Tuning the Quality System

Cost-Effective
(Statistical) Process Control (SPC)

Prof. Dr. F. Vanstapel, MD PhD
Laboratoriumgeneeskunde
UZ KULeuven

Teaching Goals

What is the Added Value of internal Quality Control ?
What does it do for You ?

Does internal Quality Control
bring Attained Quality closer to the Desired Quality ?
improve Process Capability ?

Teaching Goals

- understand process control
- understand requirements of standards
- understand the relationship between process control and the value chain under control
- analyze internal quality control (iQC) as a diagnostic process
- understand operational characteristics of iQC
  - pending failure *versus* unacceptable failure
  - urgency & window of opportunity
  - false/true positive/negative detection rates
- avoid useless counterproductive schemes
- nuts and bolts of iQC (statistical techniques)
- cost-effective method selection & validation & iQC-deployment

Teaching Goals

We cover in separate modules
the diagnostic process
uncertainty of measurement
This is prerequisite knowledge for this unit on
statistical process control

We provide supplemental material with
definitions and skills needed to understand this unit
α- and ß-error and power curves

Operational Definition

Statistical Process Control (SPC) : iQC

SPC / iQC by itself is purposeless.
The objective is
to keep a target process under control.

"Expectations is the place you must always go to before you get to where you're going. …"

from The Phantom Tollbooth
by Norton Juster

N. Juster

Operational Definition 1/5

Statistical Process Control (SPC) : iQC

The objective is
to keep a **target process** under control
= Process **Care**

Statistical = an attribute
Statistic = an estimate
Control = **Action**
stemming from
evaluation of a statistic
to exert Control

SPC = to exert control over a process, using statistical diagnostics
≠ Statistics of a Process that is Monitored

Operational Definition 2/5

Intelligent Process Control

- **Monitor** parameters signifying
  - **OUTCOMES of the process under control**
  - **factors key to performance** of that process

The proof of the pudding is in the eating

Key technical requirements

- Using statistical techniques
  interpret whether set targets are achieved (appraisal)

- **Use** that evaluated information
  - in a **continuous Shewart-Deming-cycle**
  - to **tune / adjust :**

the process under control & the control-process

- achieve set targets
- reset targets
- analyze & reduce variability

- operational characteristics
- optimize estimates
- …

Create Value / **Minimize Burden & Cost**

Create Value / Minimize Burden & Cost

Quality =
Value for your Money =
Cost Effectiveness

Operational Definition 3/5

Appraisal & Prevention =

Diagnosis = Decision Making

No decision = Waste without added value
Decisions = CAVE: Error and unwanted costs

Optimize operational characteristics of your diagnostic process

Medical Diagnosis

- by its symptoms
- identify a disease
- to decide about
  - treatment of the patient

Statistical Process Control

- computation based on *observations*, which provide
  - information on the past & current behavior of the system
- algorithms and techniques to determine
  - whether behavior of the system is correct
  - which kind of fault we are facing
  - which part of the system is failing
- decision making
  - to prevent or to remedy failure

= 3-part definition identical to medical diagnosis

Operational Definition 4/5

Cost-Effective Investment in OUTCOME CONTROL

Shewart-Deming cycle articulates with cost objects

Investment in compliance with the specs

- Prevention	Adds Value (*i.a.* up-time) ← Main Focus
- Appraisal	Unwanted costs (*e.g.* false rejection)

Costs of remedies for failures

- Internal Costs	Undesirable costs of scrap and rework
- External Costs	Cost of damages and losses

→ Domain of iQC

Less is better = cutting waste ←

Prevention by fail-proof design is better than quality by inspection

Operational Definition (Summary)

Statistical Process Control (SPC) :

Intelligent system :

- **place a value on an objective at which to aim**
- **measure and keep account of effects of a given action**
  through feedback loops that return
  a message signifying **the suitability of the outcome**
- **interpret the feedback**
  (conceptualize & build model with requisite variety)
- **integrate, decide & act, timely**
- **learn** (from your mistakes)

We translate here the Shewart-Deming cycle
into corresponding **mental processes**.
Lack of intelligence
is often the weak point in current iQC-practice.

Requirements of the Standards

ISO-15189:2003

5.6.1 The laboratory shall design **internal quality control** systems that **verify the attainment** of the **intended quality of results**.
It is important that the control system provide staff members with clear and **easily understood information** on which to base **technical and medical decisions**.
Special attention should be paid to the elimination of mistakes in the **process of handling samples, requests, examinations, reports**, etc.

5.6.2 The laboratory shall **determine the uncertainty** of results, ...

5.6.3 A programme for calibration of measuring systems and **verification of trueness** shall be designed ...

Current iQC-practice
biased towards end-of-line control / rejection actions

ISO-15189:2003

5.6.1 The laboratory shall design **internal quality control** systems that **verify the attainment** of the **intended quality of results**.
It is important that the control system provide staff members with clear and **easily understood information** on which to base **technical and medical decisions**.
Special attention should be paid to the elimination of mistakes in the **process of handling samples, requests, examinations, reports**, etc.

5.6.2 The laboratory shall **determine the uncertainty** of results, …

5.6.3 A programme for calibration of measuring systems and **verification of trueness** shall be designed …

Standard focuses on (is biased towards ?) analytical specs

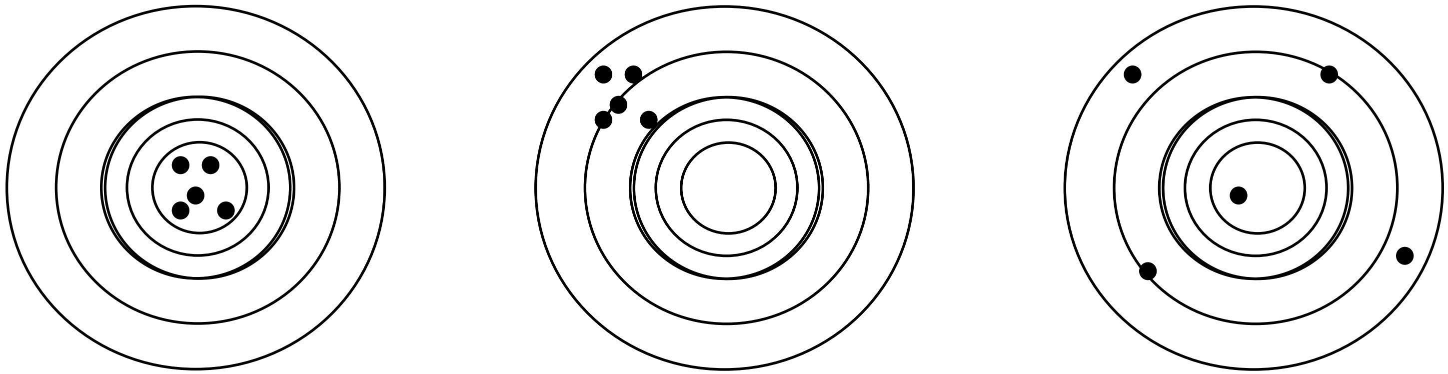
Voice of user / intelligence has to be consequential

Include pre- en post-analytical phases

Defining
the Process of Process-Control

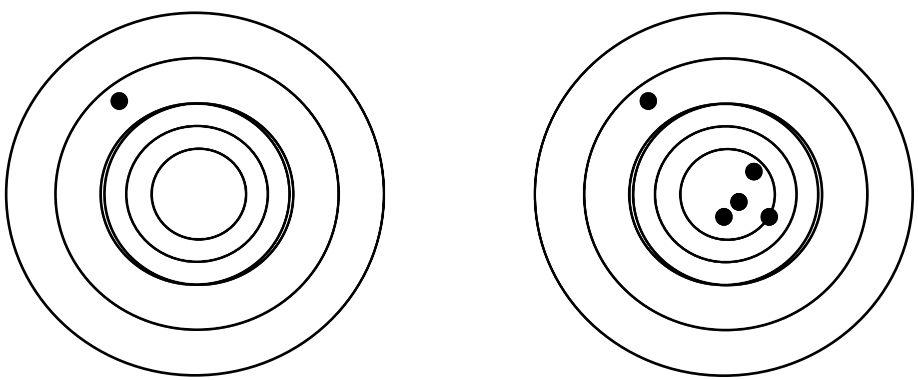
Defining “ Control ”

Shooting-in a gun : descriptive **statistics**



on target inaccurate imprecise

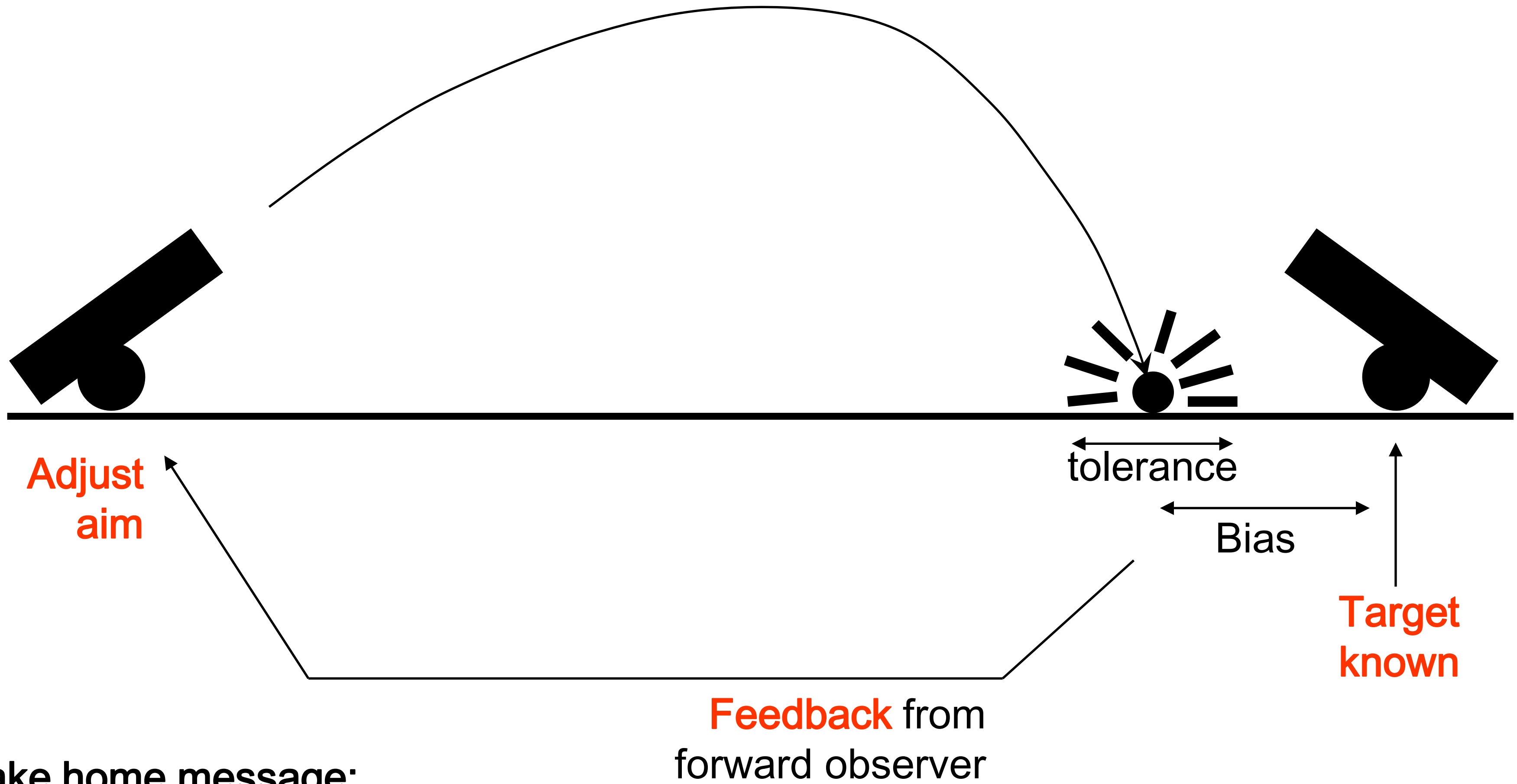
Marksman : **control**



first shot & next shots

CONTROL =
learn from the past
to affect the future

Field artillery: process control



Take home message:

