EQALM symposium

Berlin 1 July 2009

Quality requirements and quality goals

A review of the Stockholm consensus on analytical quality specifications

Per Hyltoft Petersen

Per.Petersen@isf.uib.no

NOKLUS, Bergen, Norway

Background

Competing approaches to goal-setting in clinical biochemistry

Analytical quality specifications based on

- clinical and biological use of measurements
- expert groups
- EQAS and PT
- state of the art

Quality Specifications ? History ? Imprecision

1963 *David Tonks*

ALE = 2CV = [1/4 reference range/mean] x 100% {biological}

1968 *Roy Barnett*

"Medically significant CV" - opinions of clinicians and laboratory specialists {clinical}

1970 *Cotlove, Harris and Williams*

Biological variation - tolerable analytic variability
CV < 1/2 $CV_{within\text{-}subject}$ {biological}

1976 *CAP Aspen Conference* (1977) {biological}

1978 *Wiveka Elion-Gerritzen*

"Medically significant CV" - opinions of clinicians {clinical}

Quality Specifications ? History ? Bias

1988 *Elizabeth Gowans*
Specifications for acceptable bias {biological}

1980s Analysis of clinical situations [Nordic countries] {clinical}

1991 Sverre Sandberg
"Medically significant CV" - opinions of patients {clinical}

1997 *Callum G. Fraser*
Levels of quality {biological}

1990s *EGE-Lab Working Group*
Biological variation and state of the art {biological}

European EQA Organisers Working Group {biological}

ISO TC 212/WG3 ISO 15196
"Analytical Performance Goals Based on Medical Needs" {clinical}

The Cotlove goal for imprecision

Analytical quality specifications for coefficient of variation:

$$CV_{Analytical} < 0.5 * CV_{Within\text{-}Subject}$$

Cotlove *et al. Clin Chem* 1970;16:1028-32

Influence of imprecision on monitoring CV

Influence of analytical variation, $CV_{Analytical}$, on the total variation, $CV_{Total\text{-}Monitoring}$, during monitoring

$$CV_{Total\text{-}Monitoring}^2 = CV_{Within\text{-}Subject}^2 + CV_{Analytical}^2 =$$

$$CV_{Within\text{-}Subject}^2 + (0.5 * CV_{Within\text{-}Subject})^2 =$$

$$1.25 * CV_{Within\text{-}Subject}^2 \Rightarrow 1.12 * CV_{Within\text{-}Subject}$$

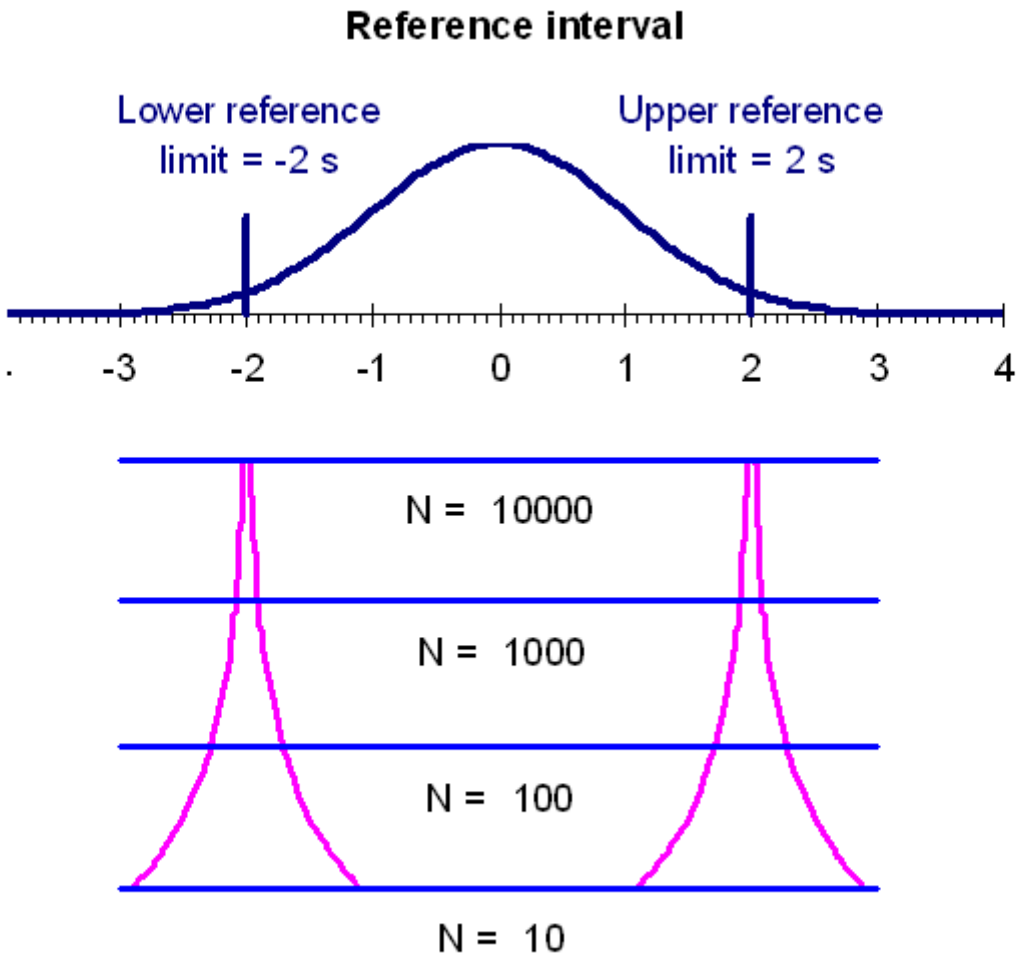
The Gowans goal for analytical bias

Specifications for Bias:

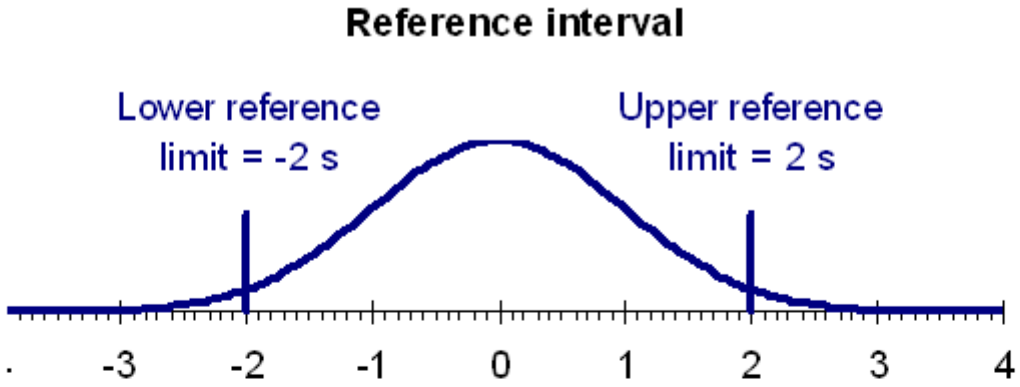
$$|\text{Bias}| < 0.25 * CV_{\text{Population}}$$

