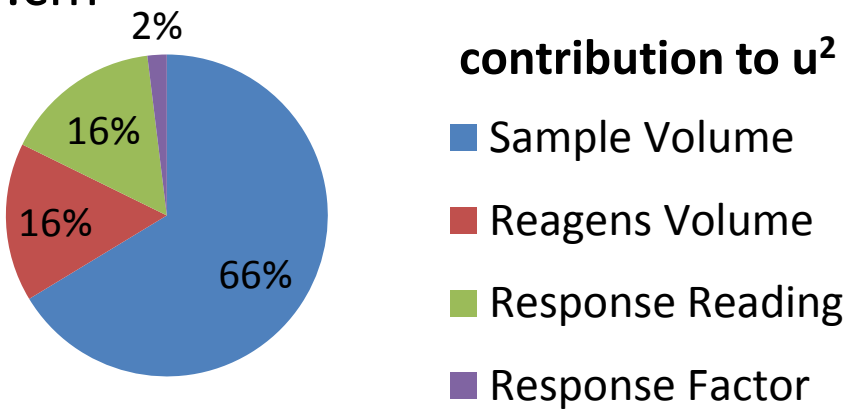
Model		Effect of partial disturbances		
X = A*B/C		u_a	u_b	u_c
A	a	$a + u_a$	a	a
B	b	b	$b + u_b$	b
C	c	c	c	$c + u_c$
Evaluation	x	x_a	x_b	x_c
	$u_{x,n}^2$	$(x_a - x)^2$	$(x_b - x)^2$	$(x_c - x)^2$
	$\sum u_{x,n}^2$			u^2

Propagation of error: An Example

A sample is diluted in a reagent mix. At the end point of the reaction the color yield is read, and divided by a calibration factor.
That result is multiplied with the dilution factor.

		Value	SI	u_i/x_i	
y	**$= (x_3/x_4)(x_1+x_2)/x_1$**	**359**	**mM**	**0.1%**	
x_1	Sample volume	5	µL	4%	from pipet control program
x_2	Reagens volume	200	µL	2%	from pipet control program
x_3	Response Reading	0.7	cm^{-1}	0.05%	from equipment specs
x_4	Response Factor	2000	10^6.L. $mol^{-1}.cm^{-1}$	0.02%	Epistemic variable (literature)

Simulator



Message #6

Don’t be scared ☺ of a Taylor Series ☺

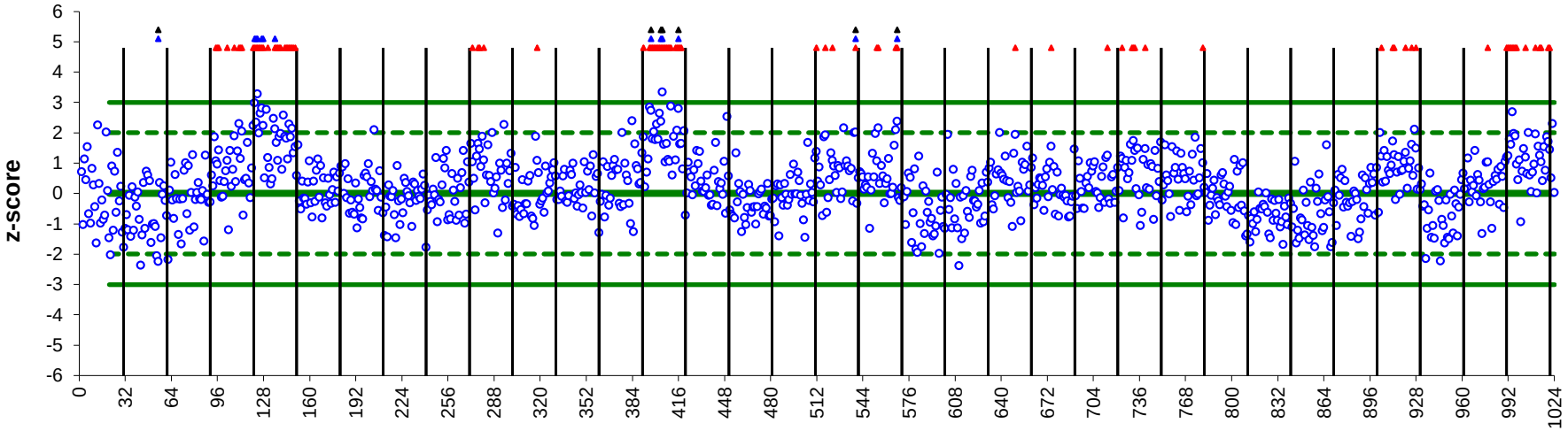
The analysis tells you what to focus on ☺

Approach #3a

Measurement Uncertainty From Internal Quality Data

Internal Quality Control Data

Typical immunometric assay



There are episodes where results consistently differ from the majority of results?