there is ONLY process CONTROL when

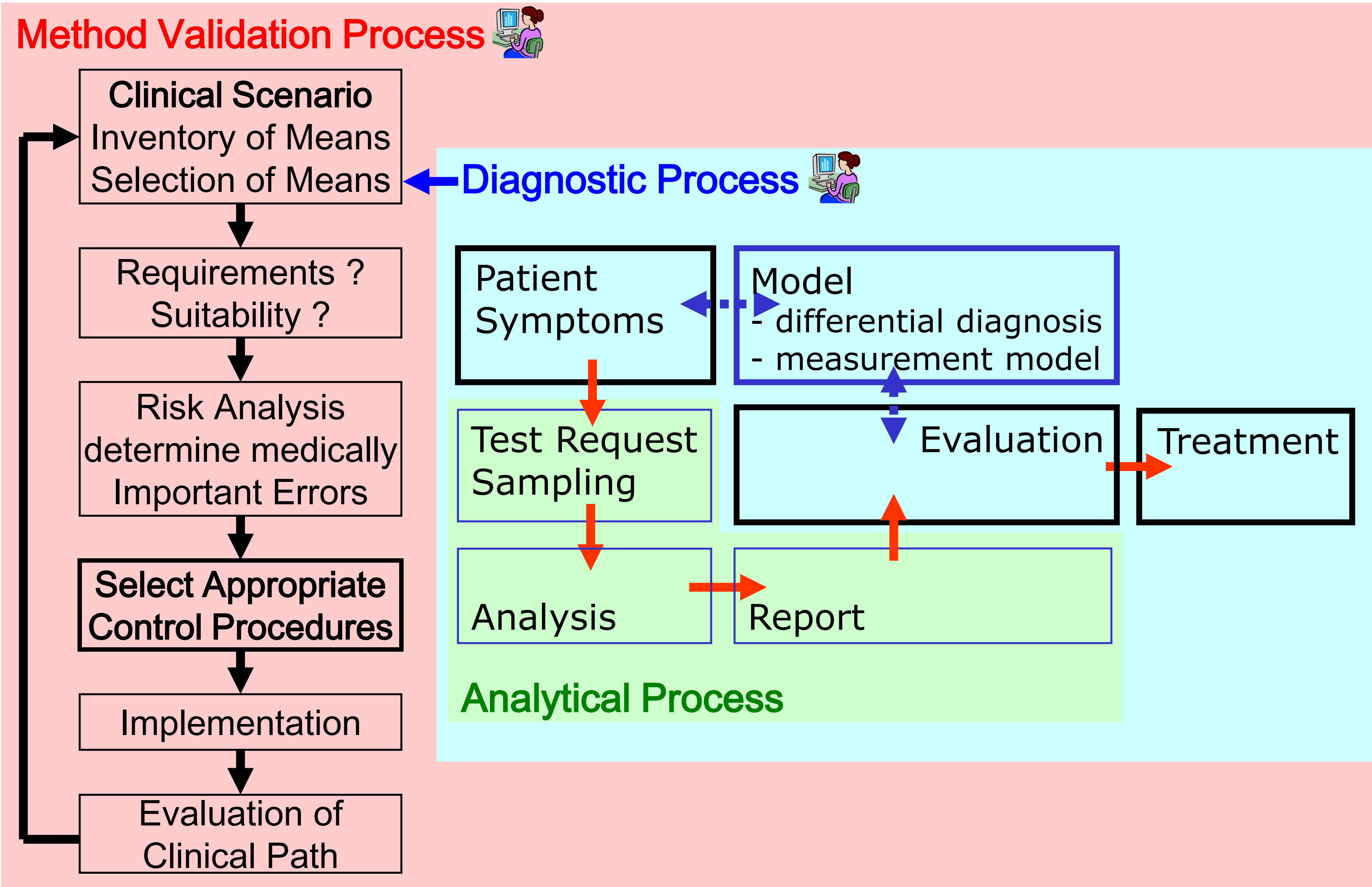
- the TARGET is DEFINED
- & INFORMATION from the past
  is CONVERTED into appropriate ACTION to influence the future

The Process of
Control of the Diagnostic Process

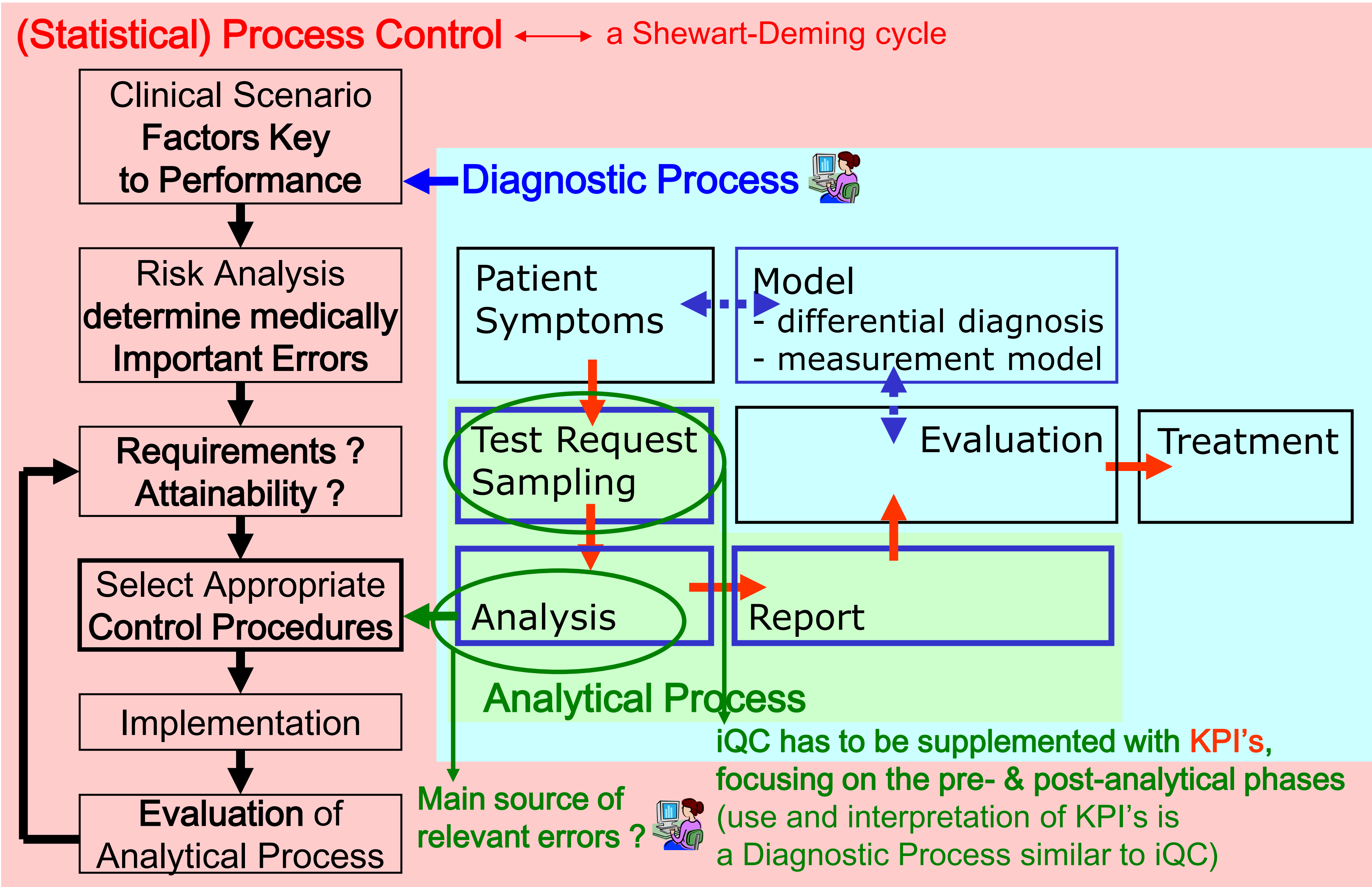
The Broader Panorama (Method Validation)

SPC analyzed as a Diagnostic Process

METHOD VALIDATION: PROCESS FLOW CHART



CONTROL of the ANALYTICAL PROCESS: FLOW CHART

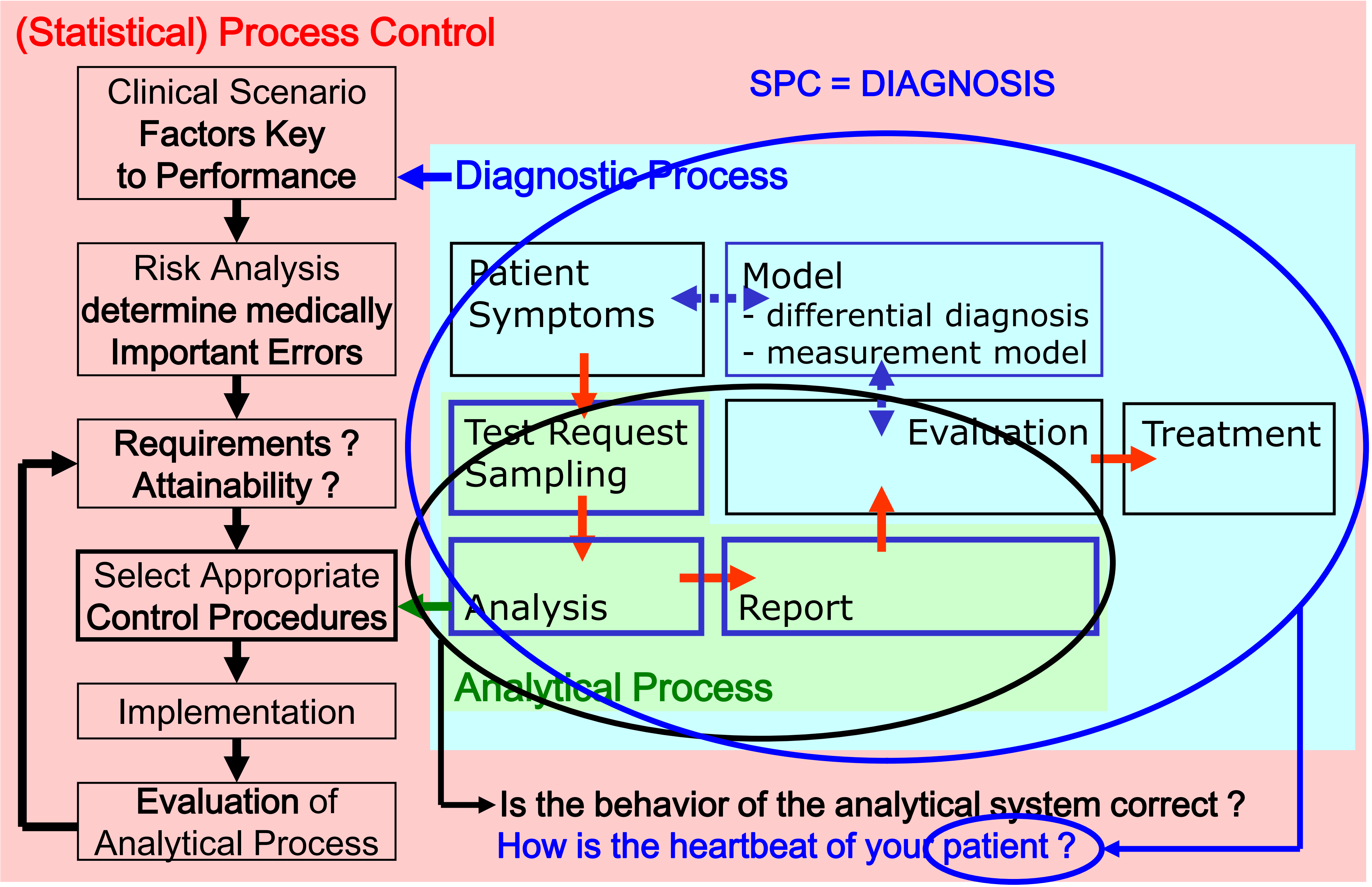


The Process of Control of the Diagnostic Process

The Broader Panorama (Method Validation)

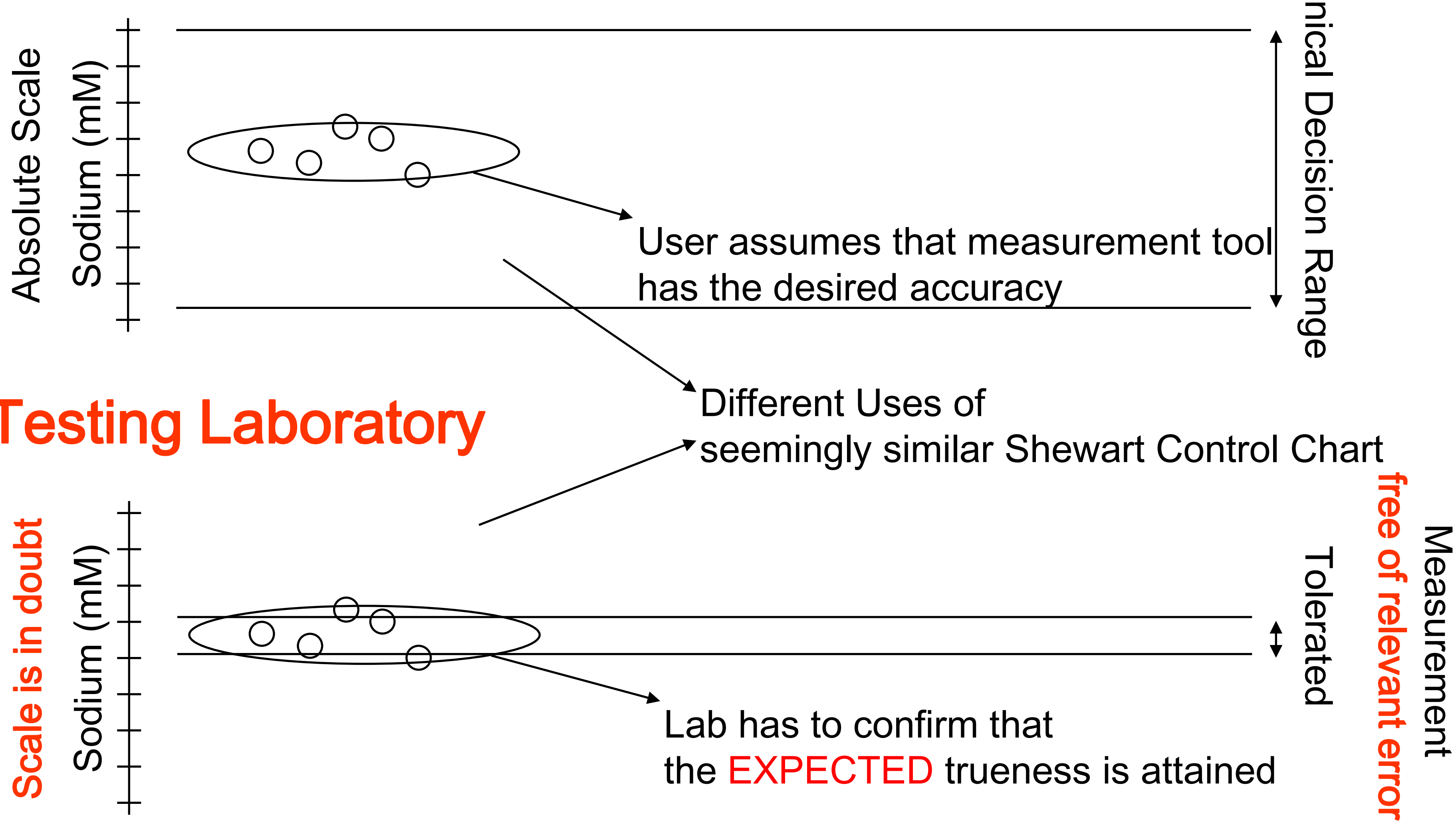
SPC analyzed as a Diagnostic Process

CONTROL of the ANALYTICAL PROCESS: FLOW CHART



Quality Control

Patients Chart



A Course in Medical Pathology

1. General physiology
2. Organ systems
3. Pathology & Propedeutics

A Course in Statistical Process Control

1. The diagnostic process 
2. Uncertainty of measurement 
3. This unit

DIAGNOSIS

- computation based on *observations*, which provide
  - information on the past & current behavior of the system
- algorithms and techniques to determine
  - whether behavior of the system is correct
  - which kind of fault we are facing
  - which part of the system is failing
- decision making
  - to prevent or to remedy failure

But we treat the patient, not his heartbeat !

