Critical Test Appraisal

Validation of Analytical Performance

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

Operational Definition

- (Analytical Test) Validation :
  - Systematic evaluation and appraisal of (analytical test) validity
- (Analytical Test) Validity :
- (Analytical specs: range, precision profile, specificity, …)
  - Verification of methods & raw data !
  - Sound ?
    - Sound experimental design ?
    - Conclusions justified ?
  - Applicable ?
    - **fit for purpose = function of the intended use**

Proper starting point & end point of this think experiment

Operational Definition

Test Validation

= to assure that the test has all necessary characteristics for optimum/desirable outcome of care *

= Purposive optimization :
Logistics & Risk management

- cost-effective care
- compliance in the greatest possible group
- minimize adverse treatments

* The “ value ” of a laboratory test in the therapeutic “ value chain ” and the consequences of this analysis for test selection and validation is discussed in a separate teaching module : the diagnostic process

Teaching goals

- Understand requirements of standards
- Make sure that the exercise adds value
- Find relevant information
    - or what your IVD supplier should have done for you
- Life cycle of a method
    - validation plan and method validation
    - verification of implementation on the workfloor
    - continuous learning
- Efficient data collection and analysis
    - LAG toolbox explained

Analytical Test Evaluation
Requirements of Standards

Praktijkrichtlijn 6.1 i.v.m. KB *Art 22 §1 Art 23 §1*

6.1. Aanschaf van goederen
NORM
(Een) daartoe aangewezen medewerker(s) moet(en) er op toezien dat de ten behoeve van het onderzoek toegeleverde of in eigen beheer geproduceerde onderzoeksmiddelen alsmede andere voor het onderzoek benodigde goederen aan de daarvoor vastgelegde criteria voldoen. Deze criteria dienen te zijn vastgesteld onder rechtstreekse verantwoordelijkheid van de biologen.

Praktijkrichtlijn 10.7 i.v.m. KB Art *25-29*

10.7. Validatie

(klinische validatie ++)
Sommige hoofdstukken en onderafdelingen zijn gemerkt met een vermelding "(++)". Dit betekent dat de desbetreffende eisen verder gaan dan wat gevraagd wordt in de norm EN 45001.

…

NORM

Validatie van resultaten moet worden uitgevoerd op twee niveaus:
- een analytische validatie …
- een klinische validatie die uitsluitend behoort tot de bevoegdheid van de bioloog

…

Bij de klinische validatie:

… moet aandacht besteed worden aan een aantal algemene punten zoals:
- Oorsprong van de referentiewaarden

...

VEREISTEN

o een procedure voor het in dienst nemen van nieuwe apparatuur
o een procedure voor het valideren van nieuwe methoden

ISO-15189:2003

5.5 Examination procedures

5.5.1 The laboratory shall use examination procedures, … which meet the needs of the users … . Preferred procedures are those that have been published in established/authoritative textbooks, peer-reviewed texts or journals, or in international, national or regional guidelines.
If in-house procedures are used, they shall be appropriately validated for their intended use and fully documented.

5.5.2 The laboratory shall use only validated procedures … The methods and procedures selected for use shall be evaluated and found to give satisfactory results before being used for medical examinations.

ISO-15189:2003

5.5.3 … The procedure … based on the instructions for use (e.g. package insert) written by the manufacturer, provided that they are in accordance with 5.5.1 and 5.5.2 … Any deviation shall be reviewed and documented. … Each new version of examination kits with major changes in reagents or procedure shall be checked for performance and suitability for intended use. …

… documentation should include, when applicable, the following:

a) purpose of the examination;

…

c) performance specifications (e.g. linearity, precision, … uncertainty of measurement, detection limit, measuring interval, trueness of measurement, analytical sensitivity and analytical specificity);

d) primary sample system (e.g. plasma, serum, urine);

e) type of container and additives;

…

g) calibration procedures (metrological traceability);

…

j) interferences (e.g. lipaemia, haemolysis, bilirubinaemia) and cross reactions;

…

l) biological reference intervals;

m) reportable interval of examination results;

n) alert/critical values, where appropriate;

o) laboratory interpretation;

…

q) potential sources of variability.