Gowans *et al. Scand J Clin Lab Invest* 1988;48:757-64

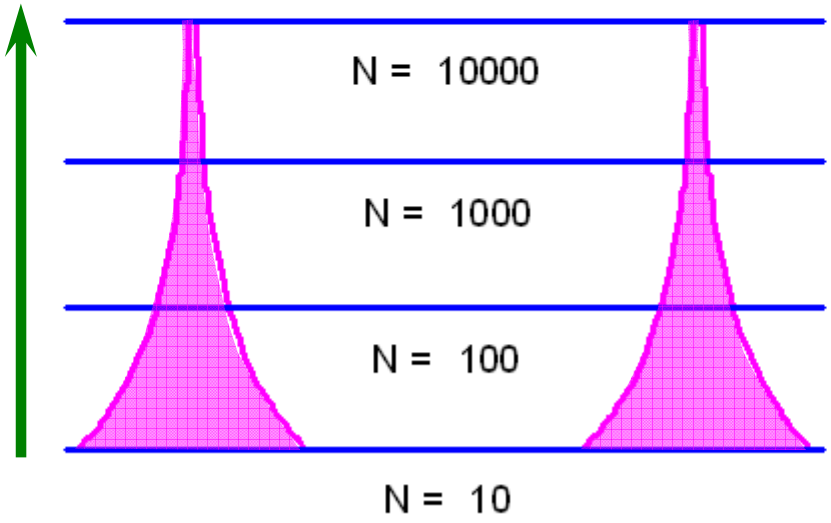
Reference interval with limits



Confidence intervals for reference limits



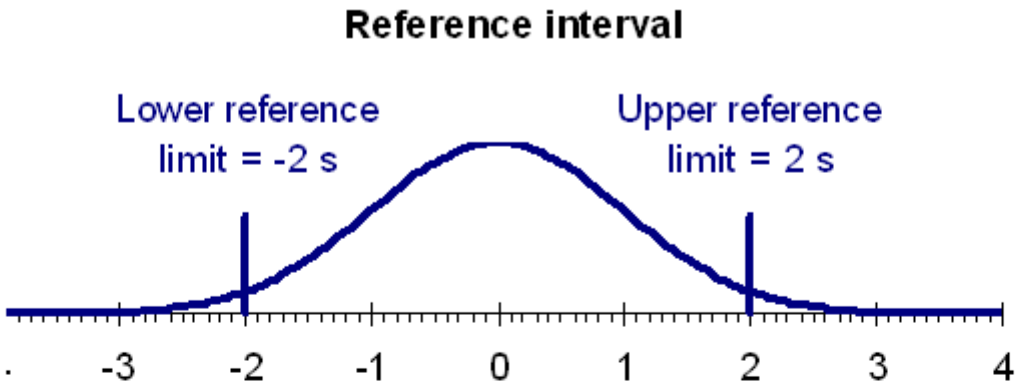
Decreasing CI for increasing N



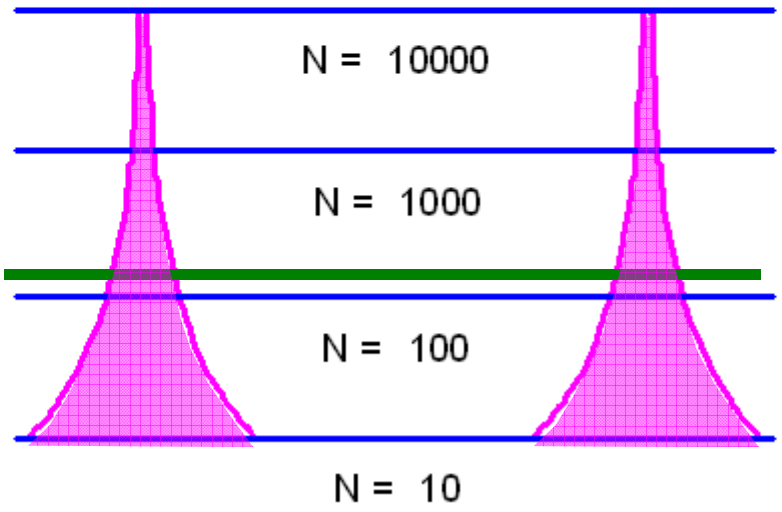
90 % CI 90 % CI

IFCC recommendations

IFCC recommendations



Sample size, N, should be > 120



N = 120

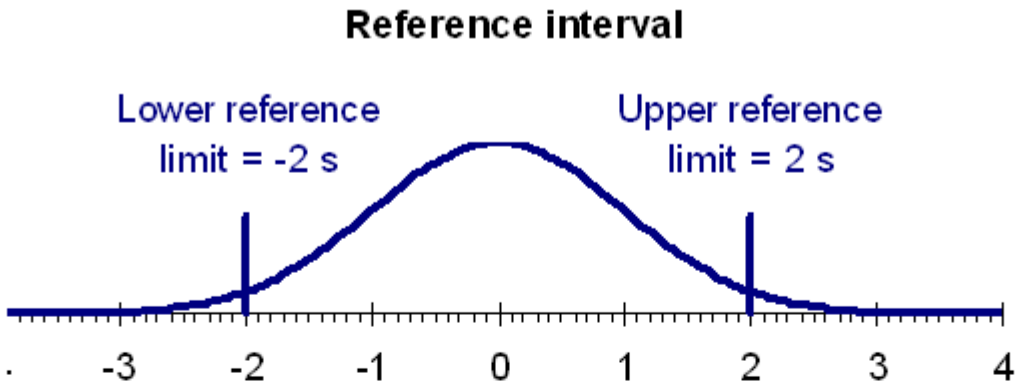
90 % CI

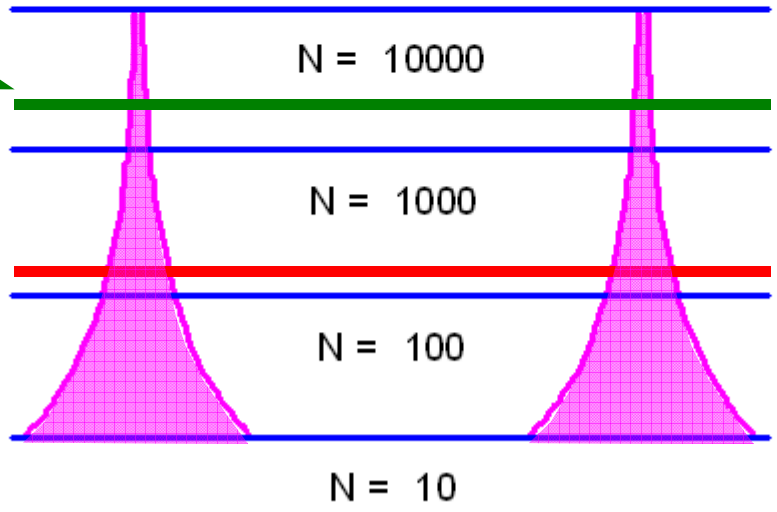
90 % CI

The Gowans goal for analytical bias

Sample size
N > 3000

Acceptable bias





N = 120

90 % CI 90 % CI

Gowans *et al.*
Scand J Clin Lab Invest
1988;48:757-64

The Gowans goal for analytical bias