Statistical Process Control (SPC) : iQC Flow Chart

Define indicators relevant to **outcome** or **key to performance**

Analytical Process

Model (of measurement process)

$x = X + \sum_i [\Delta + e\ \text{random}(0,1)]_i$

Diagnostic Process

Sampling
- Materials
- Patient results
- Levels
- Time frame

Evaluation
- Risk evaluation
- Rules (1-3S, …)
- Operational Characteristics
- …

Actions

Process Control
- Prescribed actions
- Recall
- …

- Fixed
- Sliding
- Cumulative
- …

Estimates

- z-Score / Levey-Jennings
- …

Report

Statistical Control = an appreciation of the past, resulting in effective & efficient actions to influence the future **= purposive**

iQC Analytical Process

Statistical Process Control (iQC)

Concentrating efforts on accurate statistics, disconnected from purposive thinking, is often the weak spot in iQC-practice

Requirements of the Diagnostic Process
Factors Key to SPC-Performance

- Relevant	: USER ASKS for a DIAGNOSTIC PROCEDURE to - detect (pending and nothing but) Relevant Failure - identify Nature of that Failure
- Accurate	: sampling design / data processing
- Timely	: timing iQC / TAT of evaluation
- Accessible	: publication of results / conclusions
- Understandable	: graphical interface / procedures & definitions
- Comparable	: over methods / levels / time frames
- Coherent	: bench marking / coherent time frames
- Complete	: identification of lacking / censored data
- Right price/costs	: high value / low burden to users of procedures

Statistical Process Control

Physiology
1. measurement
2. relevant characteristics

Pathology
3. errors
4. medically relevant errors

Propedeutics
5. semeiology
6. diagnostic power
7. sampling / evaluation / reporting

Principles of Cybernetics

- Systems can be treated as machines
- Control = feedback loops
- Your analysis is as good as your model

  (law of requisite variety =
  model has to be as complex as is needed,
  and not more complicated than that)

W. R. Ashby

The Measurement Model

Subject to Control

Reference

Model

$$x = X_{D-} + \Delta_D + [S_{D,G}\,\text{random}(0,1)] + [S_{D,I}\,\text{random}(0,1)] + \Delta + [e\,\text{random}(0,1)]$$

Bias

Biol variability

between individuals

within individuals

Uncertainty
of Measurement

dependent
on Pathology

Epistemic variables

≠ Constants

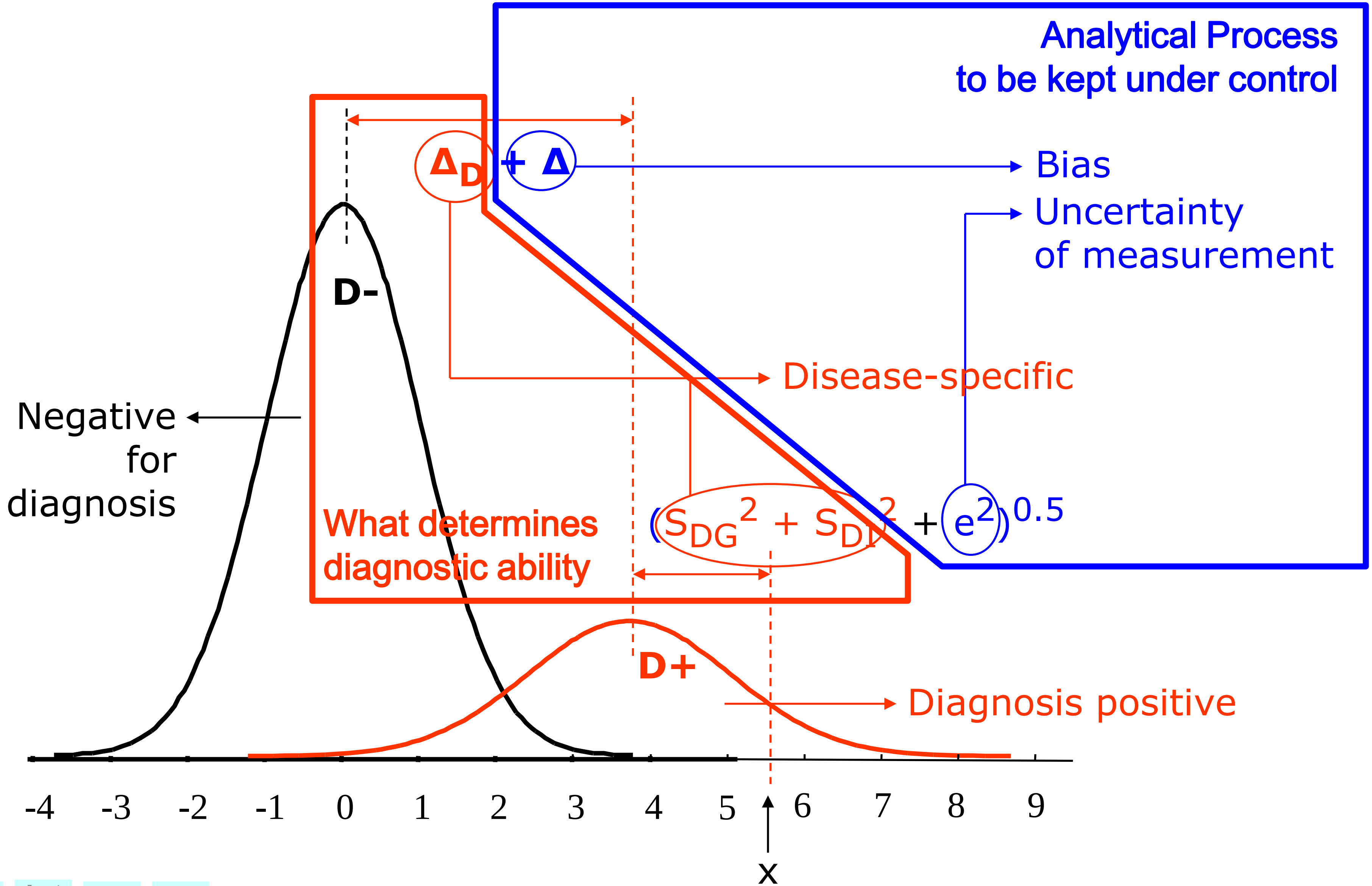
Aleatory variables

Subject to
estimation

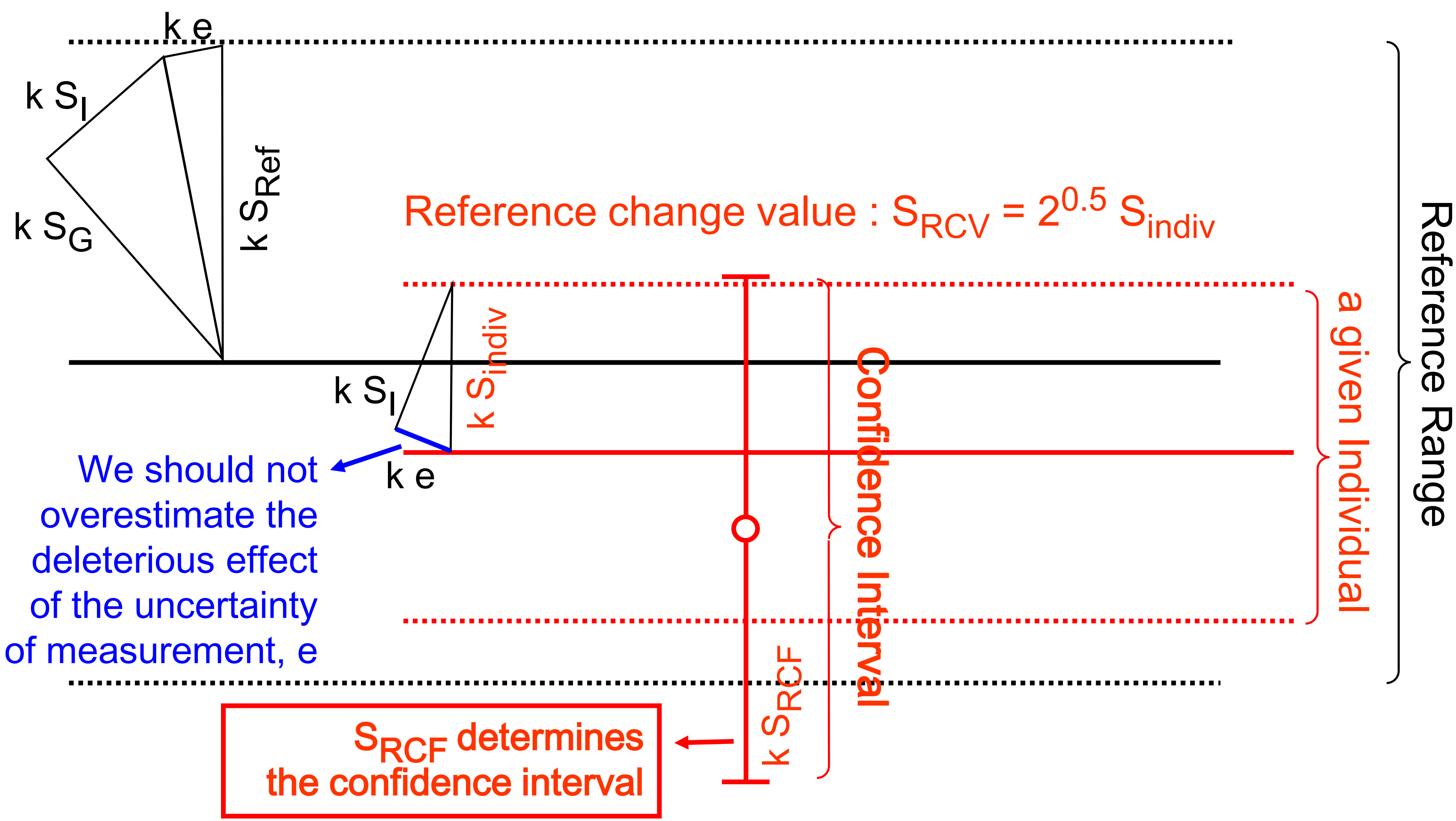
Forbidden zone :
not allowed to act upon

Known in principle, but never perfectly
= Epistemic Uncertainty

The Measurement Model



The contribution of uncertainty of measurement to the uncertainty of interpretation



Process capability of your analytical process

uncertainty of interpretation 

uncertainty of measurement 

Statistical Process Control

Physiology

1. measurement
2. relevant characteristics

Pathology

3. errors
4. medically relevant errors

Propedeutics

5. semeiology
6. diagnostic power
7. sampling / evaluation / reporting

Relevant characteristics of errors

analytical characteristics ← **Analysis of uncertainty of measurement** 

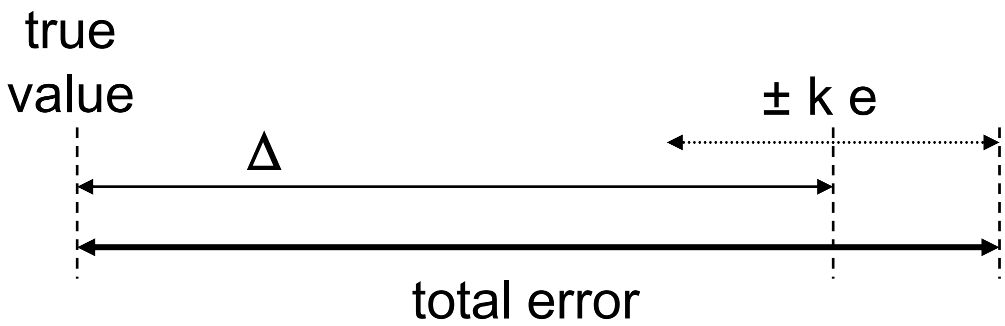
magnitude

random or bias

frequency : intermittent or persistent

sudden or gradual

Total error concept



total error $= \Delta + k\,e$

$k\,e$: Expanded uncertainty of measurement

Δ: Bias

However,

- At best, your method is traceable and gives the best estimate for the truth
  whence $\Delta = 0 \pm$ the epistemic uncertainty
- At worst, bias exists,
  but Δ remains unknown, till it is uncovered,
  and by that fact ends to exist

The bias conundrum

Analytical specs of your actual method

$$e_{Tot} = e_{Within} + e_{Between}$$

Variance from calibrations, reagent batches, …

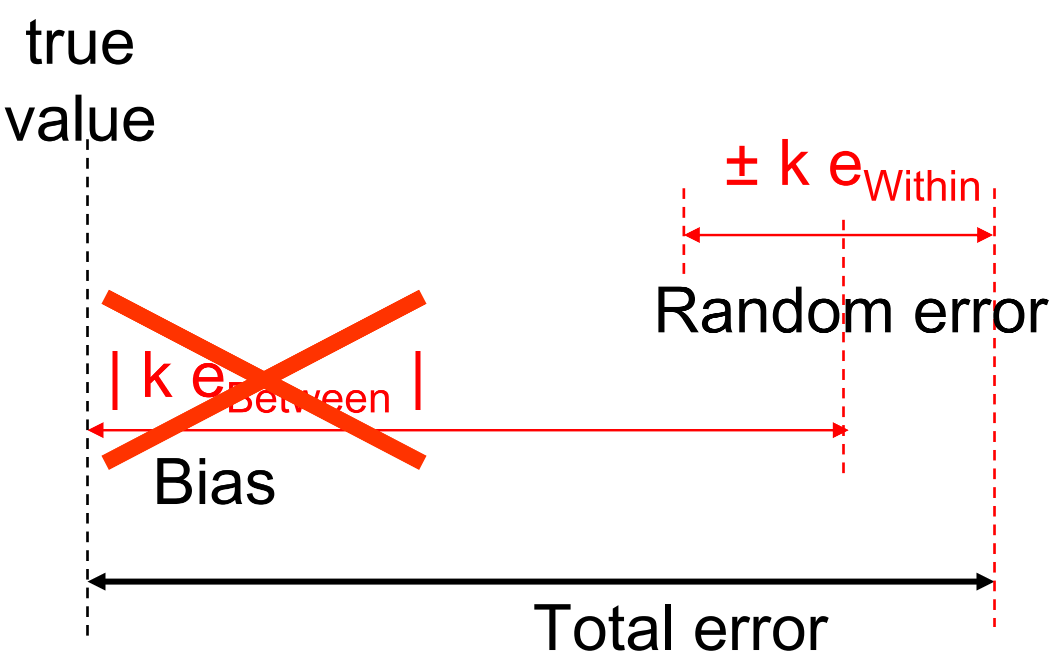
Short-term analytical variance

Long-term total variance

Can you predict the sign ?