ISO-15189:2003

5.5.4 Performance specifications for each procedure used in an examination shall relate to the intended use of that procedure.
5.5.5 Biological reference intervals shall be periodically reviewed. If the laboratory has reason to believe that a particular interval is no longer appropriate for the reference population, then an investigation shall be undertaken, followed, if necessary, by corrective action. …
5.5.6 The laboratory shall make its list of current examination procedures, including primary sample requirements and relevant performance specifications and requirements, available to users of laboratory services upon request.
5.5.7 If the laboratory intends to change an examination procedure such that results or their interpretations could be significantly different, the implications shall be explained to users of the laboratory services in writing, prior to the introduction of the change.
NOTE This requirement can be accomplished in any of several different ways, depending on local circumstances. Some methods include directed mailings, laboratory newsletters or part of the examination report itself.

ISO-15189:2003

5.6 Assuring quality of examination procedures

5.6.2 The laboratory shall determine the uncertainty of results, where relevant and possible. Uncertainty components which are of importance shall be taken into account. Sources that contribute to uncertainty may include sampling, sample preparation, sample portion selection, calibrators, reference materials, input quantities, equipment used, environmental conditions, condition of the sample and changes of operator.

Analytical Test Evaluation

Added value of the exercise

- Uncertainty of measurement
  - Model : cause-effect & cause-fail analysis
  - Uncertainty budget : identification of critical steps & of main components of error
  - Design for quality: improved robustness
- Metrological traceability & commutability
  - Commutability to method used in relevant clinical study
  - Commutability when lab changes method
  - Commutability between institutions (networking)
- Reportable range adjusted to intended use
  - Simplicity of operation, avoiding undesirable sample dilutions, etc.
- Proof that lab is up to the state of the art
  - Credibility vis-á-vis customers / external auditors

Analytical Test Evaluation

Sources of information

- Uncertainty of measurement
  - EQUAS-schemes : compounds all sources of error
  - Supplier : budget & identification of main components of error
- Metrological traceability & commutability
  - EQUAS-schemes : comparison of methods
  - Supplier : Traceability of calibrators
- Reportable range adjusted to intended use
  - Supplier : Method insert sheets
- Proof that lab is up to state of the art
  - Laboratory internal quality control data
  - Participation in EQUAS-schemes

Analytical Test Evaluation

Sources of information

If your IVD supplier
cannot produce a credible UM budget,
then they did not optimize robustness of the test.

If Your IVD supplier / EQUAS scheme
do not deliver valuable information,
then you better think of selecting
a different method or a different supplier.

You better come to this conclusion early on
in the selection procedure,
before proceeding to any validation.

Analytical Test Evaluation

The validation plan

- Cycle : plan – execute – test for specs – conclude
  - objective, systematic, credible
- Specs will follow from the intended clinical use
  - cost
  - turn-around-time
  - analytical specs
- "Learning experience" : From the continuously changing tension between past experience and future expectations spring new ideas
  - You may have to amend your plan during execution
- The conclusions have to relate to the intended use of the test and derived relevant specs
  - Failure to meet the specs does not automatically have to result in rejection : allow for a work around

Analytical Test Evaluation

The validation plan

- Create value :
  - Step 1 : Weed out irrelevant elements from your plan
  - Step 2 : Produce Value (= what your stakeholders want)
    - Credibility
      Validation file (= show case) must convince
      your clinicians (and possibly external auditors)
      and therefore must "relate to what is relevant" :
      your clinician doesn't expect you to reinvent the wheel
      but wants to be assured of "good and reliable service"
    - Proactive optimization
      Analytical robustness
      Timely reporting of results
      Informative value of test results

Expectations is the place you have to go to,

1. What is the Clinical Scenario in which I use the test ?

- dichotomic classification ?
- screening or confirmation ?
- staging and/or follow-up ?
- your case-mix ?
- position in relation to complete work-up ?

2. What is the validity of published studies on diagnostic accuracy of the test ?

Applicable to your scenario ?