Do we have a problem?

Simulator

ANOVA of Internal Quality Control Data

#	Component of variation	CV	Comments: Typical sources of variability
1	Cal. Nominal Uncert.	1.0%	Applies to each new calibrator batch
2	Cal Manipulations	1.0%	Applies to manipulations and conditioning
3	Between Reagent 1	1.4%	Applies to start-up of new reagent batch
4	Within Reagent 1	0.7%	Applies to conditioning of individual reagent packs
5	Between Reagent 2	1.4%	Idem as 1
6	Within Reagent 2	0.7%	Idem as 1
7	Maintenance	1.0%	Maintenance
8	Start-Up procedure	0.0%	Start-Up procedure
9	Operator	0.0%	Operator dependent variability
10	Between Run	5.0%	Applies to long-term variability of the analytical system
10	Within Run	5.0%	Applies to short-term variabilityof the analytical system
12	iQC Manipulations	1.0%	Applies to manipulations and conditioning
Total	**UM**	**10%**	**Uncertainty of measurement**

- The effects of calibrations, lots, ... **persist**
- Certain events are **synchronous**

Message #7

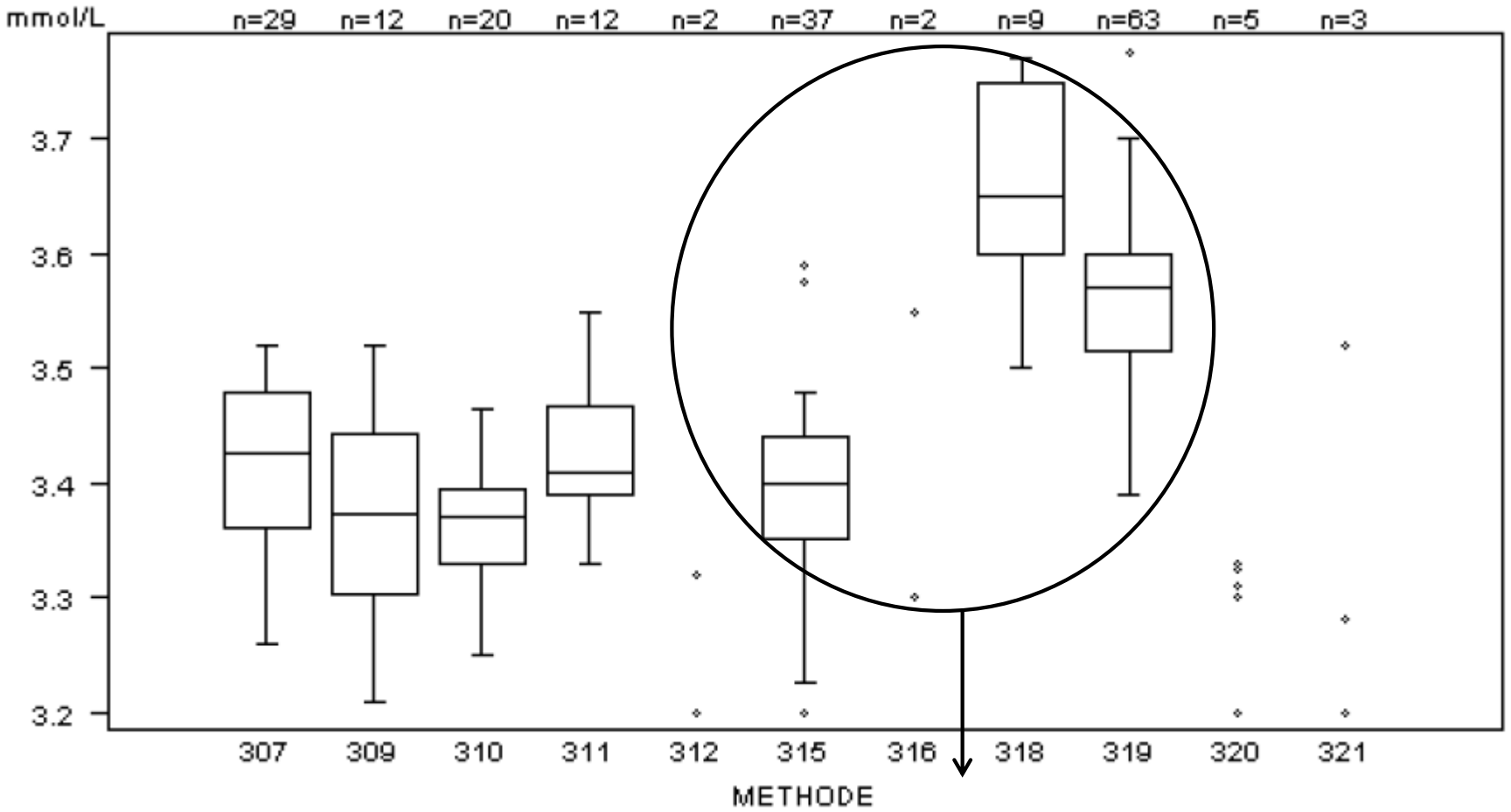
The example was a simulation:
nothing abnormal was happening
random variation within expectations ☺