With the common reference interval produced without errors and based on more than 3000

we can allow for bias and imprecision instead of sample size

Thus the common reference interval is as good for all, as if each lab had produced it according to IFCC

and with this analytical quality we can have same reference intervals for homogeneous groups in same regions

"Medically significant CV"

S-Potassium

Action limits for S-Potassium among clinicians in 12 hospitals

Each clinician gets a questionnaire with the same information about a patient and indicates for which concentration he/she will react

Reference intervals

Median of low action limits and high action limits

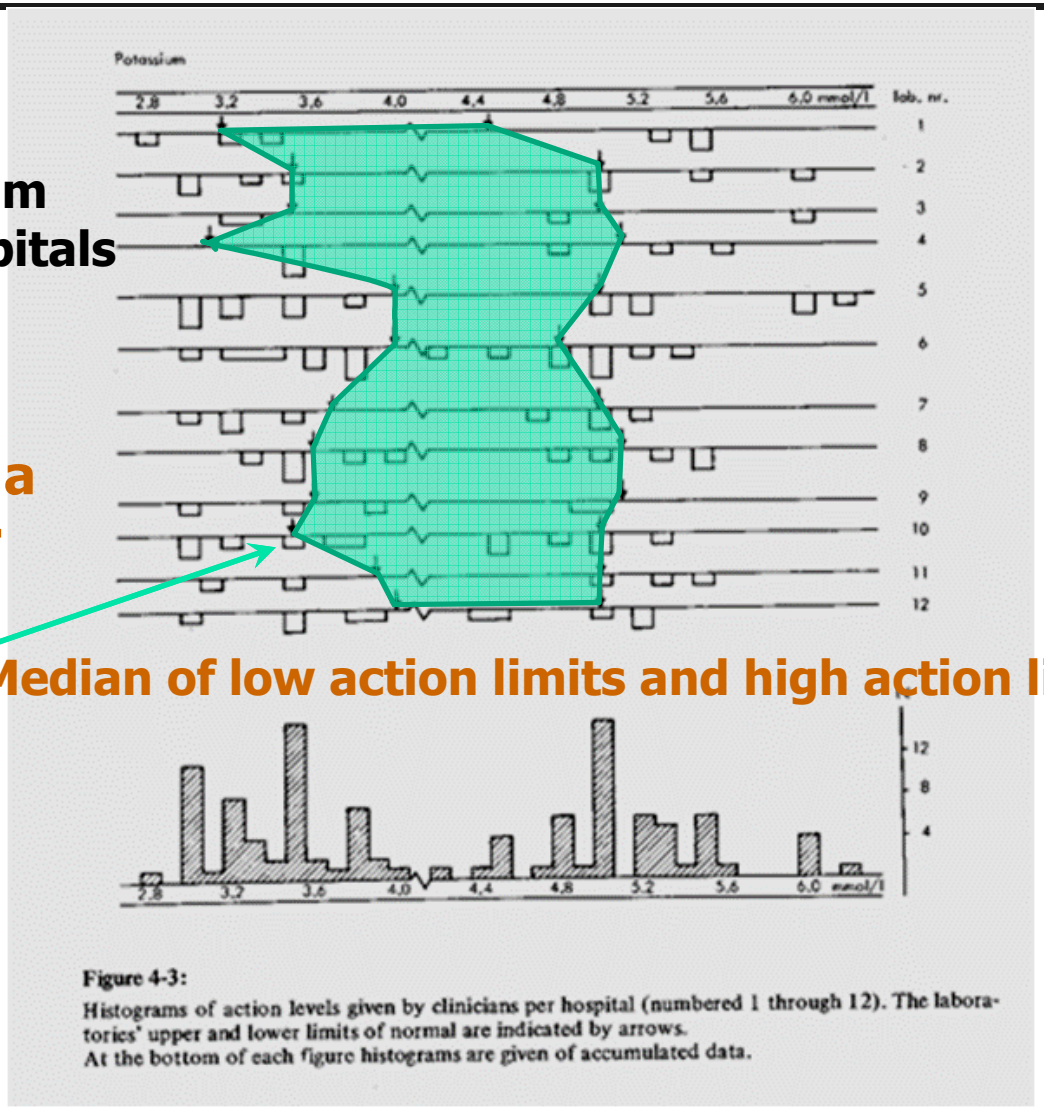
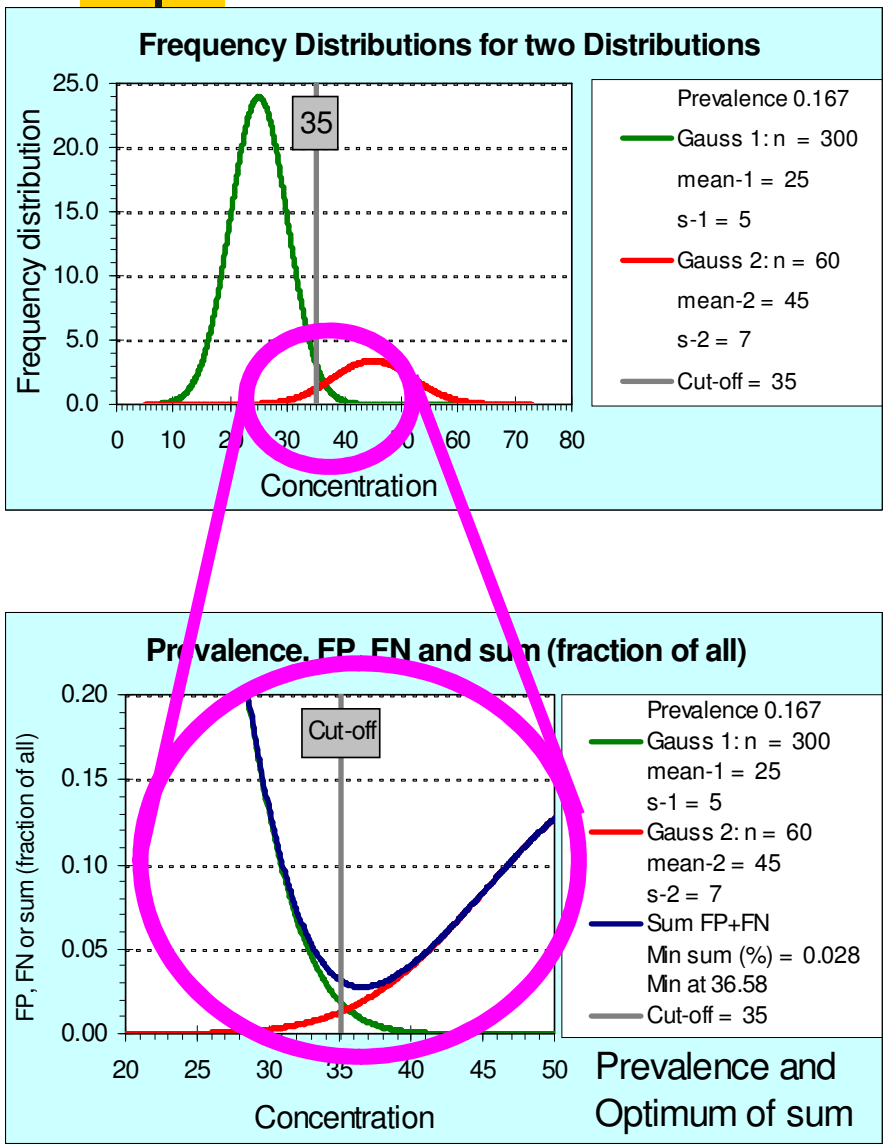