If the answer is no,

then **this is random error !**



Relevant characteristics of errors

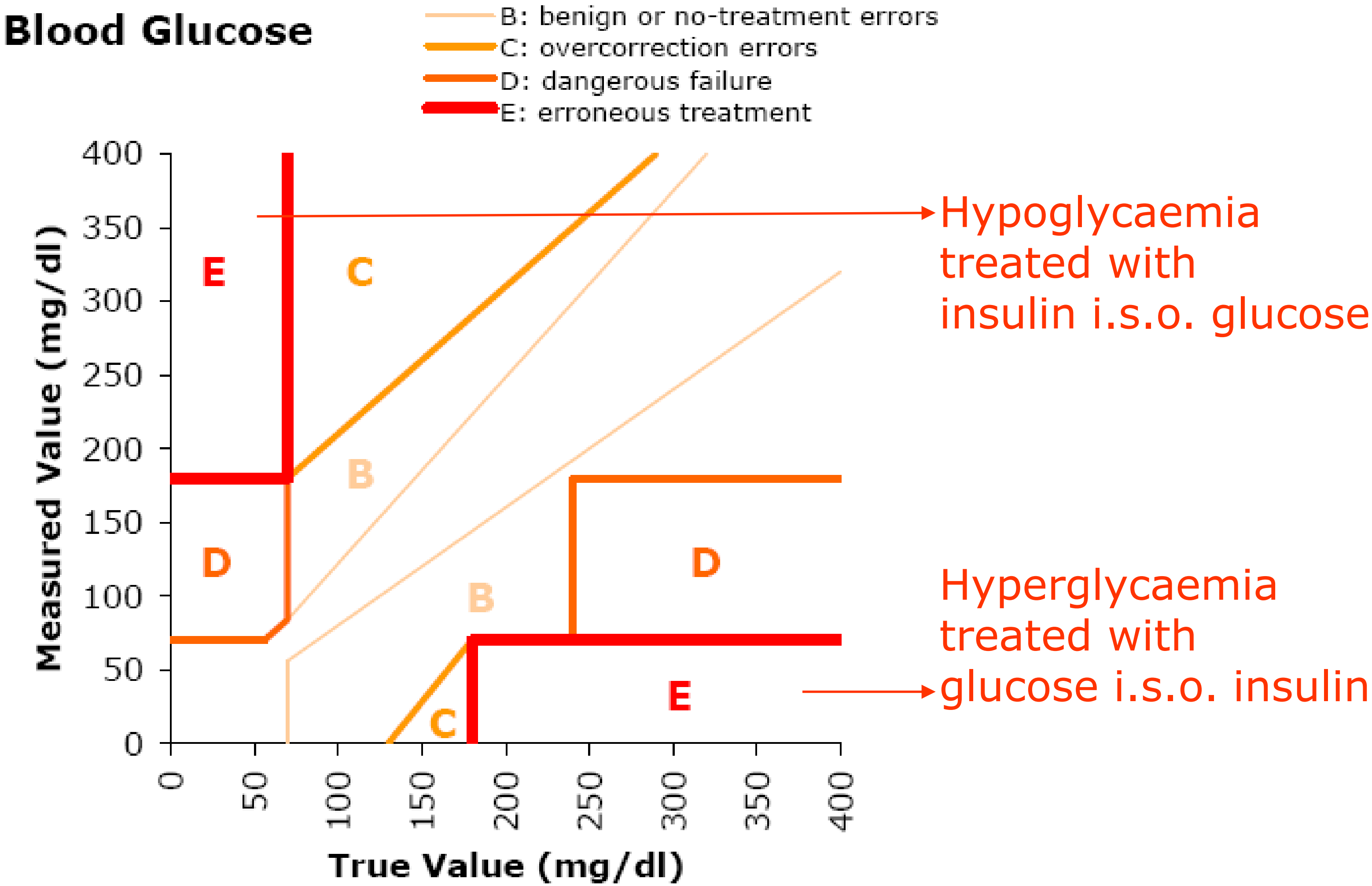
analytical characteristics ← Analysis of uncertainty of measurement

- magnitude
- random or bias
- frequency : intermittent or persistent
  sudden or gradual

medical importance ← Risk Analysis 
method validation 

- critical errors
- associated adverse effects
- frequency, urgency of and window for corrections

Risk analysis



W.L. Clarke, D. Cox, L.A. Gonder-Frederick, W. Carter, S.L. Pohl
Evaluating clinical accuracy of systems for self-monitoring of blood glucose
Diabetes care 10:622-28 (1987)

Statistical Process Control

Physiology
1. measurement
2. relevant characteristics

Pathology
3. errors
4. medically relevant errors

Propedeutics
5. semeiology

6. diagnostic power
7. sampling / evaluation / reporting

Semeiology

statistic relevant to process steering
or the iQC-test of choice depends on

level of measurement
diagnostic scenario
nature of the failure to be detected

Different levels of measurement translate into corresponding models & quality acceptance criteria

Nominal scale
measured property falls into one of 2 classes

Diagnostic Scenario

How many lymfocytes I have to count in order to be confident about the lymfocytosis ?

Ratio scale
The reported result falls into [0,1]

Measurement Scenario
Question = “ **central theorema** ” rephrased for the **binomial distribution** : *cfr.* simulator

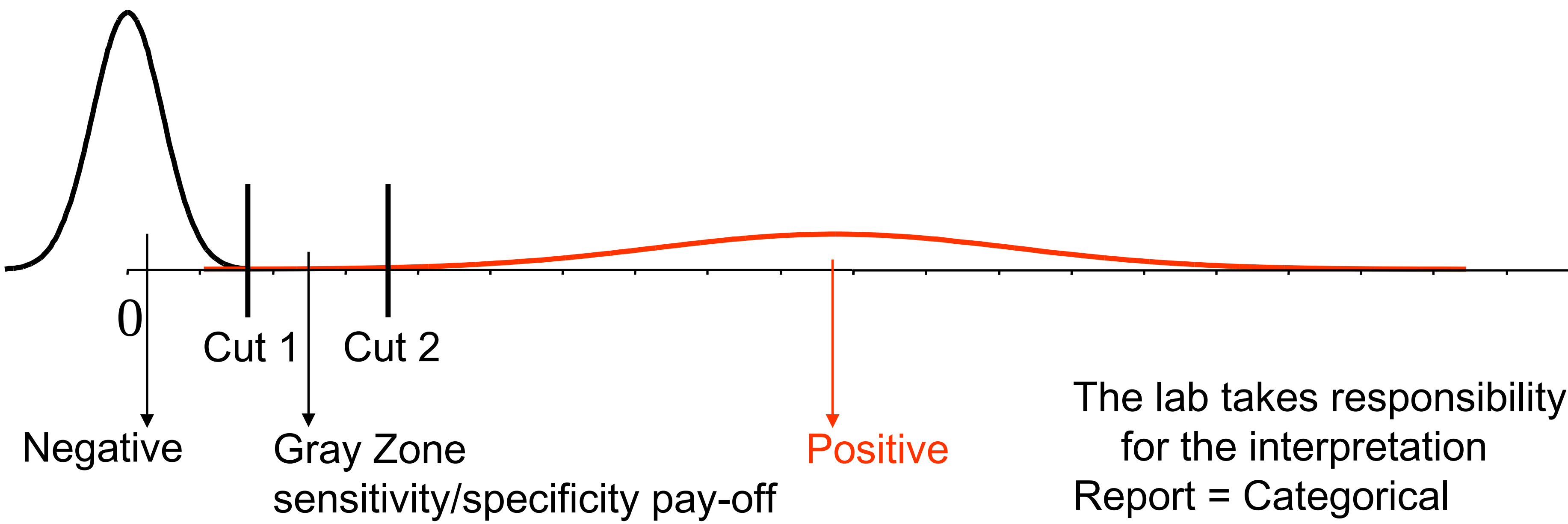
Quality acceptance criteria
for a given cut-off,
an agreed level of confidence is guaranteed
by specifying the **number** of cells **to count**

What remains to be monitored is systematic bias of operators = competence

Agreed = within the production possibilities frontier,
profit from continued counting is offset by costs of lost opportunities

Different levels of measurement translate into corresponding models & quality acceptance criteria

Categorical Report



The lab takes responsibility for the interpretation
Report = Categorical

Categorical classification derived from **measurement of a continuous variable**

Organize iQC at the level of measured numerical variable

Semeiology

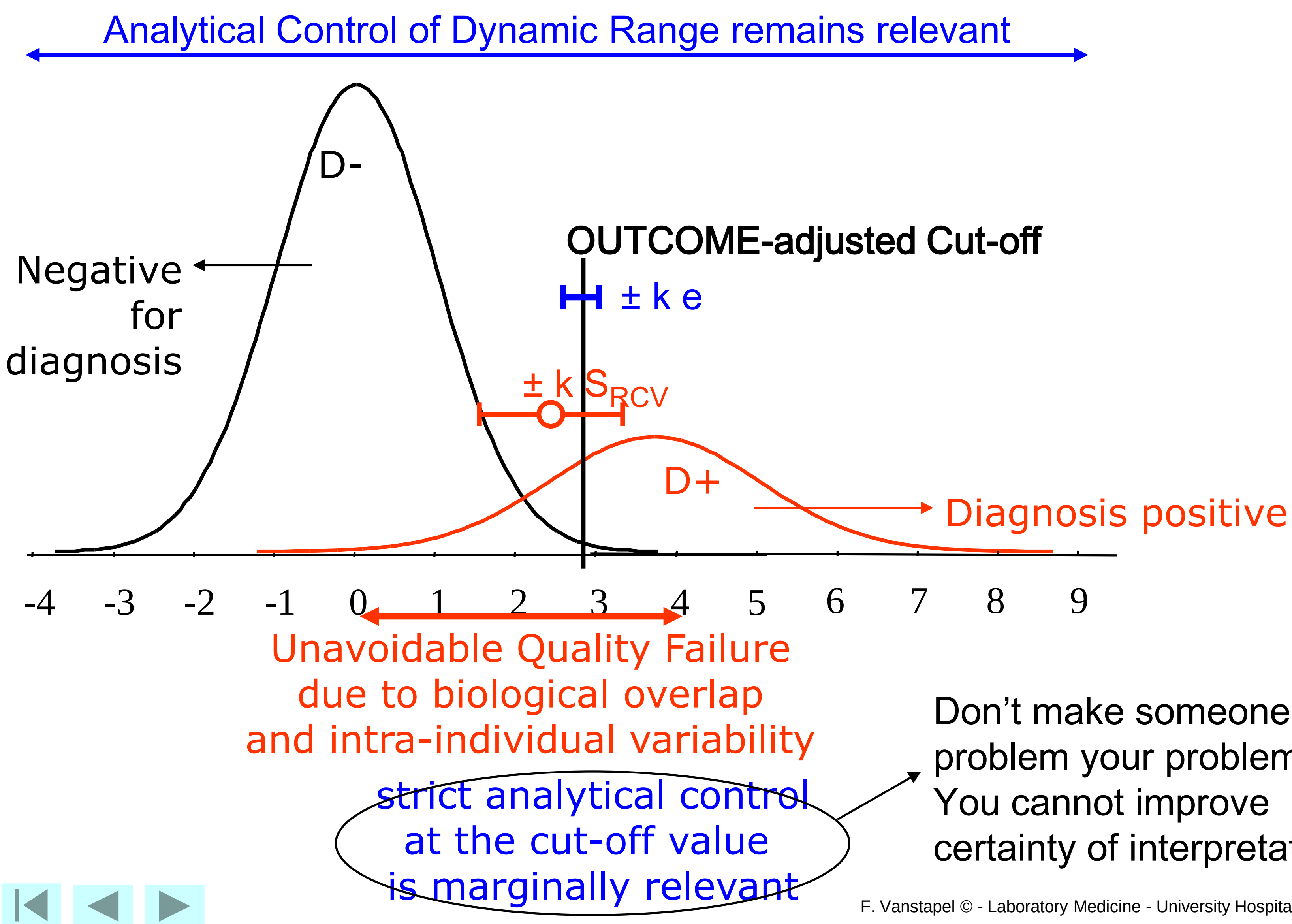
statistic relevant to process steering or the iQC-test of choice depends on

- level of measurement
- diagnostic scenario
- nature of the failure to be detected

Relevant statistics depend on the Clinical Diagnostic Scenario

Diagnostic Scenario's	**Test Characteristics**
Screening Case Finding Differential Diagnosis	Good diagnostic ability = separating power
Staging	Low S_{RCF} & Low bias
Follow-up	Good analytical reproducibility

Relevant Statistics depend on the Clinical Diagnostic Scenario



Don't make someone else's problem your problem. You cannot improve certainty of interpretation.