ANOVA of iQC data = UM Budget
helps you
to avoid panick & futile interventions ☺

Approach #3b

Measurement Accuracy
From Proficiency Testing Data

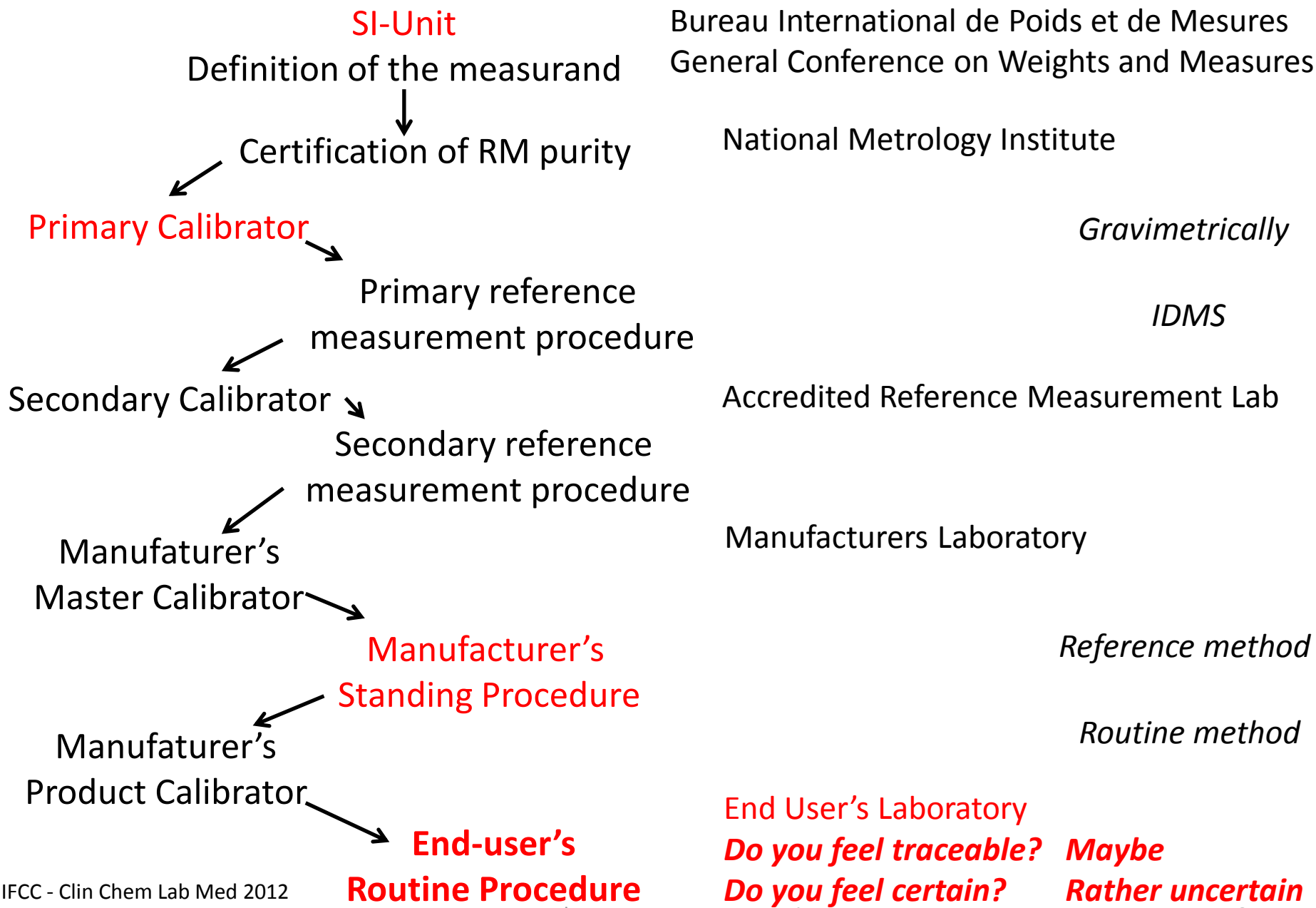
Belgian Proficiency Testing Scheme: Chemistry Calcium



EQAS commutability

single manufacturer? not OK

ISO 17511: Hierarchical scheme of traceability



IFCC - Clin Chem Lab Med 2012

Message #8

The design of proficiency testing is flawed:

- at best they measure “relative” bias
- differences may or may not be attributable to matrix effects

Hence “Proficiency” Schemes:

- document non-commutability of methods or of materials?