Figure 4-3:
Histograms of action levels given by clinicians per hospital (numbered 1 through 12). The laboratories' upper and lower limits of normal are indicated by arrows.
At the bottom of each figure histograms are given of accumulated data.

Elion-Gerritzen W, Thesis, 1978, Drukkerij J.H. Pasmans, S-Gravenhage

Clinical approach - Classification



Cut-off: 35 U/L

Two groups:
Healthy and diseased

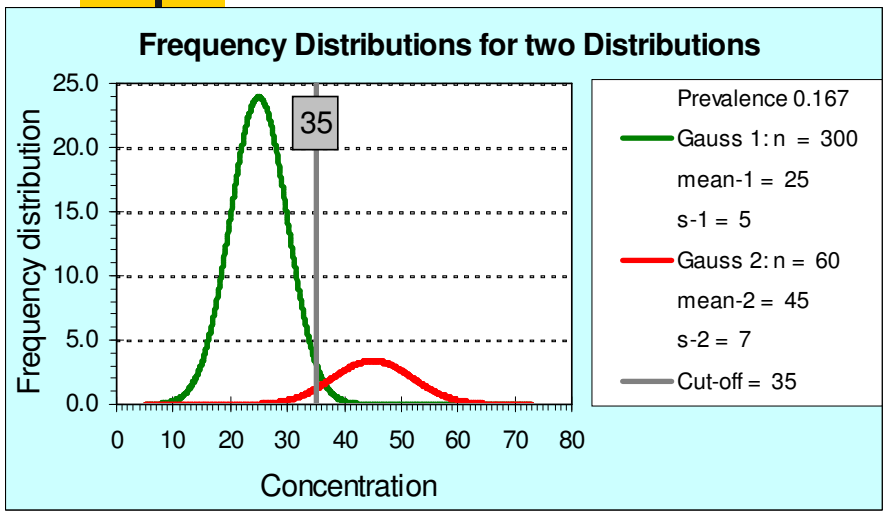
Prevalence 16.7 %

Red curve: FN as function of cut-off

Green curve: FP as function of cut-off

Blue curve: sum of FP and FN

Clinical approach - Classification

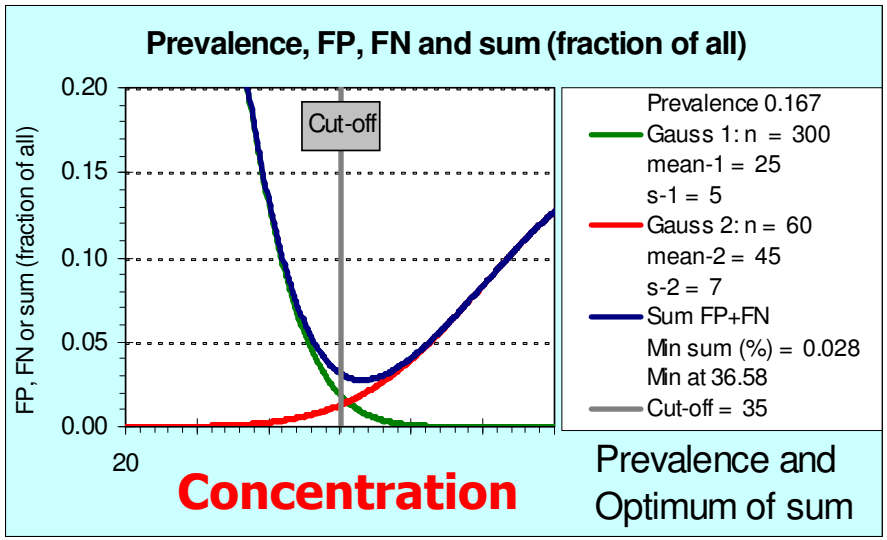


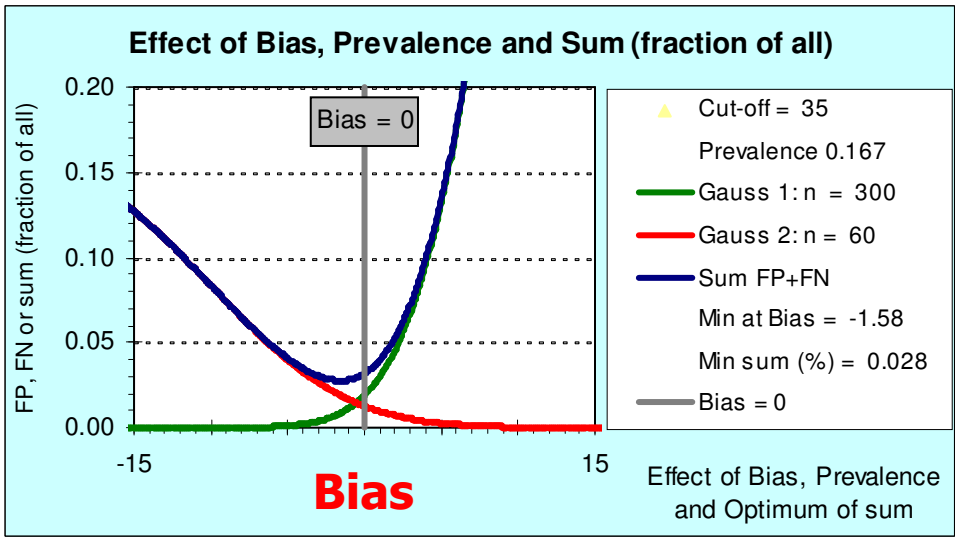
Effect of bias when Cut-off: 35 U/L

Red curve: FN as function of bias

Green curve: FP as function of bias

Blue curve: sum of FP and FN

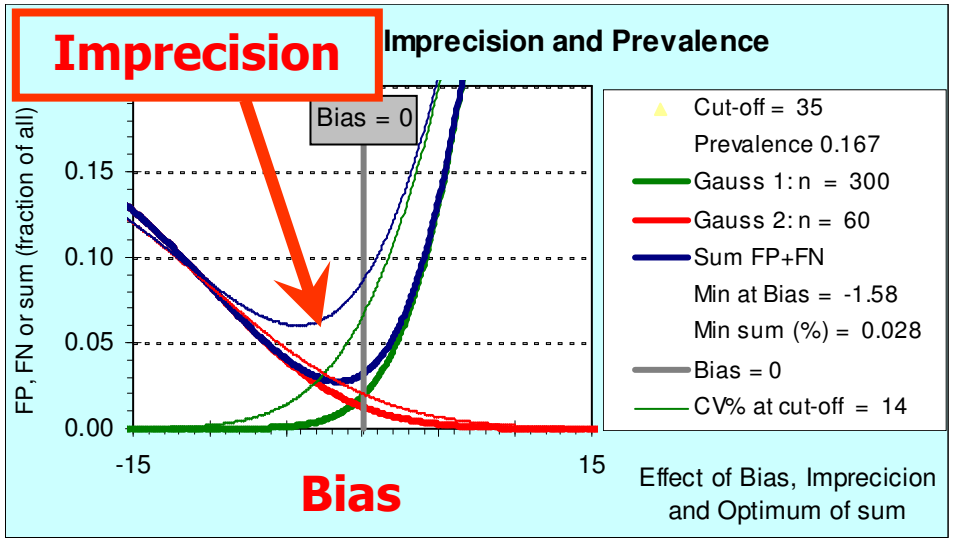


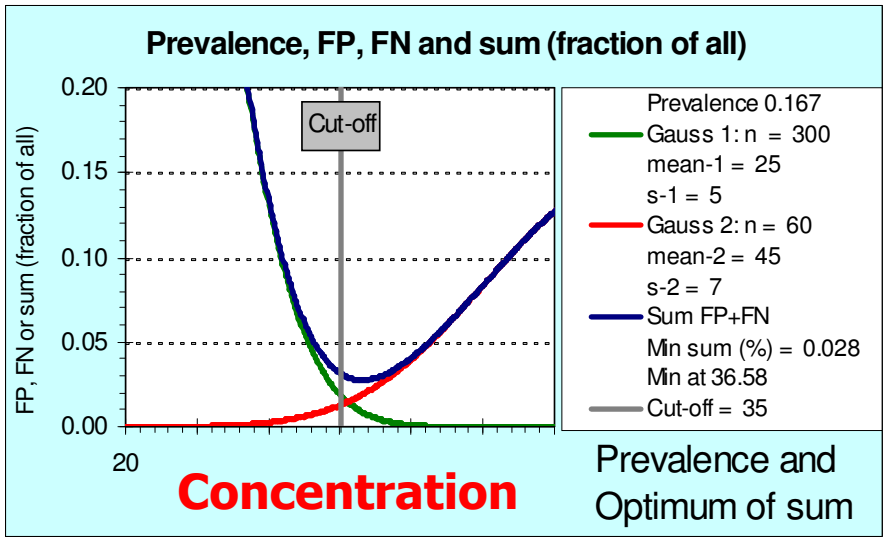


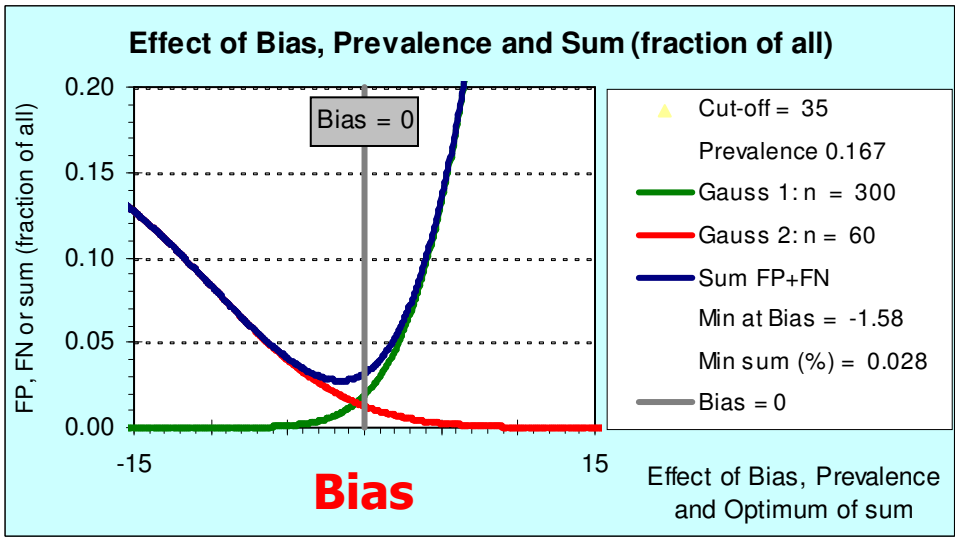
Clinical approach - Classification

Effect of bias and imprecision when Cut-off: 35 U/L

Can be advanced by use of weight factors for FP and FN



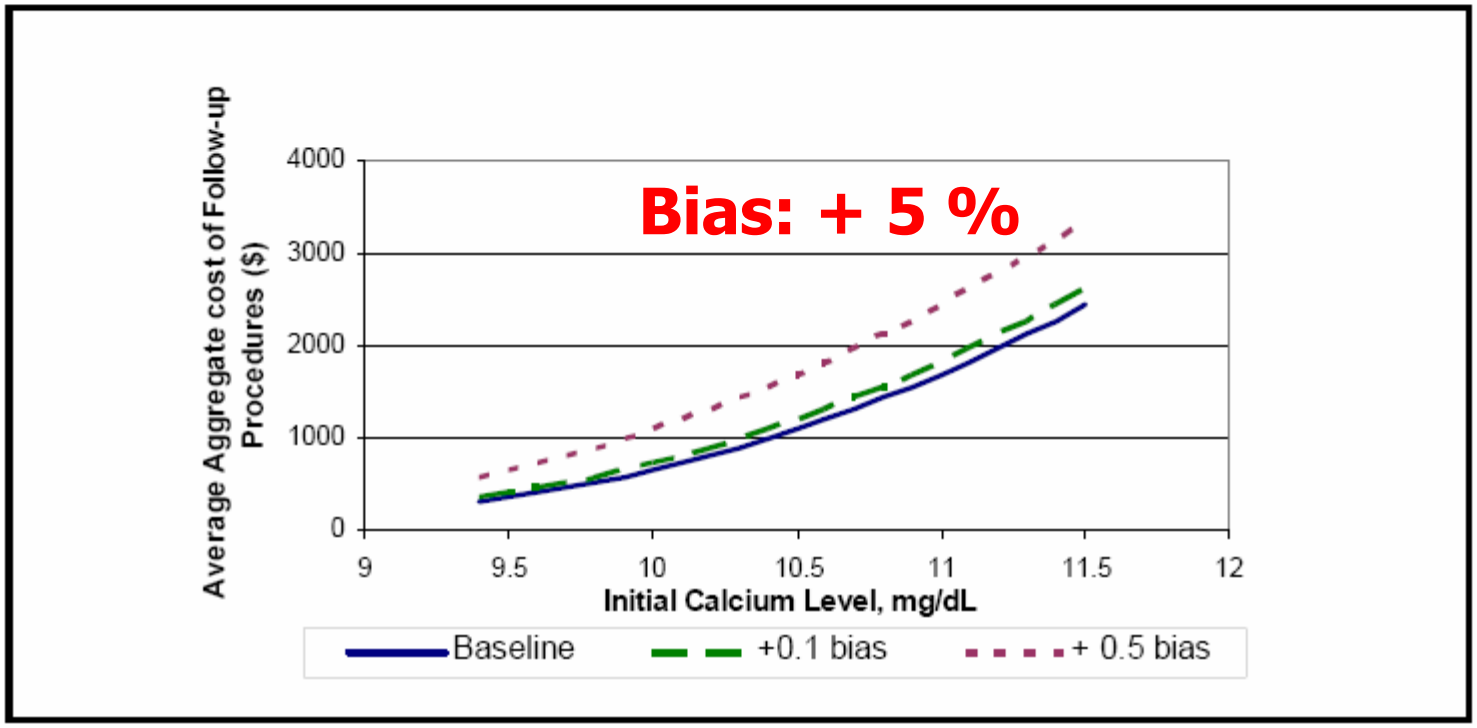




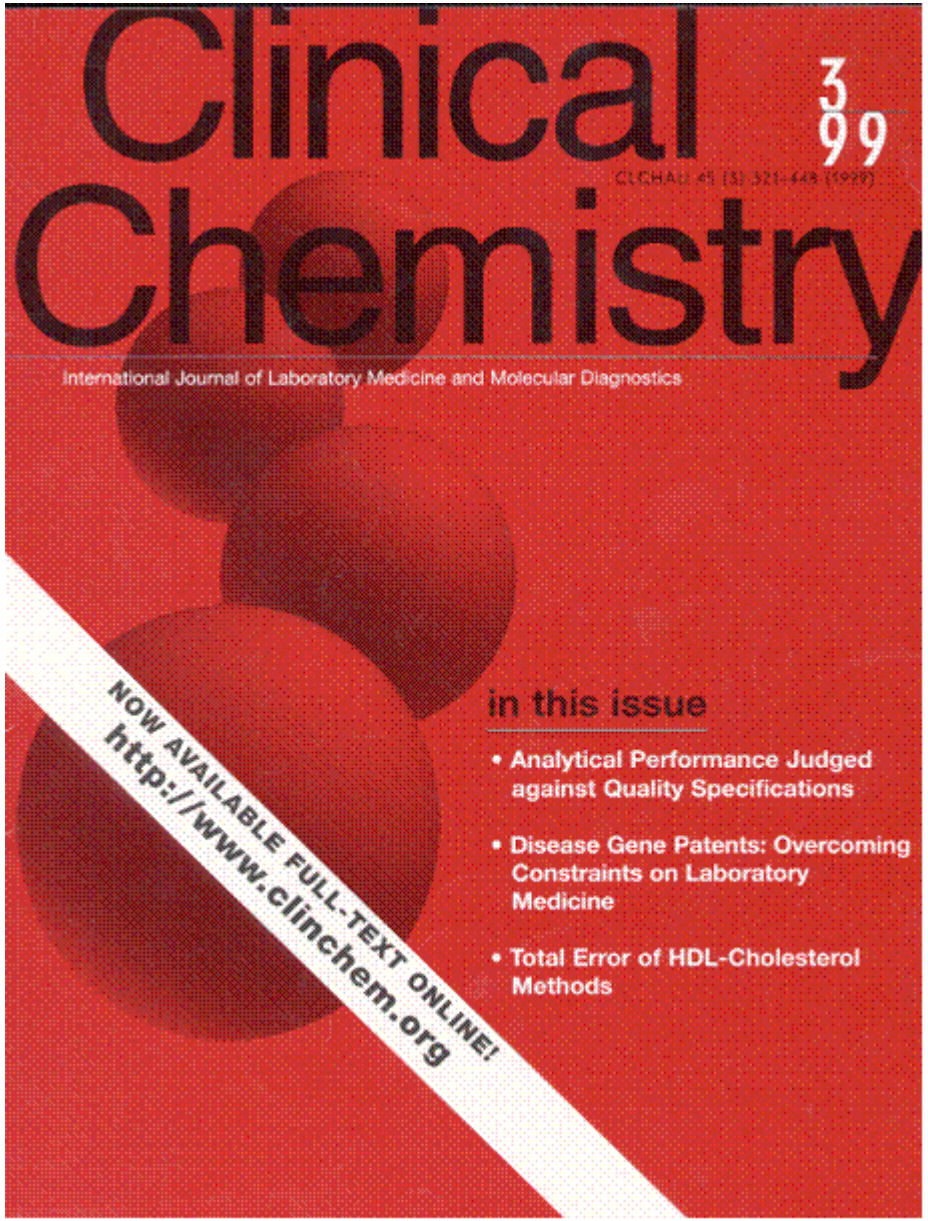
Economical approach - Costs

Cost of follow-up ($) as function of measured Ca-concentration

Figure 5-1. Shift in the Cost Function due to Analytic Bias
Private insurance patients



Gallaher MP, Mobley LR, Klee GG, Schryver P. The impact of calibration error in medical decision error. NIST 2004



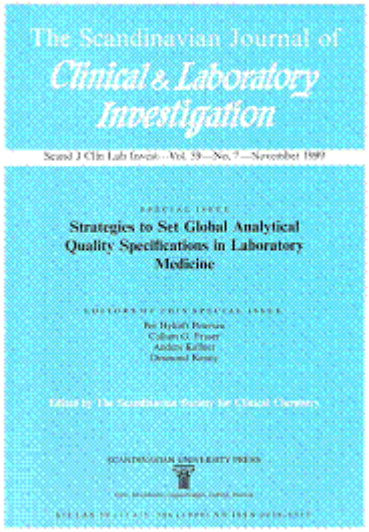
Editorial:

Fraser CG, Hyltoft Petersen P.

Analytical performance characteristics should be judged against objective quality specifications.

***Clin Chem* 1999;45:321-3**

Stockholm - Consensus Conference



Organiser: Anders Kallner

CONSENSUS STATEMENT*

The main outcome of the Conference was agreement that the following hierarchy of models should be applied to set analytical quality specifications.

1. Evaluation of the effect of analytical performance on clinical outcomes in specific clinical settings
2. Evaluation of the effect of analytical performance on clinical decisions in general:
   a. Data based on components of biological variation
   b. Data based on analysis of clinicians' opinions
3. Published professional recommendations
   a. From national and international expert bodies
   b. From expert local groups or individuals
4. Performance goals set by
   a. Regulatory bodies
   b. Organizers of External Quality Assessment (EQA) schemes
5. Goals based on the current state of the art
   a. As demonstrated by data from EQA or Proficiency Testing scheme
   b. As found in current publications on methodology

Scand J Clin Lab Invest 1999; 59: 585

Consensus agreement

D. KENNY,* C. G. FRASER,† P. HYLTOFT PETERSEN,‡ & A. KALLNER§

*Department of Clinical Biochemistry, Our Lady's Hospital for Sick Children, Dublin, Ireland; †Directorate of Biochemical Medicine, Ninewells Hospital and Medical School, Dundee, Scotland; ‡Department of Clinical Chemistry, Odense University Hospital, Odense, Denmark; and §Department of Clinical Chemistry, Karolinska Hospital, Stockholm, Sweden