Process capability of your analytical process

$$\frac{\text{uncertainty of interpretation}}{\text{uncertainty of measurement}}$$

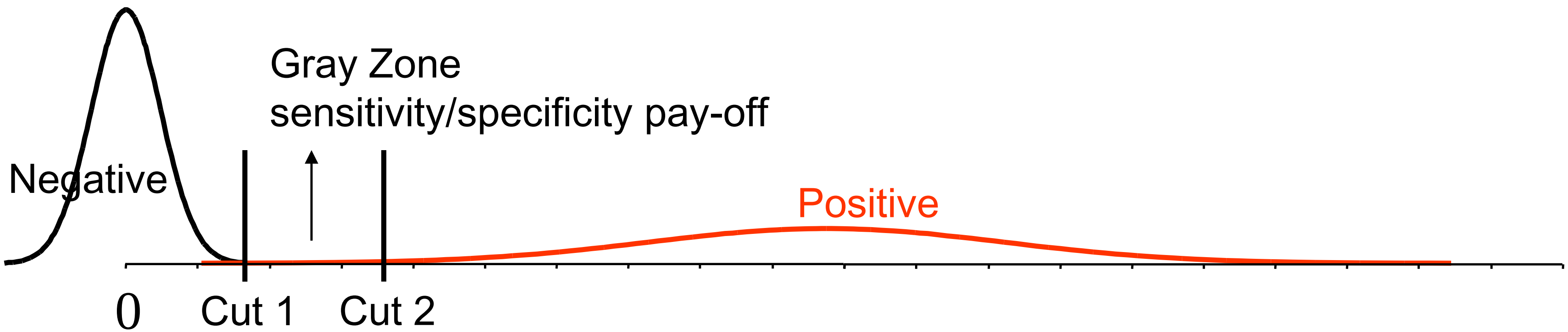

The paradox of " the poor diagnostic test " :
The poorer the diagnostic ability of a test,
the less you have to be concerned
about analytical performance. ☺

When you cannot improve the situation analytically,
shouldn't you select a different test ? ☹

Relevant Statistics depend on the Clinical Diagnostic Scenario

Diagnostic Scenario

Is your patient immunized ?
Is Ab level >> background noise in immune-negative ?



Measurement	**Reporting**	**Anal. Quality Requirement**
Ratio scale, x ELISA-type assay of Ab level	**Ordinal scale** $x \leq X_1 \longrightarrow$ neg $X_1 < x \leq X_2 \longrightarrow$ repeat later $X_2 < x \leq X_H \longrightarrow$ pos	 $\longrightarrow X_H - X_2$ for dynamic range

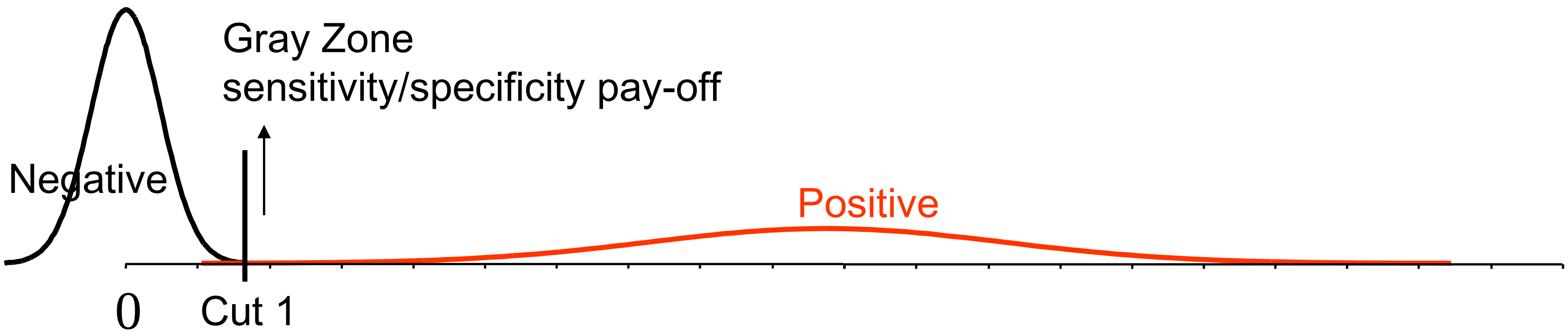
$X_1 - X_2$: grey zone
X_H: upper limit of reporting

Control **dynamic response** at **2 widely divergent levels**

Relevant Statistics depend on the Clinical Diagnostic Scenario

Diagnostic Scenario

Is your patient tumor free ?
Is tumor marker level ~ background noise in tumor-free individuals ?



Measurement

Ratio scale, x
ELISA-type assay
of Ag level

Reporting

Ordinal scale
$x > X_1 \longrightarrow$ pos

Anal. Quality Requirement

$\longrightarrow e/X_1$ for analytical sensitivity

Control at
very low level

Relevant Statistics depend on the Clinical Diagnostic Scenario

Diagnostic Scenario's	Test Characteristics
Screening Case Finding Differential Diagnosis	Good diagnostic ability = separating power
Staging	Low S_{RCF} & Low bias
Follow-up	Good analytical reproducibility

These are the clinical scenario's
where **Control** of the **Analytical Process** is most relevant
Main issue: **Commutability** in time and across the walls of the lab and institutions

Semeiology

statistic relevant to process steering or the iQC-test of choice depends on

- level of measurement
- diagnostic scenario
- **nature of the failure to be detected**

Relevant Statistics depend on Analytical Characteristics of Errors

analytical characteristics	statistical tools
Random Error	Replication
Systematic Error	Difference with target value
Drifts	e.g. Cusum (V-mask)
Periodicity	Time-series analysis
	e.g. Fourier

Statistics : objective criteria & suited for automation

: unavoidably, have limited power

Statistical Process Control

Physiology
1. measurement
2. relevant characteristics

Pathology
3. errors
4. medically relevant errors

Propedeutics
5. semeiology
6. **diagnostic power**
7. sampling / evaluation / reporting

Diagnostic Power

statistics capable of detection of
relevant quality failure

medical relevance
manageable failure
attainable quality
desired quality
operational characteristics of statistical test

We are talking about Process Control

Relevant characteristics of errors

analytical characteristics ← Analysis of uncertainty of measurement

magnitude

random or bias

frequency : intermittent or persistent

sudden or gradual

medical importance ← Risk Analysis 

critical errors

associated adverse effects

frequency, urgency of and window for corrections

manageable → Design of Control Procedure

unavoidable (aleatory) or avoidable (systematic)

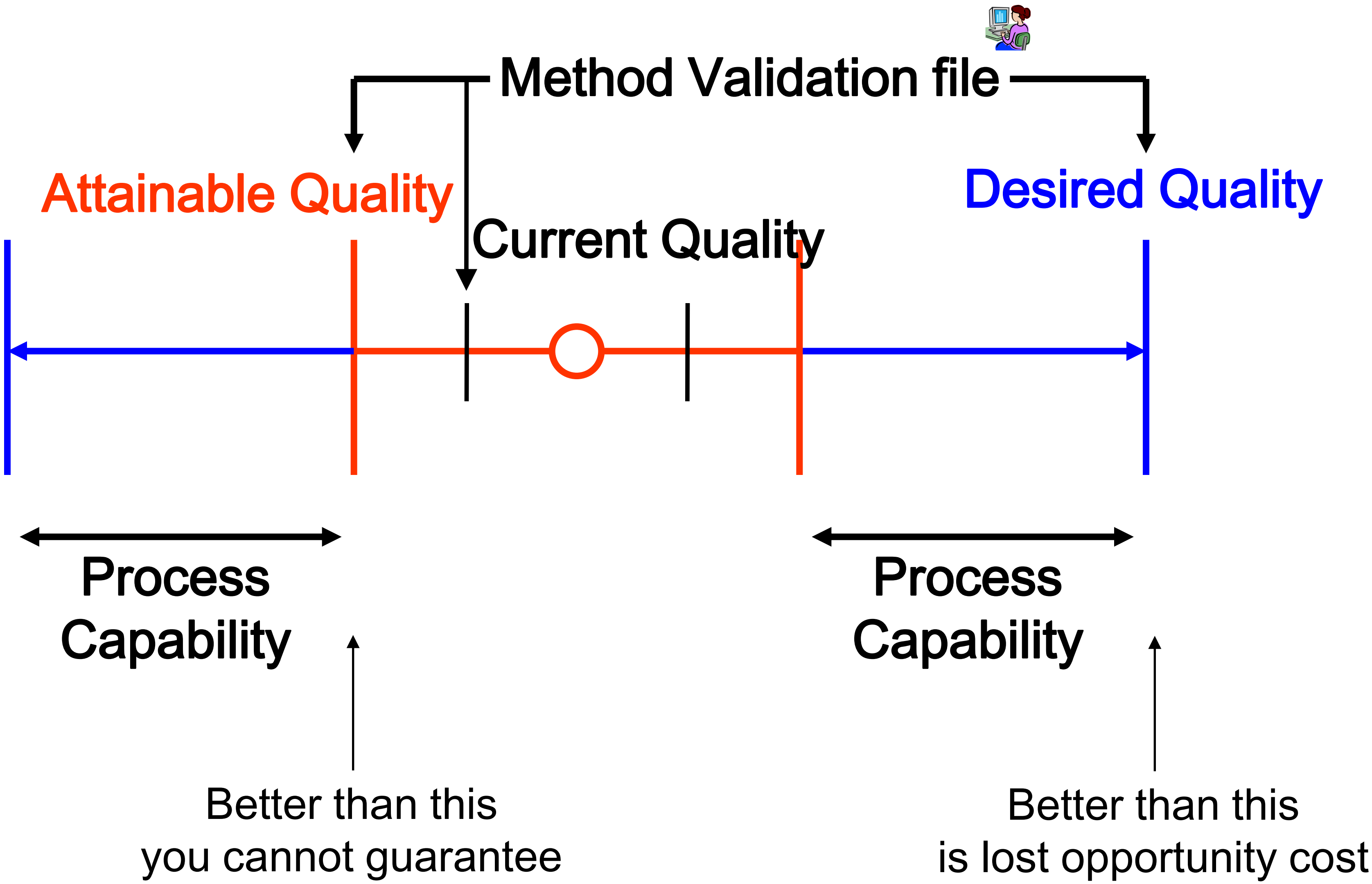
certainty with respect to min & max error frontier

analytical window of opportunity for corrections

Diagnostic Power

statistics capable of detection of relevant quality failure

- medical relevance
- manageable failure
- **attainable quality**
- **desired quality**
- operational characteristics of statistical test


Method Validation file
Attainable Quality
Current Quality
Desired Quality
Process Capability
Process Capability
Better than this
you cannot guarantee
Better than this
is lost opportunity cost

Diagnostic Power

statistics capable of detection of relevant quality failure

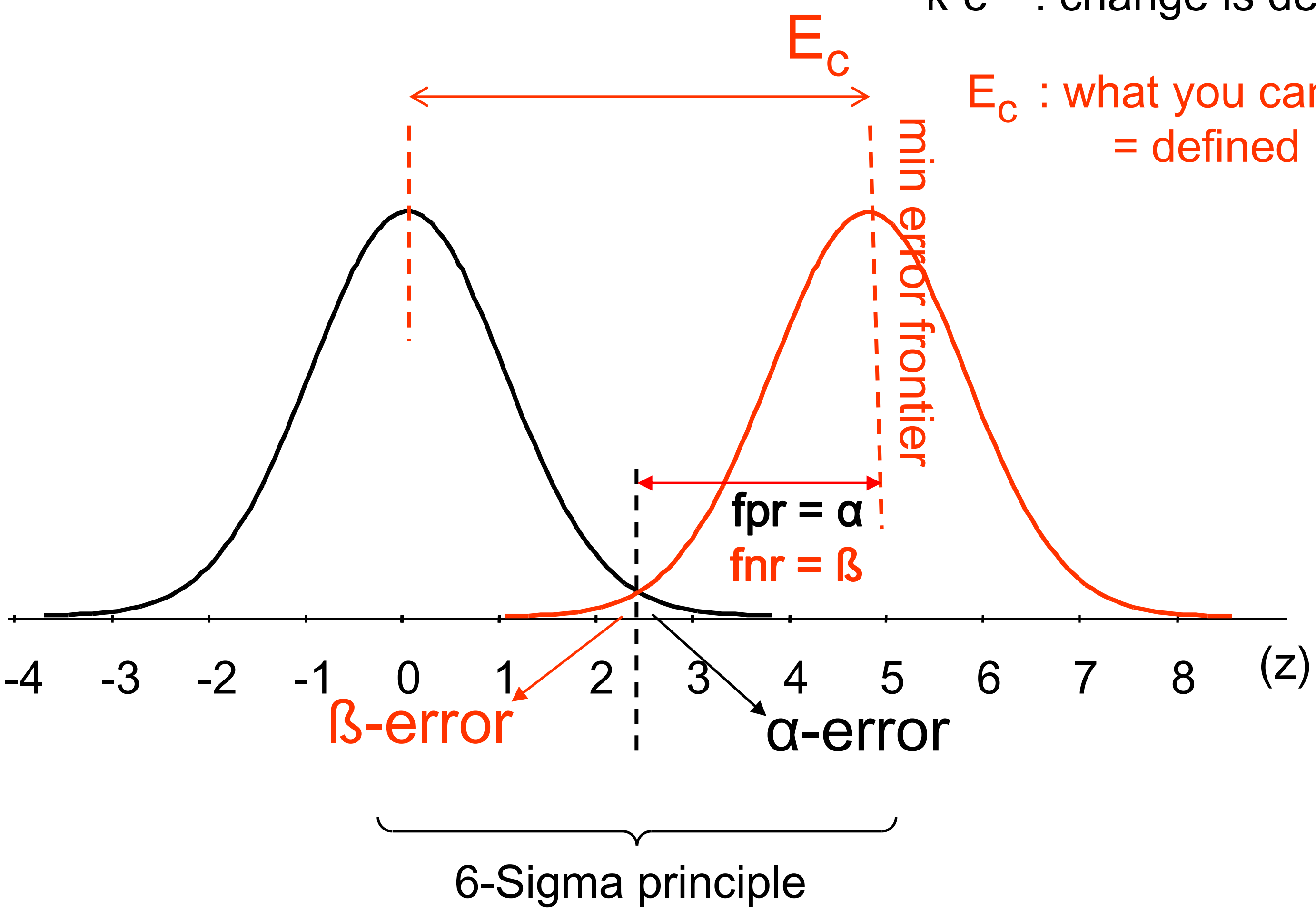
- medical relevance
- manageable failure
- **attainable quality**
- desired quality
- operational characteristics of statistical test