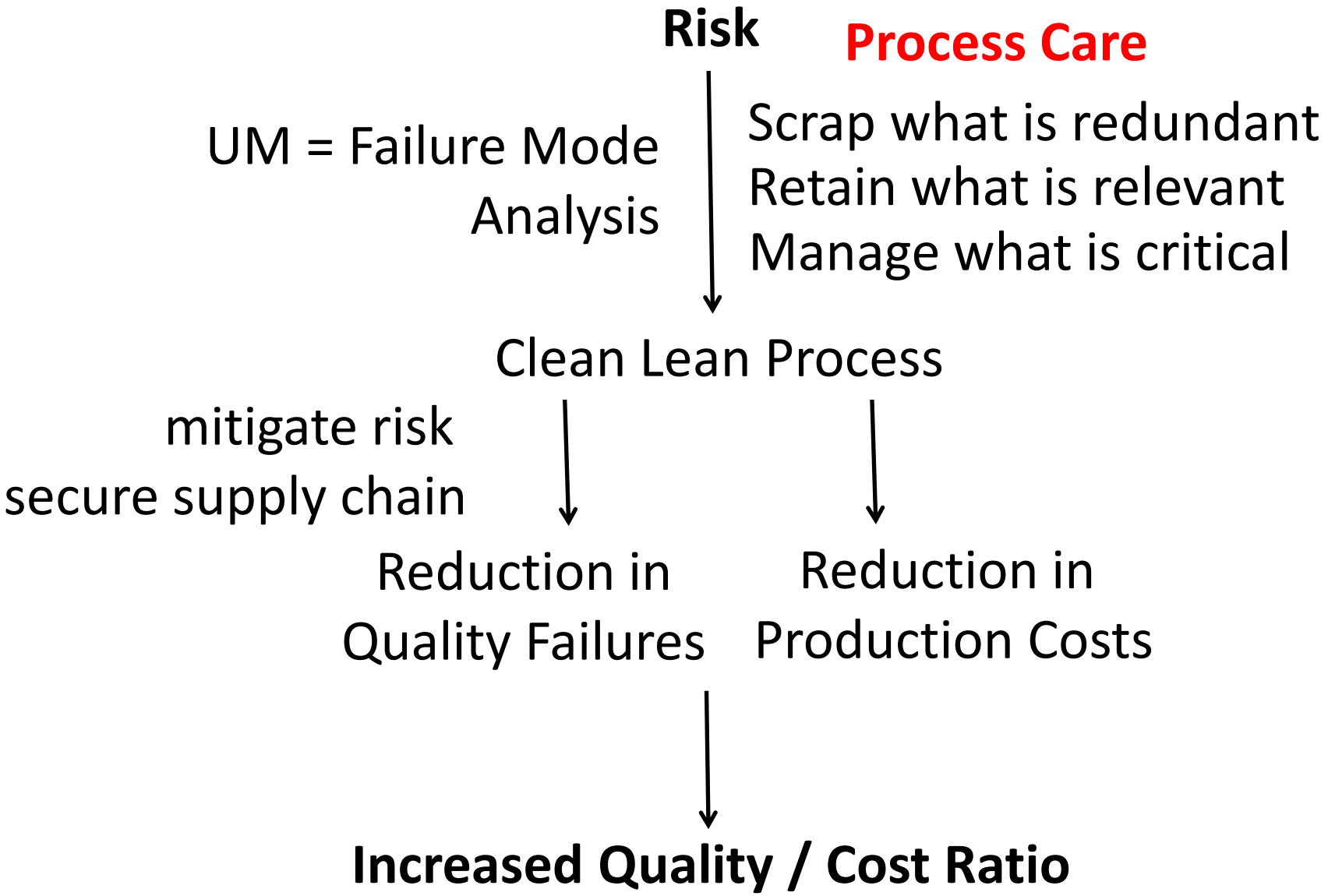
The perversion of “Matrix Effects”:

- makes it harder for the lab to switch methods

Summary

Who is paying for this?
Reaping the Fruits

The Economy of your Quality System



CONCLUSIONS

UM = Failure-Mode Analysis
= Risk Analysis Tool ☺

Risk-Mitigation
= Process-Care Tool ☺

Process-Care
= Quality at No Cost ☺

CONCLUSIONS

Understanding UM allows

Production Department:

- to focus on what matters
- not to spend resources on futile projects

Interfacing with Prescribers:

- analytical precision suffices ☺
- transmural commutability is a problem ☹
- **invest** recuperated opportunity costs **in test ordering and interpretation support**

TO DO

Develop Rugged Methods:
to reduce waste activities (iQC, etc.)

Commercialize Commutable Calibrators:
to realize
longitudinal coordinated transmural care