The Editors of this special issue of the *Scandinavian Journal of Clinical and Laboratory Investigation* and the Organising Committee of the Conference *Strategies to set Global Quality Specifications in Laboratory Medicine*, Stockholm, 24–26 April 1999, are pleased to report that this recent Conference was most successful. Over 100 participants from 27 countries actively contributed to the discussions on the 22 formal presentations. Our primary aim in organizing the Conference was to provide a vehicle for reaching consensus on the setting of global quality specifications in laboratory medicine. This objective was achieved and lively constructive debate after the presentations were complete led to agreement on the principles laid down in the following Consensus Statement.

CONSENSUS STATEMENT*

The main outcome of the Conference was agreement that the following hierarchy of models should be applied to set analytical quality specifications.

1. Evaluation of the effect of analytical performance on clinical outcomes in specific clinical settings
2. Evaluation of the effect of analytical performance on clinical decisions in general:
   a. Data based on components of biological variation
   b. Data based on analysis of clinicians opinions
3. Published professional recommendations
   a. From national and international expert bodies
   b. From expert local groups or individuals
4. Performance goals set by
   a. Regulatory bodies
   b. Organizers of External Quality Assessment (EQA) schemes
5. Goals based on the current state of the art
   a. As demonstrated by data from EQA or Proficiency Testing scheme
   b. As found in current publications on methodology

Where available, and when appropriate for the intended purpose, models higher in the hierarchy are to be preferred to those at lower levels. The concept of such a hierarchy is described in a recent Editorial in Clinical Chemistry in which the relative merits of the above models are discussed (Clin Chem 1999; 45: 321–3). This hierarchy has also been proposed by the ISO/TC 212/WG 3 subgroup on "Analytical Performance Goals Based on Medical Needs" as the basis for the ongoing revision of ISO/CD 15196. The following matters were also discussed and agreed.

- The above hierarchy includes currently available models; however, new useful concepts will undoubtedly evolve. Implementation of any of the models should use well-defined and described procedures.
- To facilitate the future debate on the setting of analytical quality specifications, there is a need for agreement on concepts, definitions and terms.
- There is a need for continuous improvement in the exchange of information on quality issues between clinical laboratory professionals and the diagnostics industry, and between clinical laboratory professionals and the users of the laboratory service.

IFCC, IUPAC and WHO kindly sponsored the Conference but it must be noted that the Consensus Statement reflects the views of the presenters and registrants who participated in the Conference and does not necessarily represent those of the sponsoring bodies.

Kenny et al. SJCLI 1999; 59:585

What was achieved in Stockholm?

The consensus from the conference describes a hierarchical structure of approaches to estimation of analytical quality specifications

What was not achieved in Stockholm?

The agreement from the Stockholm conference was not followed directly by ISO/TC 212 - ISO 15 196, because the chairman, Larry Kapland, completely changed his mind after the conference

So we gave up and didn't support the ISO-group further

However, there is now a publication, ISO 15 196, with recommendations very close to the consensus from the Stockholm conference, but we do not know who are the authors – and there is no reference to the Stockholm consensus or to the editorial which was also very close to the consensus

What was not achieved in Stockholm?