Attainable quality

e : uncertainty of measurement

k e : change is detectable

E_c : what you can guarantee
= defined by ß-error



Diagnostic Power

statistics capable of detection of relevant quality failure

- medical relevance
- manageable failure
- attainable quality
- **desired quality**
- operational characteristics of statistical test

The iQC Diagnositic Process compared with common Clinical Diagnostic Scenario's

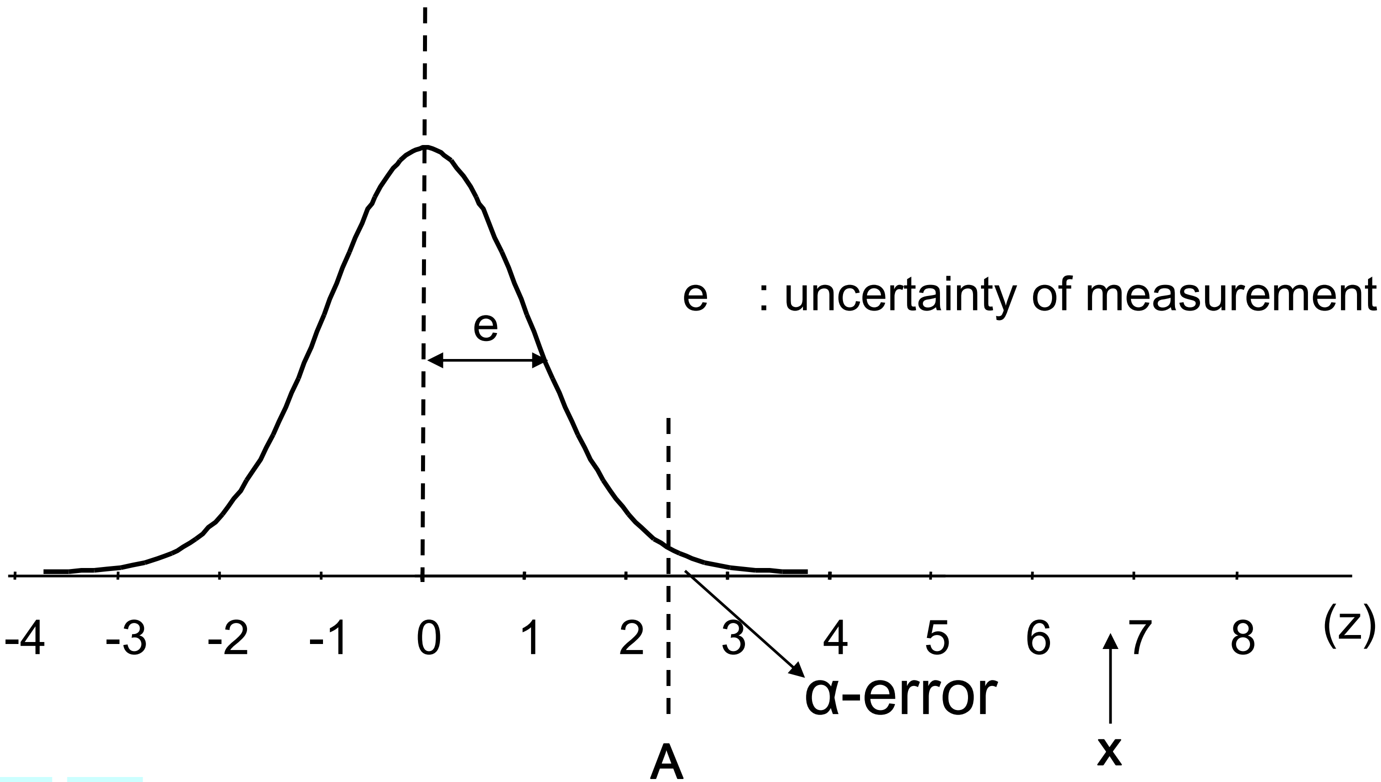
Diagnostic Scenario's	**Intent**
Screening Case Finding Differential Diagnosis	Do you know what you need ?
Staging	Normal behavior / Pending failure / Unacceptable Quality Failure
Follow-up	Detection of likely changes
	Do you know what to expect ?

The iQC Diagnositic Process compared with common Clinical Diagnostic Scenario's

You don't know what you need ?

H_0-model: x > A

P < α : statistically significant, <u>but is it relevant ?</u>



Diagnostic Power

statistics capable of detection of relevant quality failure

- medical relevance
- manageable failure
- attainable quality
- **desired quality**
- **operational characteristics of a test**
  - diagnostic ability of iQC-rules
    - & capability of the analytical process
  - the concept of power curves

iQC Diagnostic Tests

- computation based on *observations*, which provide
  - information on the past & current behavior of the system
- algorithms and techniques to determine
  - whether behavior of the system is correct
  - which kind of fault we are facing
  - which part of the system is failing
- **wheighing**
  - is the failure relevant ?

There is only PROCESS CONTROL
- when a relevant target is secured
- in a timely fashion
- with little capacity loss due to inadequacies of the control process

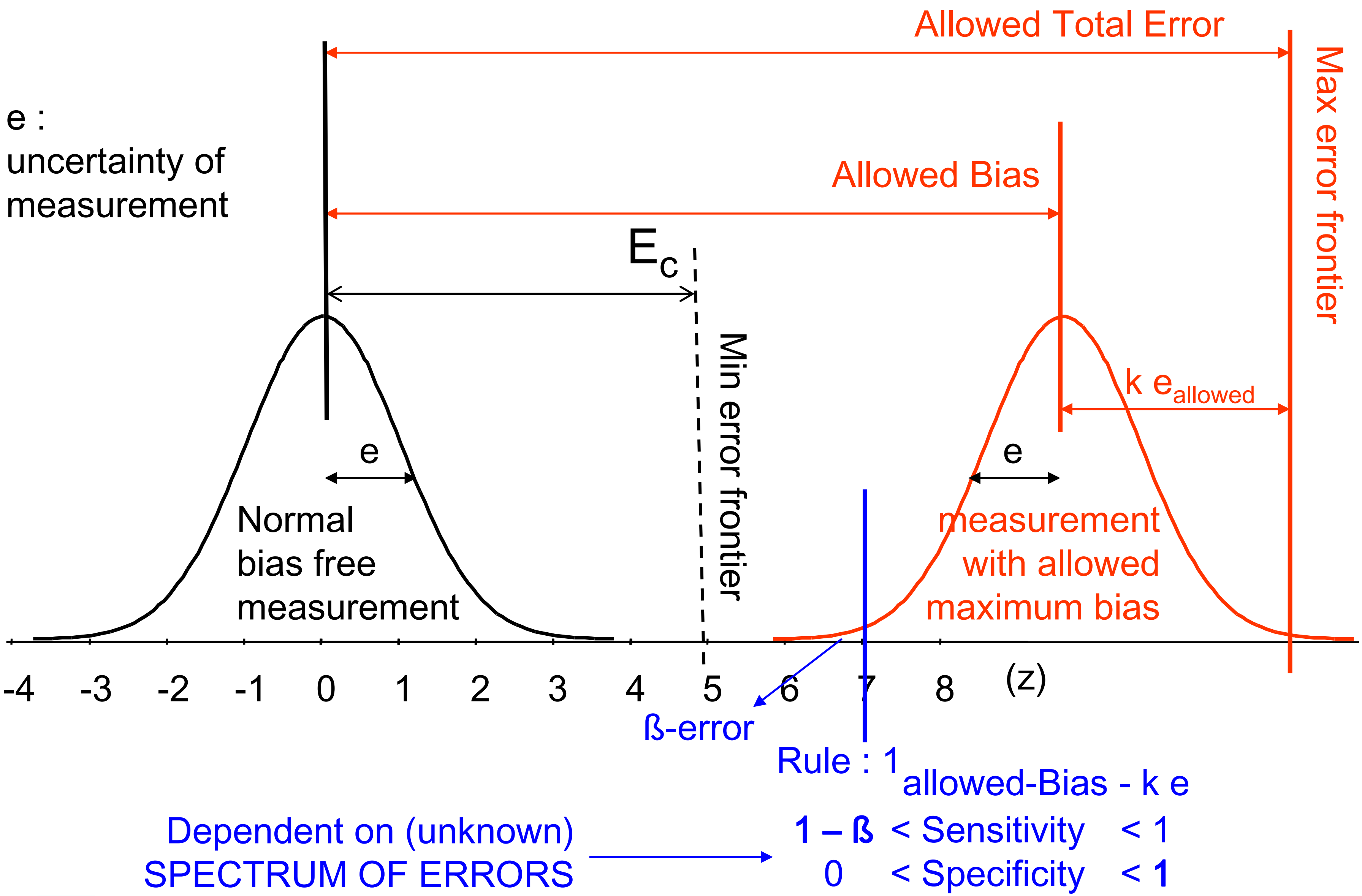
Diagnostic Ability

- **minimize miss-classification costs**

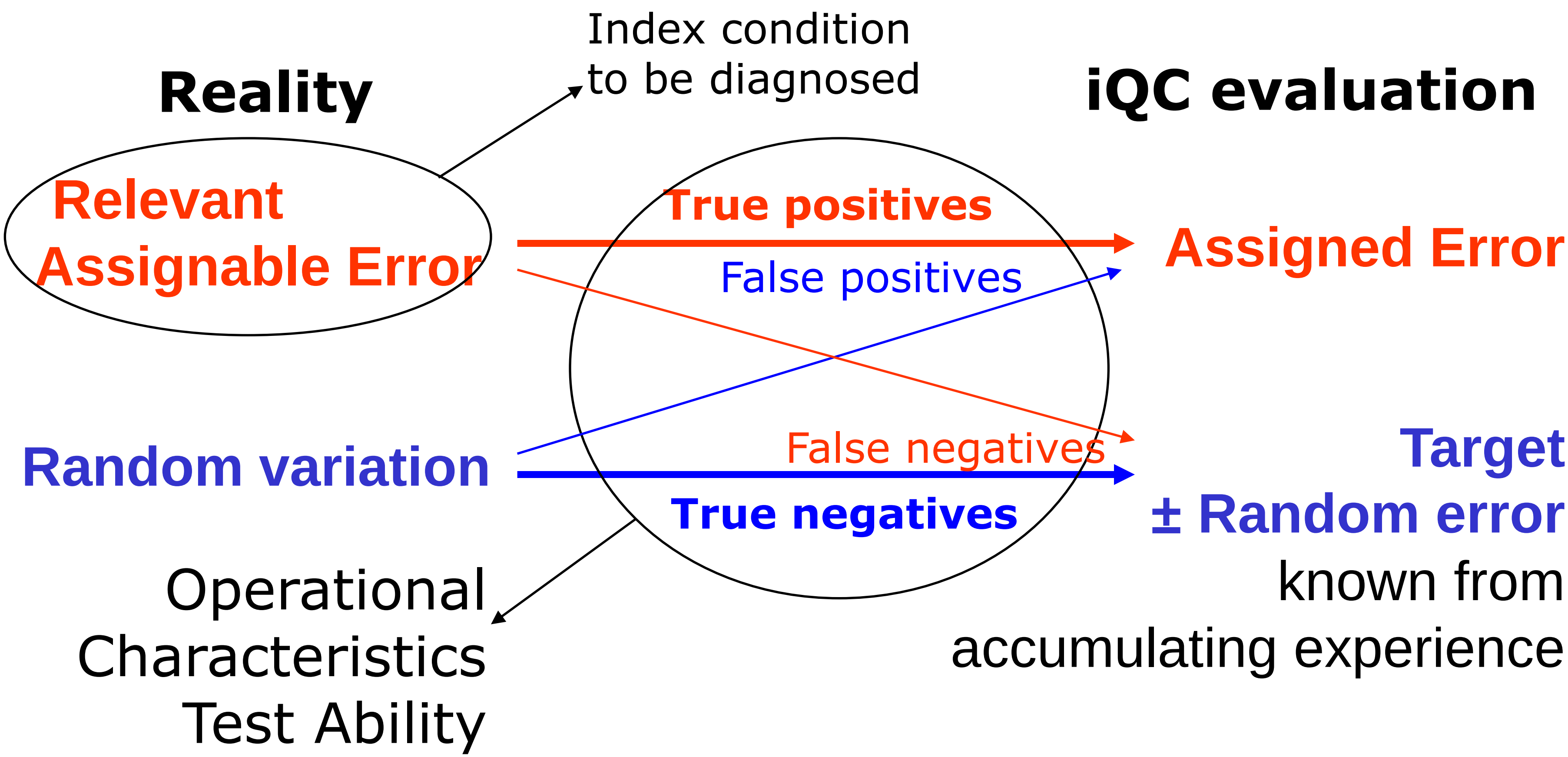
  $fdr * cost_{fd} + fnr * cost_{fn}$

The corresponding power curve only delivers a solution if an index condition is defined (H_1-model)

The ideal of worlds : Max error frontier >> Min error frontier



iQC is a diagnostic procedure



Test-drive our simulator 

The ideal of worlds : Max error frontier >> Min error frontier

Pre-test prevalence	Rule $1_{\text{allowed-Bias - k e}}$	Post-test predictive value