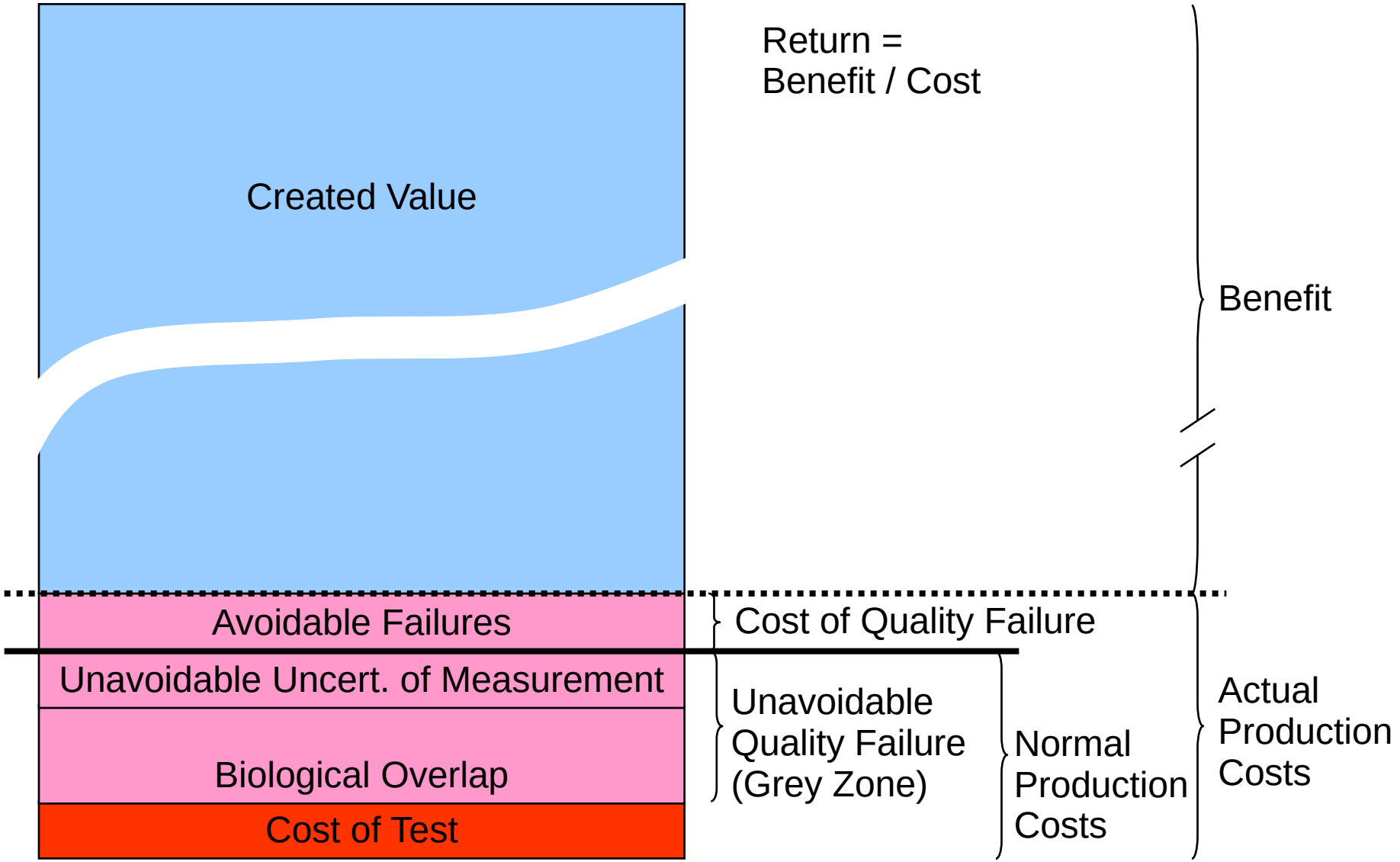
- prevalence in clinical setting often higher than in general population
  and thus the benefit (= reduction of uncertainty) of the test will be smaller
- you need to distinguish the index-condition from a clinical case mix
  and not from matched cases drawn from a healthy control group

before you get to where you are going.