There was no discussion about matrix-effects and consequently no specifications for allowable matrix

Control of INR-kit

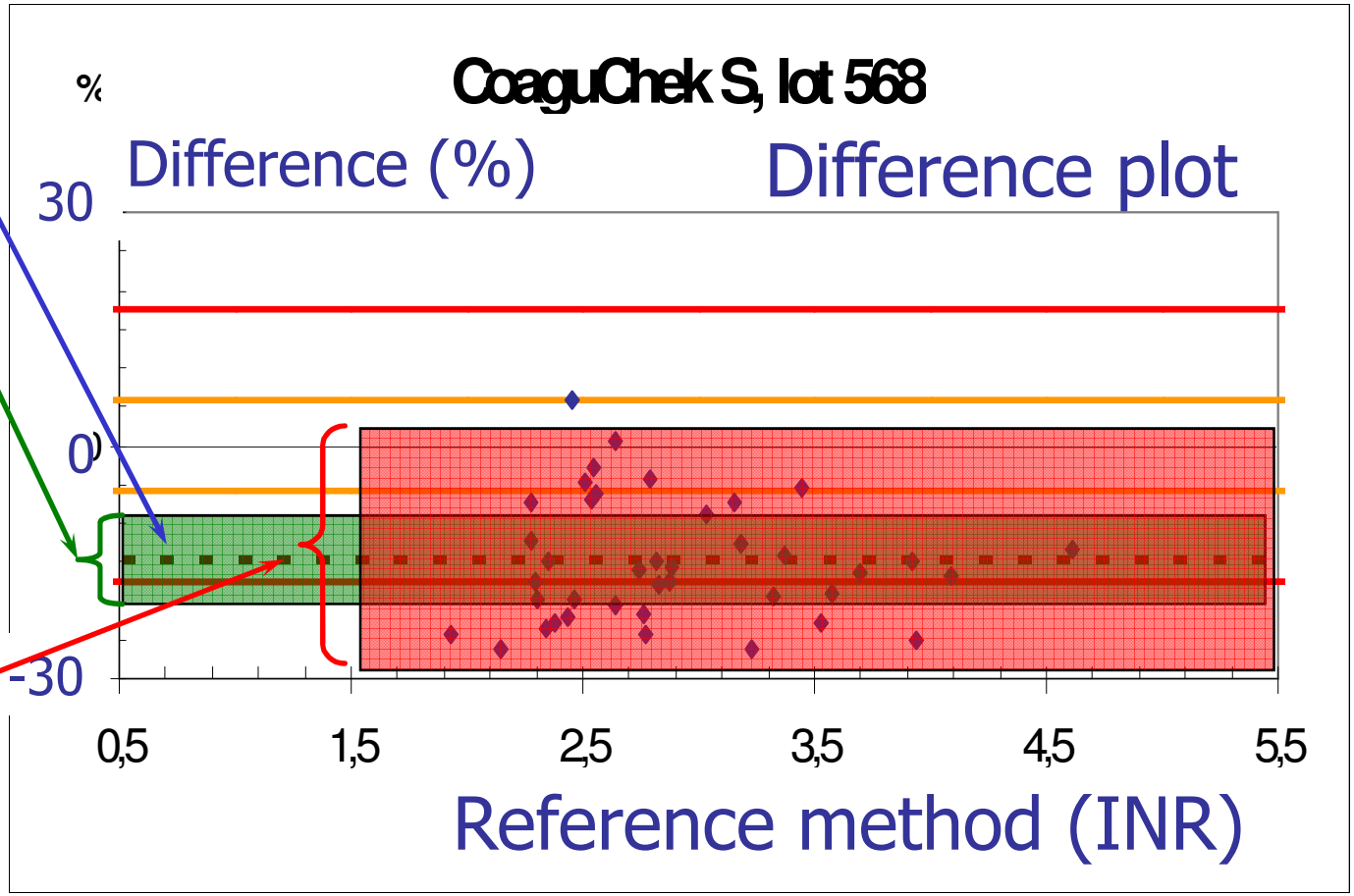
Bias = – 15 %

CV_A = 3.1 %

$CV_{Total}^2 = CV_{Matrix}^2 + CV_A^2$

CV_{Total} = 8.5 %

CV_{Matrix} = 7.9 %



Borrowed from Esther A. Jensen, Denmark

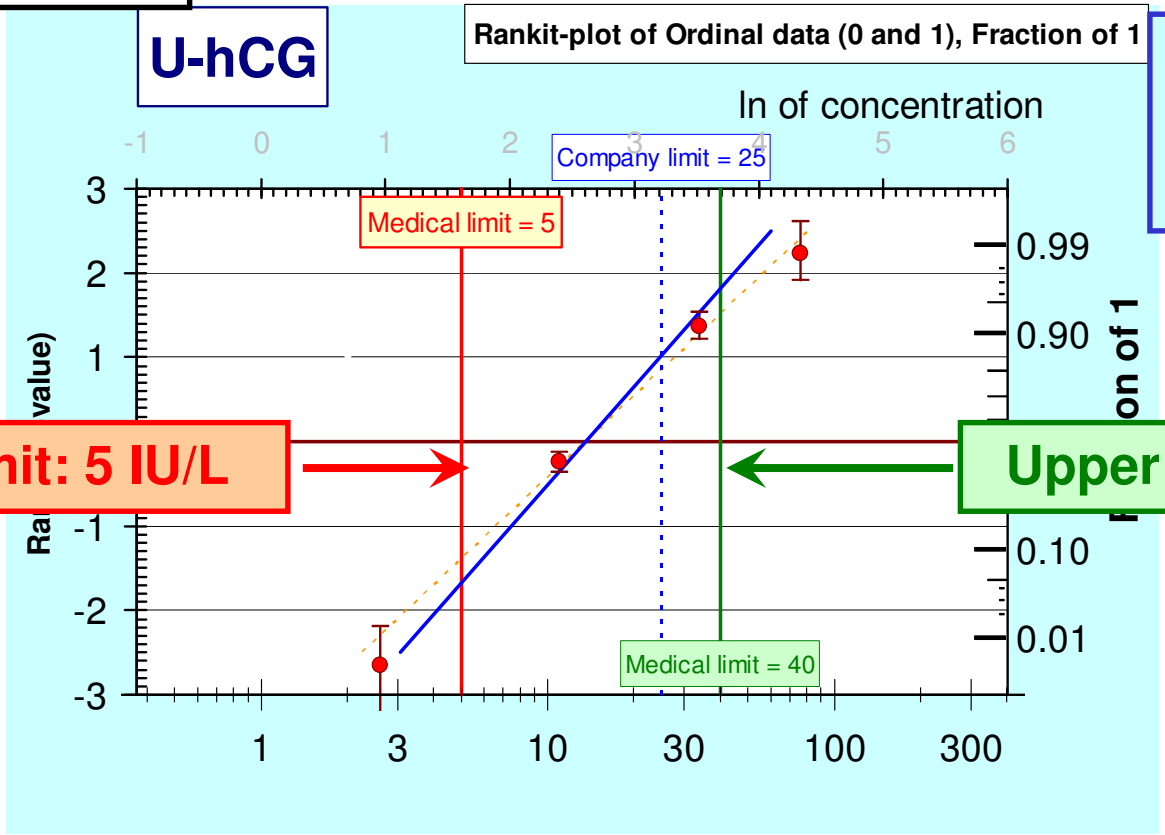
What was not achieved in Stockholm?

There was no discussion about matrix-effects and consequently no specifications for allowable matrix

There was no discussion about measurements on ordinal scale

Ordinal scale – dichotomous test

Urine-hCG



Fraction of positive

Lower limit: 5 IU/L

Upper limit: 40 IU/L

Concentration (IU/L)

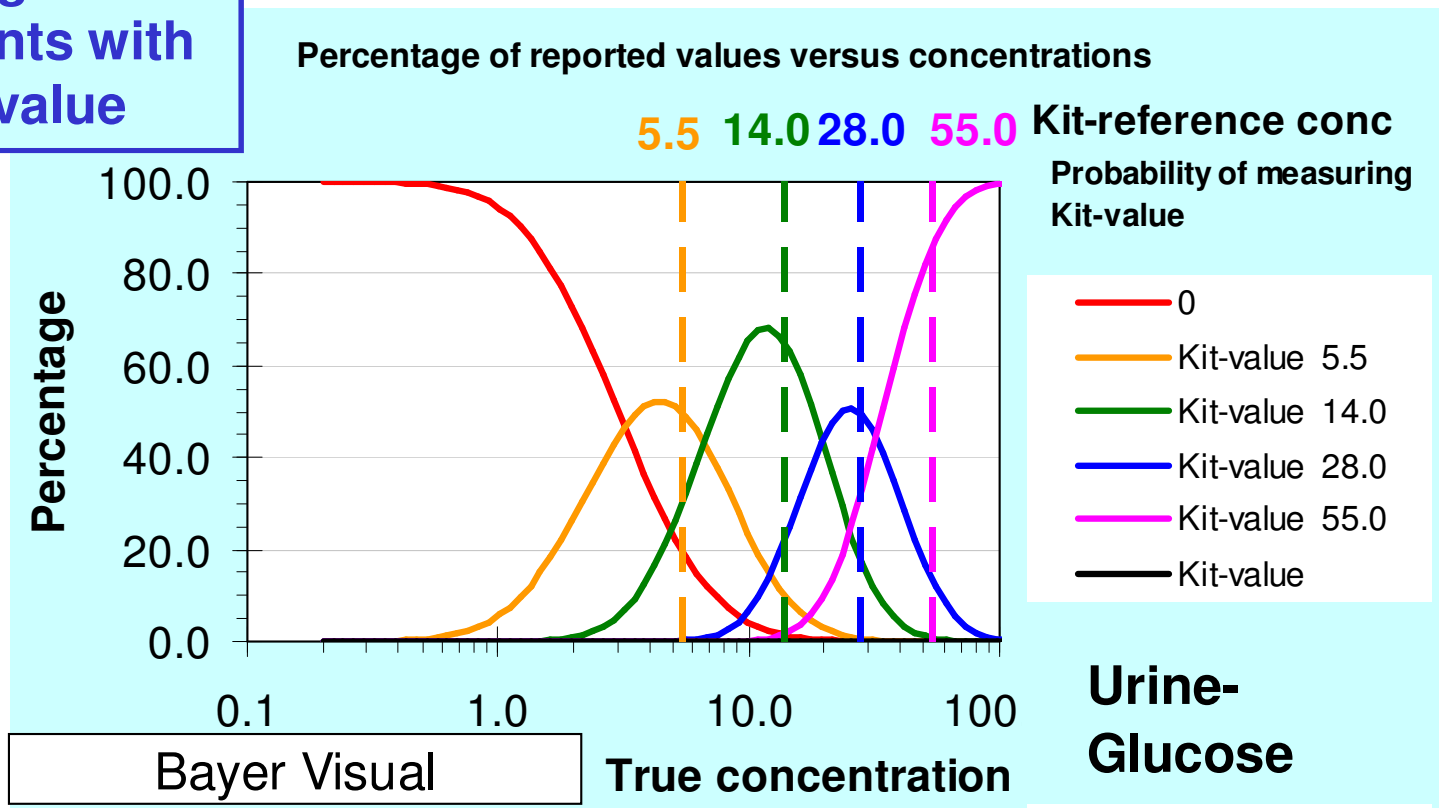
Hyltoft Petersen et al. Scand J Clin Lab Invest 2008;68:298-311