Chronic Persistent Error

5% of the measurements will fail to detect the problem

~ 100%	Sensitivity 0.95 Specificity 0.99	Neg result : ~ 0% Pos result : ~ 100%

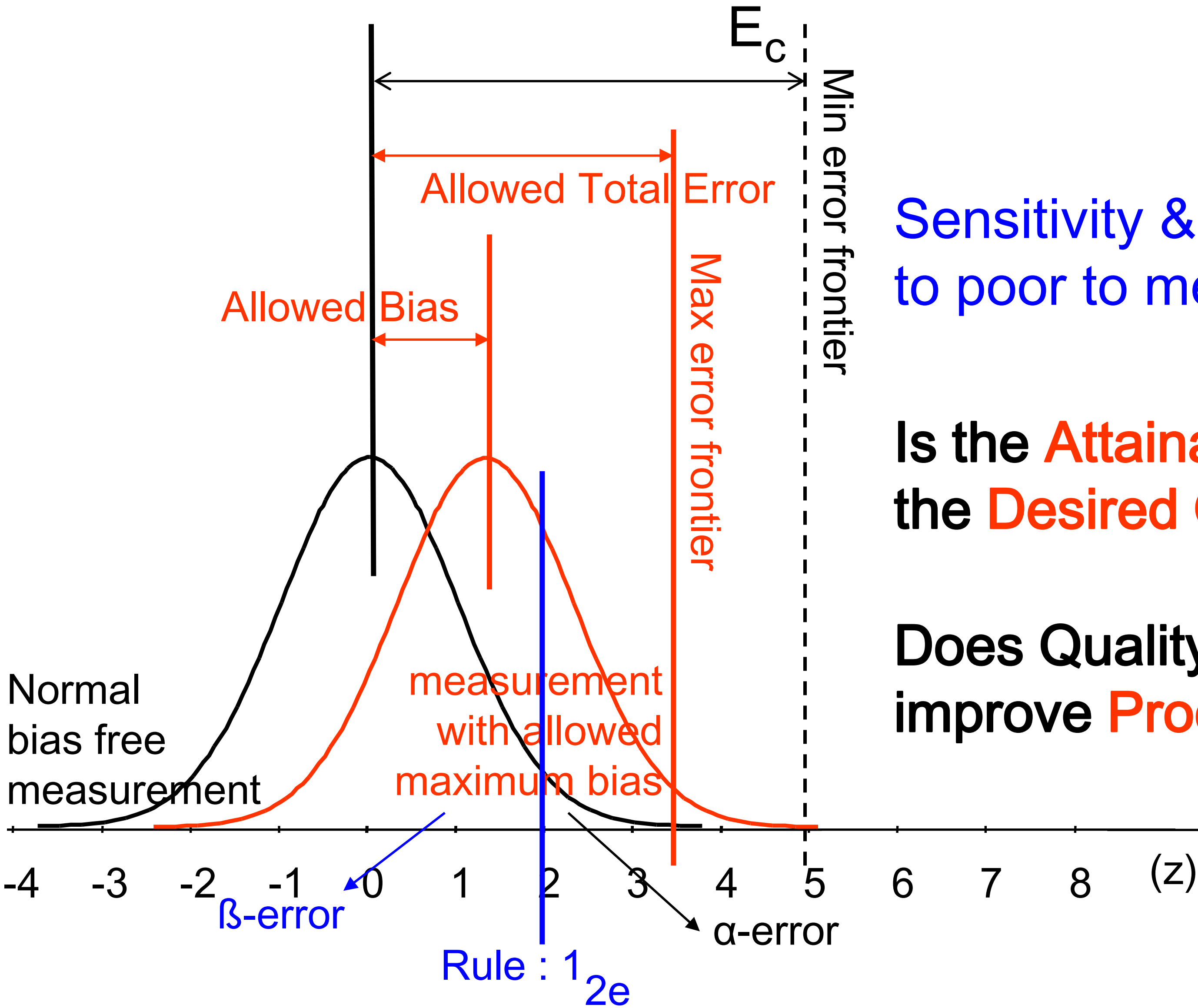
Sporadic Non-Persistent Error

~ 0.1%	Sensitivity 0.95 Specificity 0.99	Neg result : ~ 0% Pos result : ~ 10%

1% of the measurements gives rise to an alarm that will be 90% false positive

Does internal Quality Control improve Process Capability ?

The worst of worlds : Max error frontier < Min error frontier



Sensitivity & Specificity
to poor to mention

Is the Attainable Quality the Desired Quality ?

Does Quality Control improve Process Capability ?

"You'll find,
that the only thing you can do easily is be wrong,
and that is hardly worth the effort."

from The Phantom Tollbooth
by Norton Juster

Diagnostic Power

statistics capable of detection of relevant quality failure

- medical relevance
- manageable failure
- attainable quality
- **desired quality**
- **operational characteristics of a test**
  - diagnostic ability of iQC-rules
    - & capability of the analytical process
  - the concept of power curves

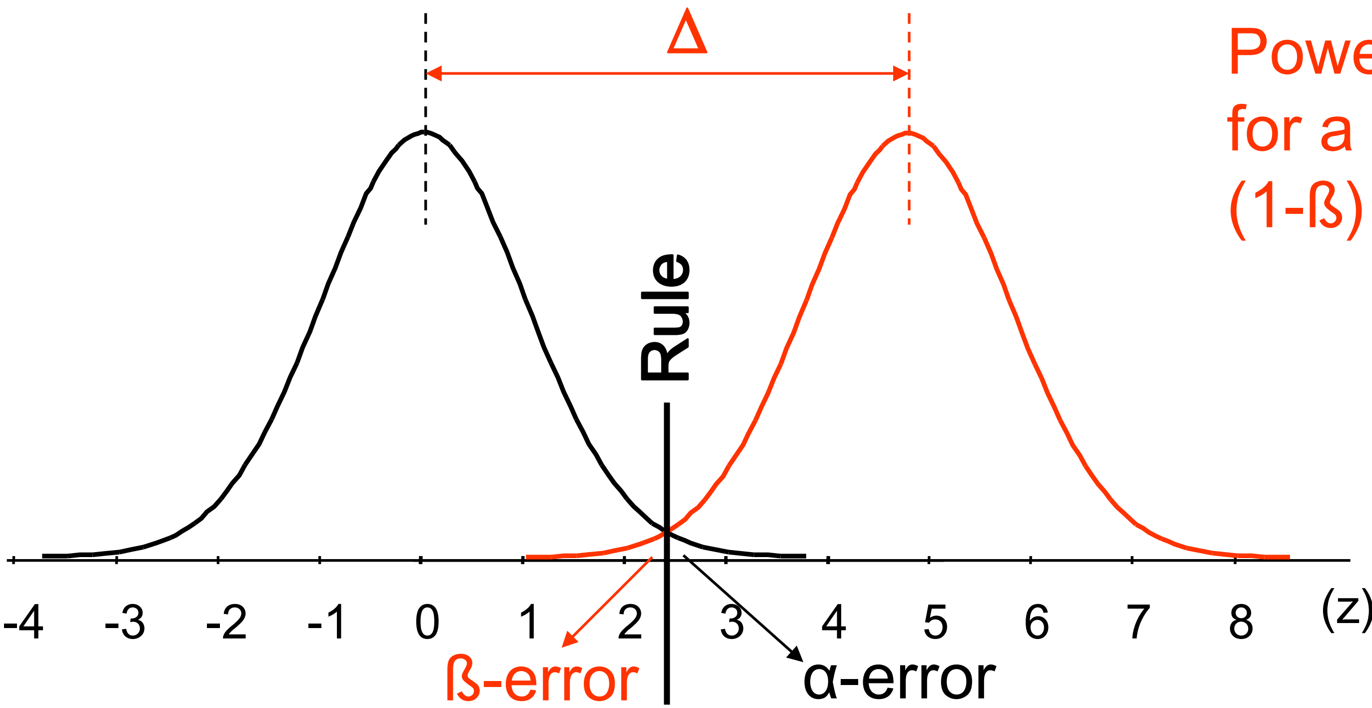
The concept of power curves

The first basic problem in any course of statistics :

Sampling design :

How big a difference (Δ) between samples can I detect ?

What is the required size of the sample ?



Power Curve
for a given Rule :
(1-ß) = f (Δ)

Test-drive our calculator 

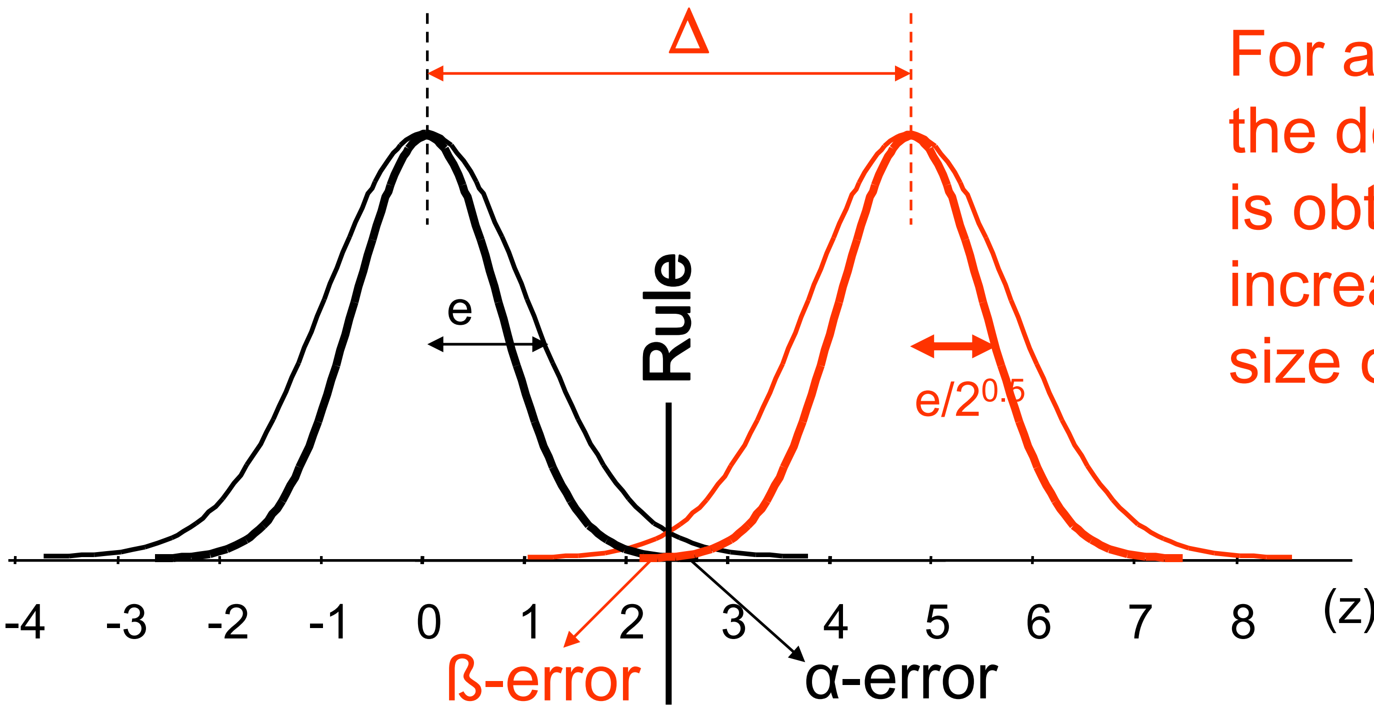
The concept of power curves

The first basic problem in any course of statistics :

Sampling design :

How big a difference (Δ) between samples can I detect ?

What is the required size of the sample ?



For a fixed Δ, the desired α & ß is obtained by increasing the size of the sample

Test-drive our calculator

The worst of worlds : detecting error near the Min error frontier

Cases :

detection of unallowable error < Min error frontier

detection of pending error

Solution : n ⇈⇈

Frequent (pseudo-)real-time sampling

chemical industry : in-process continuous monitoring
airspace industry : fly-by-wire techniques
clinical lab : statistics on patient results (e.g. AON) ☺

Frequent sampling
from the viewpoint of statistical control
= the sample size
form the viewpoint of process control
= urgency & window of opportunity
= n / time frame

The worst of worlds : detecting a sporadic error

Solution : n ⇈ ?

Increase sample size to improve power of your test

Chronic Permanent error	↔	Sporadic error
☺ n ⇈ strategy works		n ⇈ is counterproductive ☹

Statistical Process Control

Physiology
- 1. measurement
- 2. relevant characteristics

Pathology
- 3. errors
- 4. medically relevant errors

Propedeutics
- 5. semeiology
- 6. diagnostic power
- **7. sampling / evaluation / reporting**
  - **materials & sampling frequency**
  - epistemic estimates
  - bias detection
  - reporting & interpretation

Sampling : available materials

Epistemic variables to be estimated	Available materials: Patient samples	Available materials: iQC artificial materials
e uncertainty of measurement	repeat on same sample	repeats
Δ bias	average of patients (AOP)	shifts
S_R, S_D, Prevalence	spectral analysis	-

Underrated Overrated
in current iQC-practice

Sampling materials : pro & contra

Issues	Materials to select: Patient samples	Materials to select: iQC materials
Accurate ☹ Requires validation	☺ AON ± S_{ref} = constant ☹ level = f (ref range) ☺ n / time = large spectrum = relevant filtering needed ?	relevant levels selectable ☺ non-titrated stability ? ☹ handling artefacts ☹
Timely	continuously produced instantaneously available ☺	scarcely sampled ☹
Comparable between methods	directly ☺	matrix effects ? ☹
	Underrated	Overrated

in current iQC-practice

Statistical Process Control

Physiology
1. measurement
2. relevant characteristics

Pathology
3. errors
4. medically relevant errors

Propedeutics
5. semeiology
6. diagnostic power
7. sampling / evaluation / reporting
    - materials & sampling frequency
    - epistemic estimates
    - bias detection
    - reporting & interpretation

ESTIMATES – NUTS and BOLTS (1/3)

Central limit theorema :
the **best estimate** results from **continuous accumulation**

Changes in components of the measurement system
which can give rise to changes in the results
and of which **the sign of the change cannot be predicted**
have to be treated as **random error** & not as bias.
Recipe: ANOVA
Test-drive our simulator

That analysis is your analysis of measurement uncertainty
is part and parcel of your method validation
and has as purpose to reduce risk by
detecting avoidable and curing unwanted fluctuation

ESTIMATES – NUTS and BOLTS

EURACHEM Guidance Doc. No 1
WELAC Guidance Doc No. WGD 2

14.16
Once a catalogue of uncertainties is available,
the straightforward combination of the standard deviations
is appropriate, applying the general law of error propagation.
This is achieved by taking the square root of
the sum of squared contributing uncertainty components,
all expressed as standard deviations.