3. Does analytical performance affect clinical utilization ?

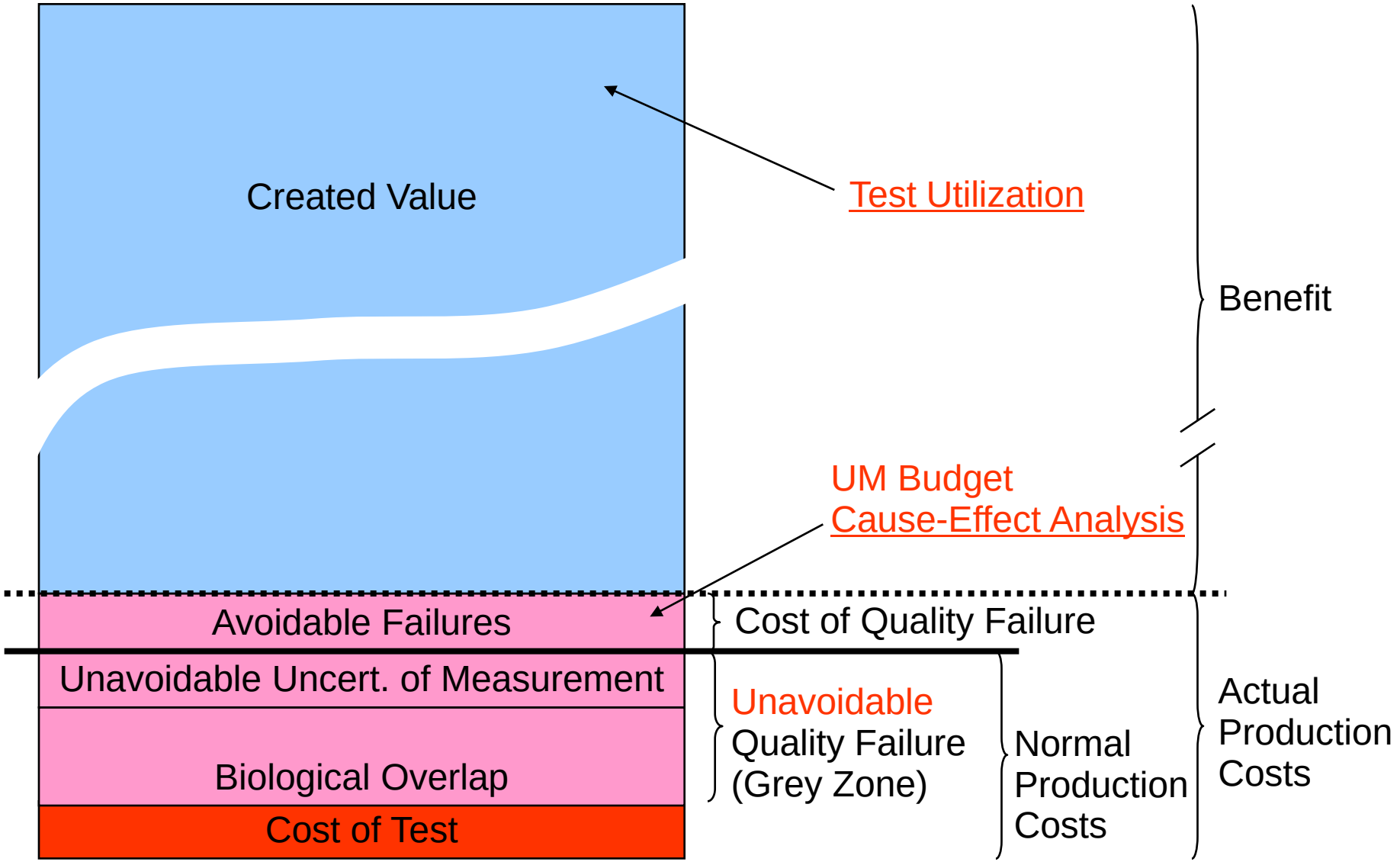
4. Is the required clinical performance realized ?

3. Does analytical performance affect clinical utilization ?



Does analytical performance affect clinical utilization ?

Where can the validation exercise create value ?



Does analytical performance affect clinical utilization ?
Where can the exercise create value ?