Ordinal scale – semi-quantitative test

Urine-glucose

Percentage of measurements with each Kit-value

Kit-values
0 mmol/L
5.5 mmol/L
14 mmol/L
28 mmol/L
55 mmol/L



Bayer Visual

Concentration (mmol/L)

Hyltoft Petersen et al.
Scand J Clin Lab Invest
2009; in press

What was not achieved in Stockholm?

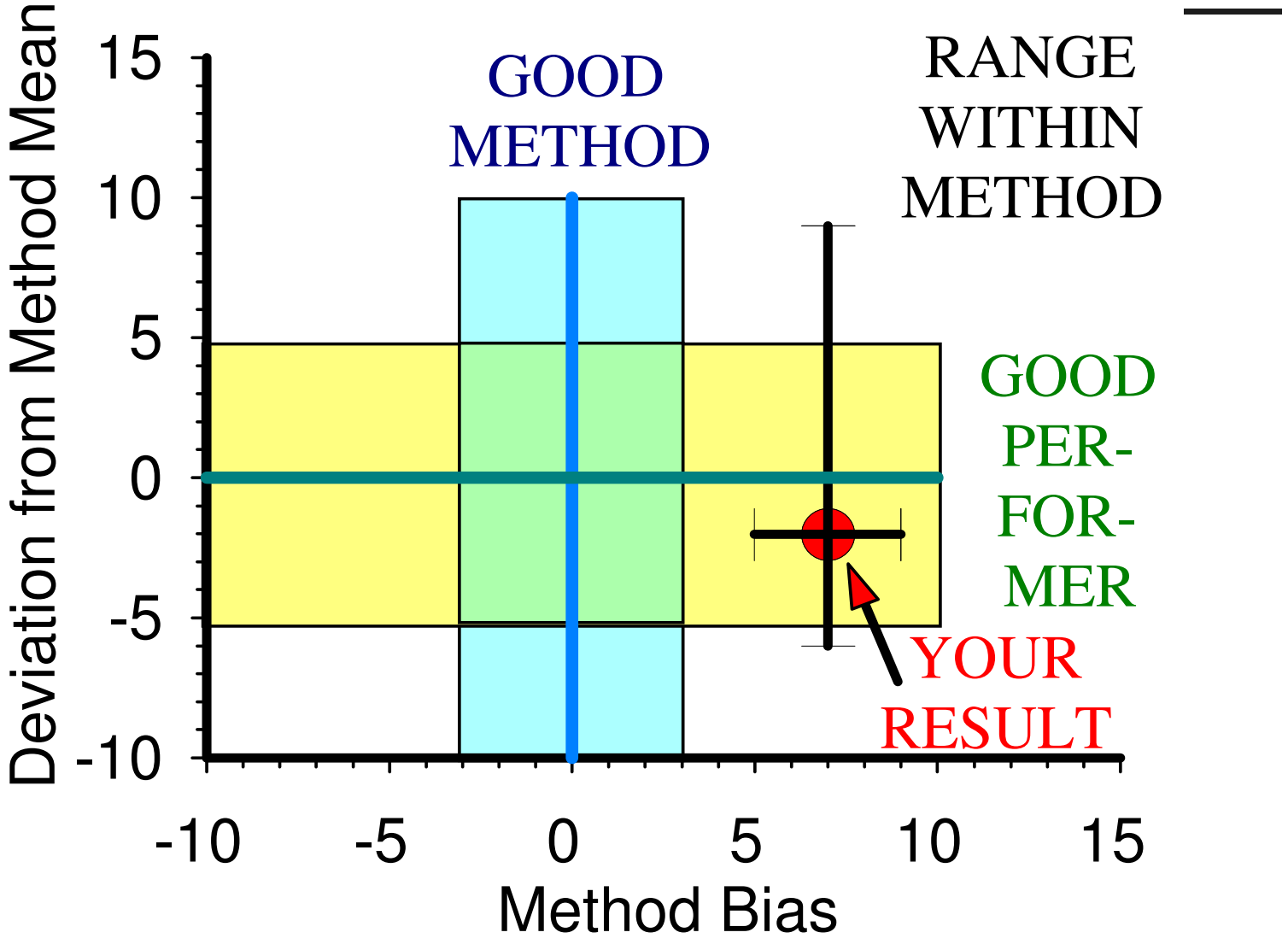
There was no discussion about matrix-effects and consequently no specifications for allowable matrix

There was no discussion about measurements on ordinal scale

There was no conclusion about absolute and relative quality:

- **Deviation from a 'true' value**
- **Deviation from the method mean**

Validation of Methods and Validation of Participants



What was not achieved in Stockholm?

There was no discussion about matrix-effects and consequently no specifications for allowable matrix

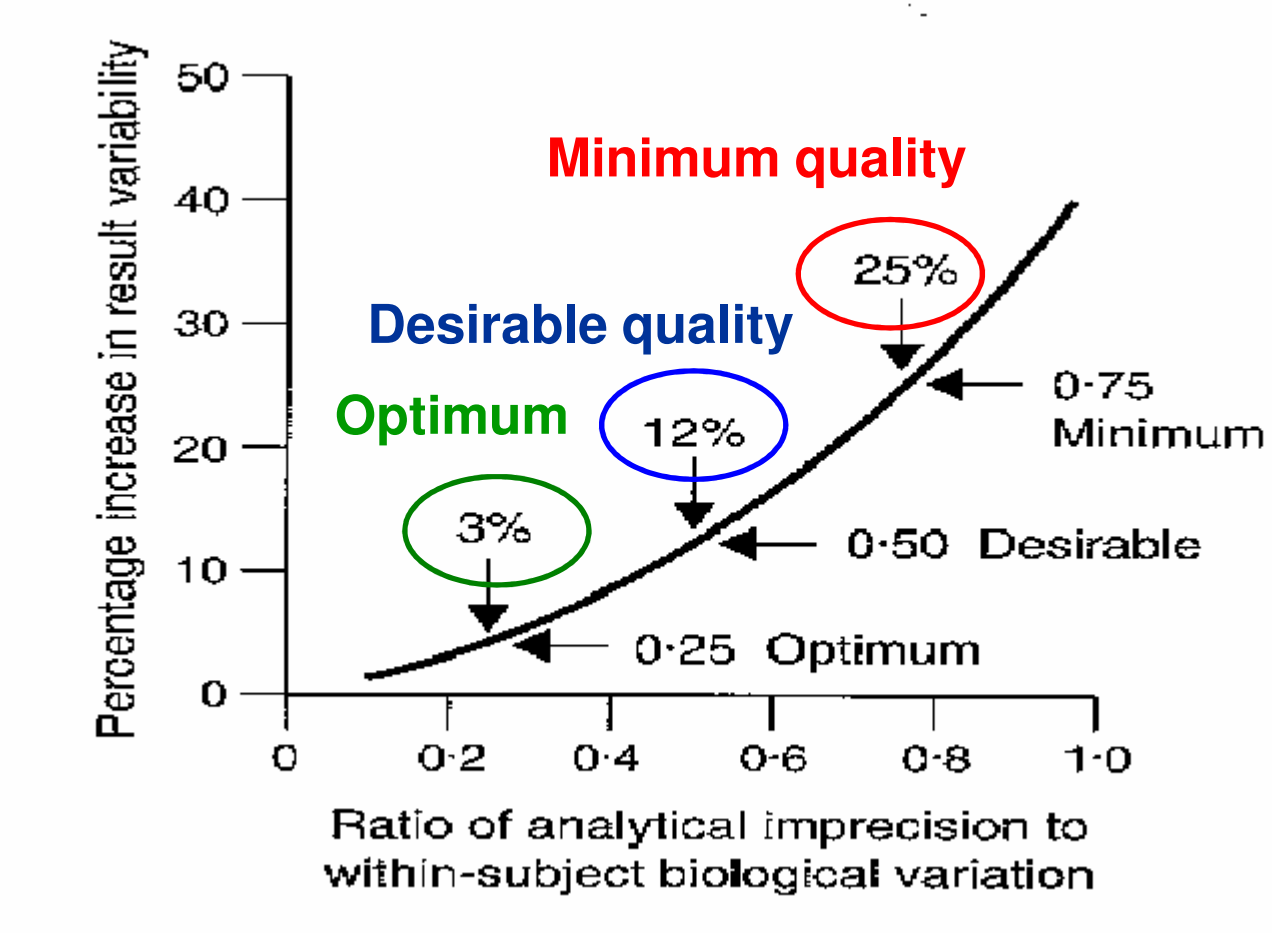
There was no discussion about measurements on ordinal scale