What do you need to realize this ?

- automated log of individual events that are source of individual error
  e.g. calibrations, lots of calibrators, lots of reagents, …
- the above ot the acquisition time as part of the raw data record

“Oh, this won’t take a minute.
I’m the official Senses Taker,
and I must have some information
before I can take your senses. … ”

from The Phantom Tollbooth
by Norton Juster

ESTIMATES – NUTS and BOLTS (2/3)

The above principles can be automated

All other (non-lean) strategies incur only error & waste

treating random error as bias
- starts from erroneous premises
- asks for attention to and action in response to
  - unavoidable variance
- & thus has to induce (avoidable) error

procedurally scheduled readjustment of target values
- calls for judgments when nothing has to be decided
- & therefore can only deliver erroneous decisions

ESTIMATES – NUTS and BOLTS (3/3)

All statistical estimates have limited power

Statistical estimates are prone to interferences

e.g. cumulative average will after a while no longer detect a constant bias

e.g. cumulative average may blanket a slow drift

e.g. cumulative estimate of variance may blanket a saw-tooth function

e.g. rules with increased sample size blanket sporadic events

Recipe :

- multiplex rules
- VISUAL INSPECTION OF THE RAW DATA

Is what I always do on audits

The bias conundrum (1/2)

Bias exists in simulations as an experimental variable.

In real life, bias exists in our minds only as a hypothetical.
As long as bias remains unknown,
it can persists, and cause havoc.
As soon as it is known,
you (mentally) compensate for it and it stops to exist.

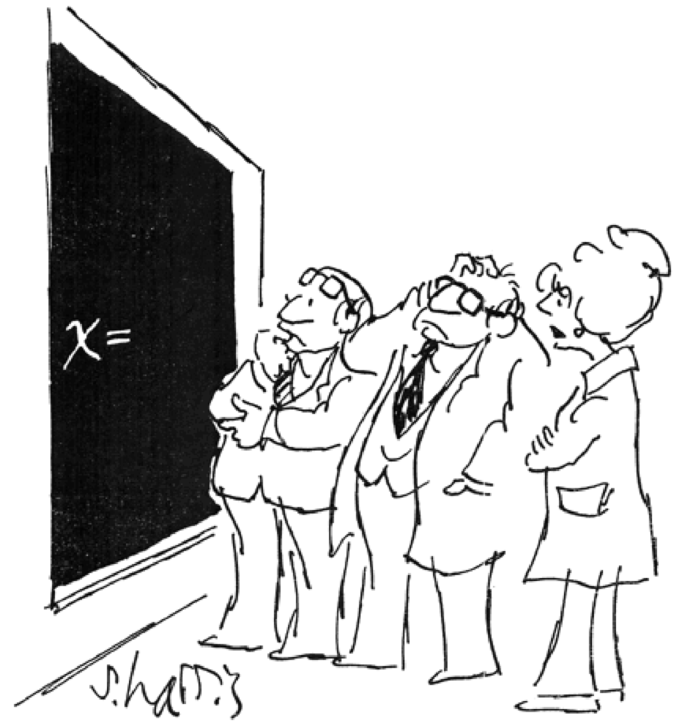
The bias conundrum (2/2)

We know the epistemic variables in our models
as constants, in principle known.

These variables are estimated
with a certain level of confidence
= epistemic uncertainty

An example from the chemical industry :
the mass per packaged unit varies (= random error)
the purity can only be < 1, and thus has a fixed sign (= bias)
the amount of reagent per packaged unit varies
around an average (= epistemic variable ± random error)

Epistemic uncertainty is a fact of life



"If something is there,
you can only see it with your eyes open,
but if it isn't there,
you can see it just as well with your eyes closed. "

from The Phantom Tollbooth
by Norton Juster

Statistical Process Control

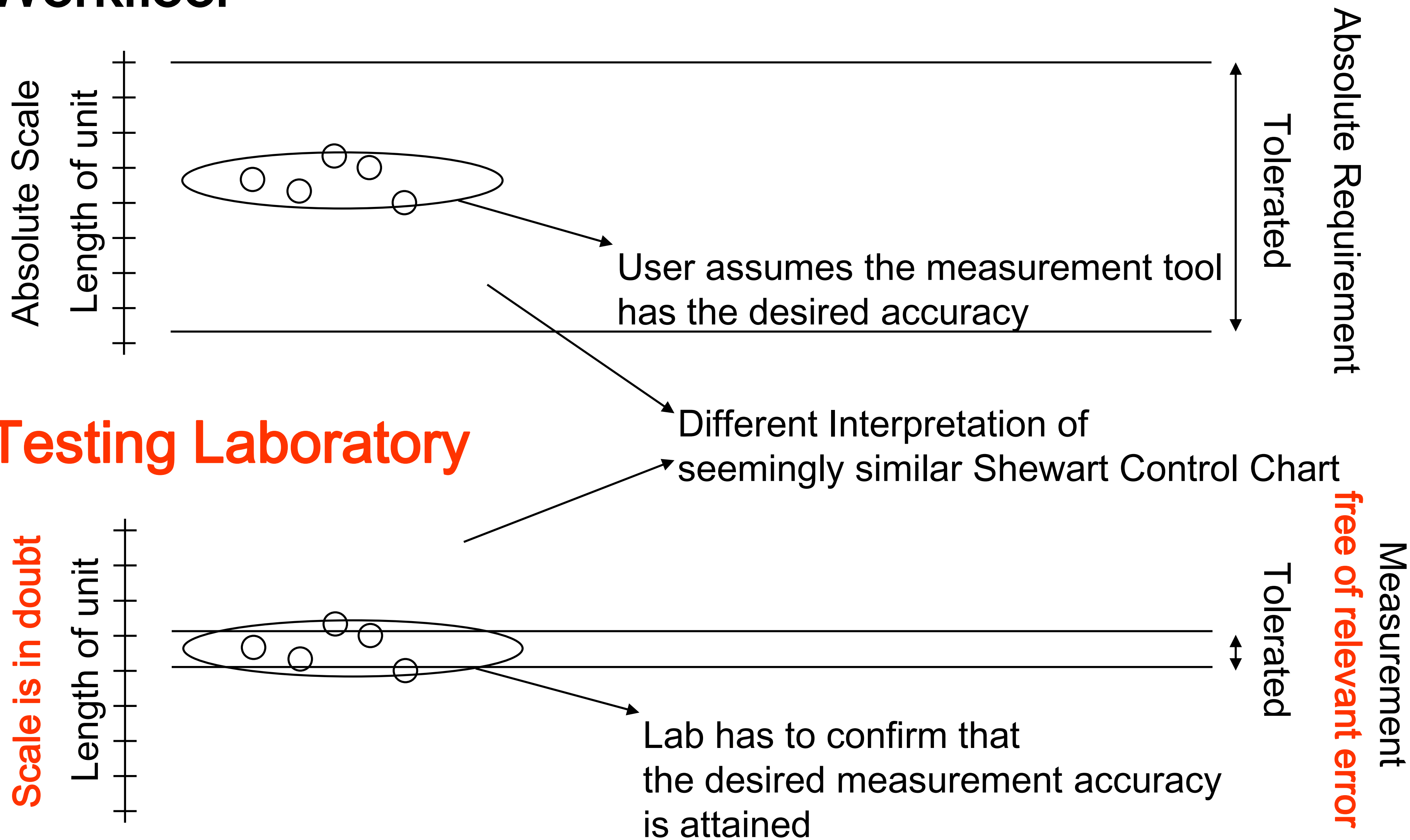
Physiology
1. measurement
2. relevant characteristics

Pathology
3. errors
4. medically relevant errors

Propedeutics
5. semeiology
6. diagnostic power
7. **sampling / evaluation / reporting**
   - materials & sampling frequency
   - epistemic estimates
   - bias detection
   - **reporting & interpretation**

Quality Control : e.g. Job Shop

Workfloor



Reporting & Interpretation (1/2)

The diagnostic scenario is equivalent to
continued monitoring of a patient

We report repeat single estimates for
a true (but unknown) quantity*

We use that single estimate (or a combination of tests)
to detect and interpret instabilities of the system

* quantity = single measurements
or single evaluations of one of the rules implemented

Reporting & Interpretation (2/2)

The last value is weighed at a prediction,
stemming from the repeated measurements in the past

We don’t know what we want ↔ We know what we want don’t want

☹ ☺

z-scores (x-X)/S and α-errors are the only format available for direct interpretation

In addition to z-scores likelihood ratio’s with respect to an agreed upon undesirable condition can be calculated

H_0-model

H_1-model

Expectations is the place you have to go to before you get to where you are going
Norman Juster

Interpretation

The lure of statistics
- objective
- amenable to automation

There are three kinds of lies:
lies, damned lies, and statistics
Benjamin Disraeli

Thermodynamics or **limited efficiency** of a statistical test :
Any statistical test will realize but a fraction (< 1)
of a maximum achievable power (sens * spec < 1)
statistically significant ≠ absolutely certain

The **futility** of statistics :
statistically significant ≠ relevant

How then do we convince ourselves ?
Visual inspection and pattern recognition
outstrips multiplex & complex rules

THE PROPER FRAME of iQC:

Not in-line post-manufacturing control
but a measurement tool part of your System of Primary Prevention

Setup-dominant	Machine-dominant	Operator-dominant	Component-dominant

Items

Setup-dominant	Machine-dominant	Operator-dominant	Component-dominant
Labeling Worklists Dispension …	Pipetting Analysis …	Adjusting Expert judgements …	Formulation Consumables …

Typical Control Procedures

Setup-dominant	Machine-dominant	Operator-dominant	Component-dominant
Precontrol First-piece inspection Attribute inspection …	Maintenance In-line periodic inspection iQC …	Acceptance inspection iQC Operator scoring …	Vendor Rating Incoming inspection Prior-operation control Acceptance inspection iQC …

After LA Seder

Proactive = Good Manufacturing Practice

Interpretation & Actions

Uncertainty of interpretation
- handling artefacts
- poor operational characteristics

Recipe
- confirmatory repeat test

Staging
- H_0: abnormal behavior of system
- H_1: predefined undesirable behavior
- H_1: unacceptable quality failure

Action
- ??? ambiguity ???
- analysis & prevention
- recall & correction

Cost-effectiveness

☺

Action
- analysis & prevention
- recall & correction

Tools
- log of reagents & events
- specific procedures

Summary

Take Home Message

Summary (1/6) Why is the doctor mad with his patient ?

Case : It is common practice in the area where Dr. X works to hand print-outs of lab results to patients. Mr. Y comes and asks what is wrong with his results for chloride and bicarbonate. The values are slightly outside of the normal range, and more significantly they are flagged on the report.

- Dr X explains that this doesn't mean anything
- He doesn't charge his patient for the results
- He stops requesting chloride and bicarbonate
- This is the fifth patient that day with similar questions: he gets mad.

Summary (2/6)
Why is our personnel dissatisfied with iQC ?

Case : Dr X is up-to date. He read about the latest test for the diagnosis of a not that uncommon ailment in his patients.
He starts requesting the test, and learns from his accumulated experience:
- most positive results, upon work-up, turn out to be negative
- positive cases go anyhow undetected

= Rock-bottom ROC-curves

- He persists
- He persists, but switches lab
- He gives up, but not without contacting the lab
- This is not his first similar experience, he gives up

Summary (3/6)

Nature of the error
- Transient
- Persistent
- Bias
- Random error

Issue : Limit yourself to what is doable
- poor operational characteristics
- best operational characteristics
- dependent on estimate of true value
- requires multiple measurements

Clinical Requirement
- Analytical sensitivity
- Dynamic Range
- Bias

Recipe : Limit yourself to what is relevant
- iQC low level
- iQC low & high levels
- AOP

Practical issue
- Poor Operational Characteristics
- Uncertainty of interpretation

Recipe
- adjust size / frequency of sample
- VISUAL INSPECTION

If you want the system to work empower your technicians & communicate

Summary (4/6)

Staging

H_0: abnormal behavior of system

H_1: predefined undesirable behavior

H_1: unacceptable quality failure

Action

???

analysis & prevention

recall & correction

Cost-effectiveness

☺

Tools

log of reagents & events

specific procedures

Cost-effective strategies depend on knowing what you want = relevance

Summary (5/6)

The concept of “ relevance ” :

The poorer the diagnostic ability of a test,
the less you have to bother about analytical performance.
Or, a poor diagnostic test cannot be improved
by being particular about analytical performance. ☺

The better the diagnostic ability of test,
the less analytical performance matters.
Or, a good diagnostic test,
is insensitive to analytical performance. ☺

Not to understand “ relevance ”
is to incur the costs of missed opportunities

Summary (6/6)

The concept of " control " :

PROCESS CONTROL = PROCESS CARE

NOT ☹
confirmation
of past performance

BUT ☺
actions to
ascertain future proper use

NOT ☹
to introduce fluctuations
by acting on "chance causes"

BUT ☺
to assign and cure
all avoidable fluctuation

NOT ☹
secondary prevention by
scrap, recall, rework

BUT ☺
primary prevention by
fail-proof design,
training & maintenance

From a conversation
between the Dodecahedron and young Milo
upon his arrival in the land of the Mathemagician:

“ But it is very accurate, and as long as the answer is right, who cares if the question is wrong? ”
“ If you want sense, you’ll have to make it yourself. ”

from The Phantom Tollbooth
by Norton Juster