Don’t lose perspective !

- **Test better be valuable**
  If so, then optimize test utilization.
  If not, then abandon the test.
- **If grey-zone phenomenon eats away benefit,**
  then consider selecting a different test,
  but don’t wail about the unavoidable.
- **No observation is possible without uncertainty of measurement**
  The loss of benefit due to this uncertainty is usually negligible.
  If this is not the case, then concentrate on selecting a different test.
- **Evaluation of uncertainty of measurement & buildup of experience**
  should result in measures improving robustness of the “diagnostic procedure”
  = from sampling, over analysis, to interpretation and treatment.

4. Is the required analytical performance realized ?

Are data of IVD supplier / literature valid ?

- Is the analytical method described in full detail ?
- Is the methodology for test validation documented ?

  Details of the protocol, participating sites, …

  Is the protocol analytically and statistically sound ?
- Can the raw data be audited ?

  Substitute: for some tests (HIV, blood banking)
  the method has to be validated by an independent authoritive organism

Are the available data relevant ?

Has the diagnostic performance been evaluated for the intended use in your clinical setting ?

  Identical differential diagnostic problem / scenario ?

  Spectrum and prevalence of study groups ?

Are the analytical specifications suited for the intended use in your clinical setting ?

Analytical Test Evaluation
Toolbox

- Comments on the concept of Uncertainty of Measurement
- Comments on the concept of Trueness
- Interpretation of standard / Expectations of BELAC
- Tools used in
  Laboratory Medicine University Hospitals KULeuven
- Literature

Uncertainty of Measurement *

= What is the reasonable dispersion of results ?

1. Uncertainty linked to measurement on the bench

- Probabilistic Model :
  - construct model with all steps in the analytical process
  - estimate / assign error for each step (see teaching unit on GUM)
  - chain of events starts with calibration:
    - uncertainty of the assigned value
    - uncertainty of the calibration procedure
  - analyze the error-propagation : identify critical steps
    principal components
  - MU = estimate of total dispersion · coverage factor

- If the test has any diagnostic power,
  - then this type of uncertainty is usually irrelevant (see slides 18-19)

* The topic is discussed in greater detail in a separate overview 

Uncertainty of Measurement

2. Measurement = Part of a Diagnostic Procedure

- Laboratory diagnosis often relies on a "surrogate-test"
  for the detection of a(n) (surrogate) index-condition
- "Measurand" = "what should appear on the report" =
  measurand(-like) substance/mass (or reactivity) in matrix x,
  sampled in particular physiological conditions (cfr. proper ref. ranges)

3. Uncertainty linked to spurious interferences

- Frequentist model / Risk analysis
  Extent, circumstances and frequency of these interferences can be known :
  - Analytical non-specificity
  - Biological / physiological responses
- When not recognized, this can/will result in misdiagnosis and mistreatment

4. Info has to be provided by IVD supplier

- The lab cannot evaluate uncertainty stemming from calibrators and reagents
  - requires the evaluation of the effect of many batch changes
  - requires logging of data in a format fit for ANOVA
  - currently, both of these conditions are not regularly met by the industry
- EQUAS can middle out simultaneously the effect of many variables
  - but is compounded by handling of the test material at multiple sites

Trueness

= Relationship to the true value

1. Traceability to Certified Reference Materials (CRM)
   - responsibility of the IVD supplier

However, our business is not the truth, but is the best treatment !

2. Trueness refers to adequacy of decision limits
   - refers to the sources of these limits & the tests used in establishing them
   - but, don't lose your perspective :
     the main determinant of "diagnostic accuracy" can be
     "metrological traceability / commutability"

but certainly also is :
- prevalence
- spectrum or case-mix

and
- effects of sampling and sample handling
- biological / physiological interferences

- These factors are bound to a particular clinical setting / path

and therefore shall be guarded by the laboratory physician !

RISK Analysis

Our business is minimizing avoidable harm

Manageable Risk refers to
- Impact of errors : medical & quality of care delivery
- Frequency of important errors
- Timeliness of detection of these errors
  - how long will error persists before being detected ?
  - how much time do I have to prevent / revert harm ?

Risk Analysis

- Starts from the clinical scenario
- Delivers
  - Operational & Analytical quality targets
  - Appropriate & cost-effective Control Procedures
    = **PRO-ACTIVE** CONTROL

Tools: the standard as a checklist

ISO-15189 5.5.3

a) purpose of the examination;
…
c) performance specifications (e.g. linearity, precision, … uncertainty of measurement, detection limit, measuring interval, trueness of measurement, analytical sensitivity and analytical specificity);
d) primary sample system (e.g. plasma, serum, urine);
e) type of container and additives;
...
g) calibration procedures (metrological traceability);
…
j) interferences (e.g. lipaemia, haemolysis, bilirubinaemia) and cross reactions;
…
l) biological reference intervals;
m) reportable interval of examination results;
n) alert/critical values, where appropriate;
o) laboratory interpretation;
…
q) potential sources of variability.

Tools: the standard as a recipe book

BELAC 2-102

Algemeen

…
gewerkt met commerciële kits en systemen.
De fundamentele validatiecriteria (lineariteit, kruisreacties, interferenties, detectielimiet,…) moeten door de producent worden aangetoond .
(see also: slide 22 point 4)
…
dient het laboratorium alleen de validatie van de implementatie uit te voeren in haar eigen omgeving.
(see also: slide 23 point 2)

Tools: the standard as a recipe book

BELAC 2-102

Validatie van de implementatie

Precisiegegevens:

intra- en interrun CV's

De intrarun CV's zijn een maat voor de meetonzekerheid …

Voor de intrarun CV

zijn minimum 10 bepalingen nodig.

De interrun CV's

kunnen afgeleid worden van gegevens van de interne QC.

Tools: the standard as a recipe book

BELAC 2-102

Validatie van de implementatie

Juistheid

t.o.v de targetwaarde voor de gebruikte methode.

Ze kan afgeleid worden uit :

- gegevens van de externe QC
- testen op controlematerialen met targetwaarde voor deze methode
- testen op gecertificeerde referentiematerialen
- methodevergelijking met andere methodes …

een voldoend aantal monsters

over het gehele klinisch relevante gebied

Tools: the standard as a recipe book

BELAC 2-102

Validatie van de implementatie

Verificatie van referentiewaarden

Het is aan elk laboratorium om voor elke test
de meest geschikte referentiewaarden op het protocol te vermelden.
De oorsprong ervan moet wel bekend zijn.
… gepubliceerde referentiewaarden
(uit tekstboeken, uit kitbijsluiters, publicaties).
Het laboratorium moet een verificatie uitvoeren om na te gaan
of de gebruikte referentiewaarden gelden voor zijn situatie.

Tools: Laboratory Medicine - UZ – KULeuven *

1. Standard Operating Procedure (Dutch)
   - PRC-132 - Structuur van het validatiedossier
   - available upon request : laboratoriumgeneeskunde@uz.kuleuven.be

2. Checklists used for method selection & communication with the suppliers
   - CHK-201 - Checklijst: Voorstel voor invoeren van een nieuwe test
   - CHK-202 - Checklijst: Bevraging van toestel- en test-specificaties

3. Spreadsheets used for evaluation of selected test(s)
   - PRG-125 - Methode validatie: Precisie gegevens
   - PRG-126 - Methode validatie: Methode vergelijking
   - PRG-127 - Methode validatie: Verificatie van referentie ranges

4. Checklist used to assure correct implementation on the workfloor
   - CHK-204 - Checklijst: Implementatie van nieuwe of gemodifiëerde test

* List corresponds to status of the LAG quality system on nov 8 2005
for extra-muros readers: hard-copy versions can be requested
at laboratoriumgeneeskunde@uz.kuleuven.be

Tools: Laboratory Medicine - UZ - KULeuven

Optimize efficiency of data collection and calculus

- PRG-125 - Methode validatie: Precisie gegevens
  - Daily duplo measurements : ANOVA for within day and between day CV
- PRG-126 - Methode validatie: Methode vergelijking
  - Data from low – normal – high, with 20 / group
  - Deming orthogonal analysis with estimation of imprecision
    - (cfr. FDA/CLIA-waiver, on estimate of UM)
  - Project reference-ranges of old and new method on the regression line
    - for validation / adjustment of such ranges
- PRG-127 - Methode validatie: Verificatie van referentie ranges
  - Automatic evaluation of percentiles and outliers

Tools: Laboratory Medicine - UZ - KULeuven

Lab documentary system 

- SOP (Standard Operating Procedure) :

  How do we do it ? ⟶ *Intended Public: Lab Technicians*

- Logbook :

  What happens on the workfloor ? ⟶ *Intended Public: Lab Techn. & Staf*

  Log significant events &
  actions & conclusions

- Validation file :

  Why do we do it this way ? ⟶ *Intended Public: Requesting Physicians*

  Method Validation

  How good are we at it ? ⟶ *Intended Public: Staf & Customers*
  External Auditors

  Implementation validation (CHK-204)
  Continuous learning : Significant conclusions form
  Internal Quality Control
  EQUAS

Tools: Laboratory Medicine - UZ - KULeuven

Exploit the contract with your IVD-supplier :

- Validation encompasses
  - *- installation, operational validation of technical settings*
  - *- implementation validation*

- At installation and acceptance testing time :
  - *- This is normally done by the supplier*
  - *- Supervise it to verify the raw data & soundness of procedure & conclusions*
  - *- That way you avoid to repeat this at your own expense without extra cost to the supplier*

- At installation and training time :
  - *- Part of the training is usually done by the supplier*
  - *- Use the training to simultaneously evaluate whether specs are met & whether personnel is competent*
  - *- That way you maximize the return of your & their efforts*

Special case ? Microbiological Testing

European Cooperation for Accreditation of Laboratories EAL-G18 guidance document (1996)

8.1. Each laboratory will have particular requirements for the performance … in order to demonstrate suitability for the intended purpose. … Essential features of any method … the 'correct' answer with respect to specified limits of detection, selectivity, repeatability and reproducibility.

8.2. For official methods, or methods from recognized … standard organisations, … introduce it by a documented training programme. Basic parameters like variation … can be found in … publications, books and manuals for media.

8.3. Commercialised test systems (kits) may not require further validation if validation data from alternative sources, *eg* based on collobarative testing, is available. Laboratory should seek from manufaturers validation data …evidence. …

In summary: *nil novi sub sole*

- specify requirements
- demonstrate that requirements are met
- introduce the method, with attention to training and meeting the requirements
- make use of inter-personnel evaluations, EQUAS, …
- demonstrate your attention by a formal conclusion

Take home messages

Test Validation IS NOT an
EXERCISE in POLITCAL CORRECTNESS

FIND THE RELEVANT INFORMATION

Test Validation IS KNOWING WHAT YOU WANT
based on a CLINICAL SCENARIO

Test Validation IS VERIFYING that your IVD-SUPPLIER
did what is needed to SET UP a ROBUST & USEFUL TEST

AFTER MAKING AN INFORMED SELECTION

Test validation IS NOT TEST RESEARCH
you should not have to repeat what others should have done (be)for(e) you

Test Validation IS DEMONSTRATING your ABILITY
to perform the test properly

Test validation IS IMPLEMENTATION suited for the INTENDED USE
a health-economic analysis may help you to focus on what counts (slide 17)

Literature

CLSI (formerly NCCLS) :
EP10 - Quantitative clinical laboratory methods (1998)
EP12 - Qualitative test performance (2001)

Eurachem :
The fitness for purpose of analytical methods (1998)

EAL :
G18 – Accreditation for laboratories performing microbiological testing (1996)

CUMITECH report
Verification and validation of procedures in the clinical microbiology laboratory
American Society for Microbiology (1997)

Vassault A, Grafmeyer D, de Graeve J *et al.*
Analyses de biologie médicale : spécifications et normes d'acceptabilité á l'usage de la validation de techniques
Ann. Biol. Clin. 57:685-95 (1999)

White GH and Fraser CG
The evaluation of a kit for clinical chemistry: a practical guide for the evaluation of methods, instruments and reagents kits
J. Automatic. Chem. 6:122-48 (1984)