There was no conclusion about absolute and relative quality:

- Deviation from a 'true' value
- Deviation from the method mean

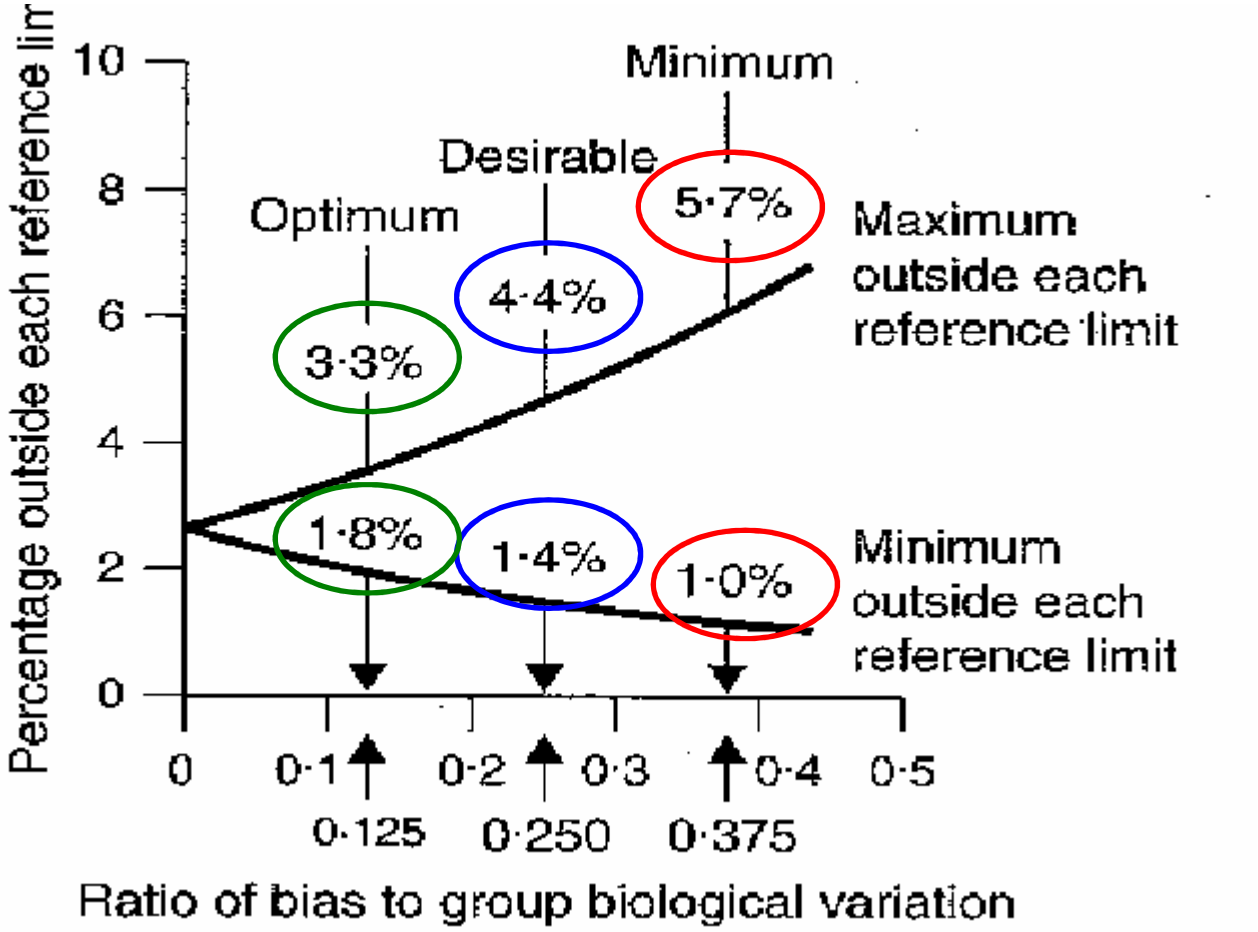
There was no agreement on which level of quality should be achieved

The effect of imprecision on test result variability



Fraser et al. *Ann Clin Biochem* 1997;34:8-12.

Effect of bias on reference values



Fraser et al. *Ann Clin Biochem* 1997;34:8-12.

What was not achieved in Stockholm?

There was no discussion about matrix-effects and consequently no specifications for allowable matrix

There was no discussion about measurements on ordinal scale

There was no conclusion about absolute and relative quality:

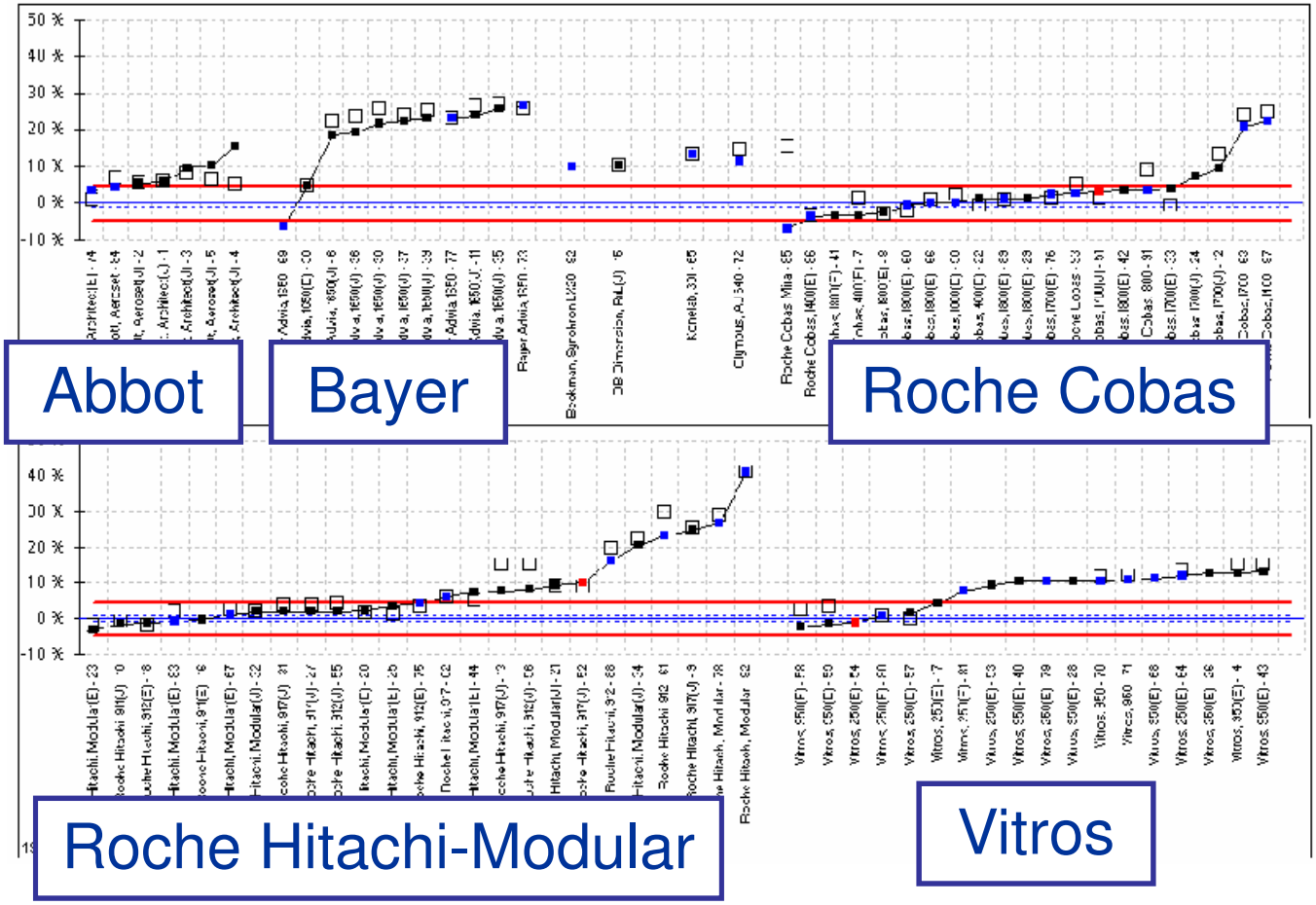
- Deviation from a 'true' value
- Deviation from the method mean

There was no agreement on which level of quality should be achieved

There was no agreement on consequences of poor quality

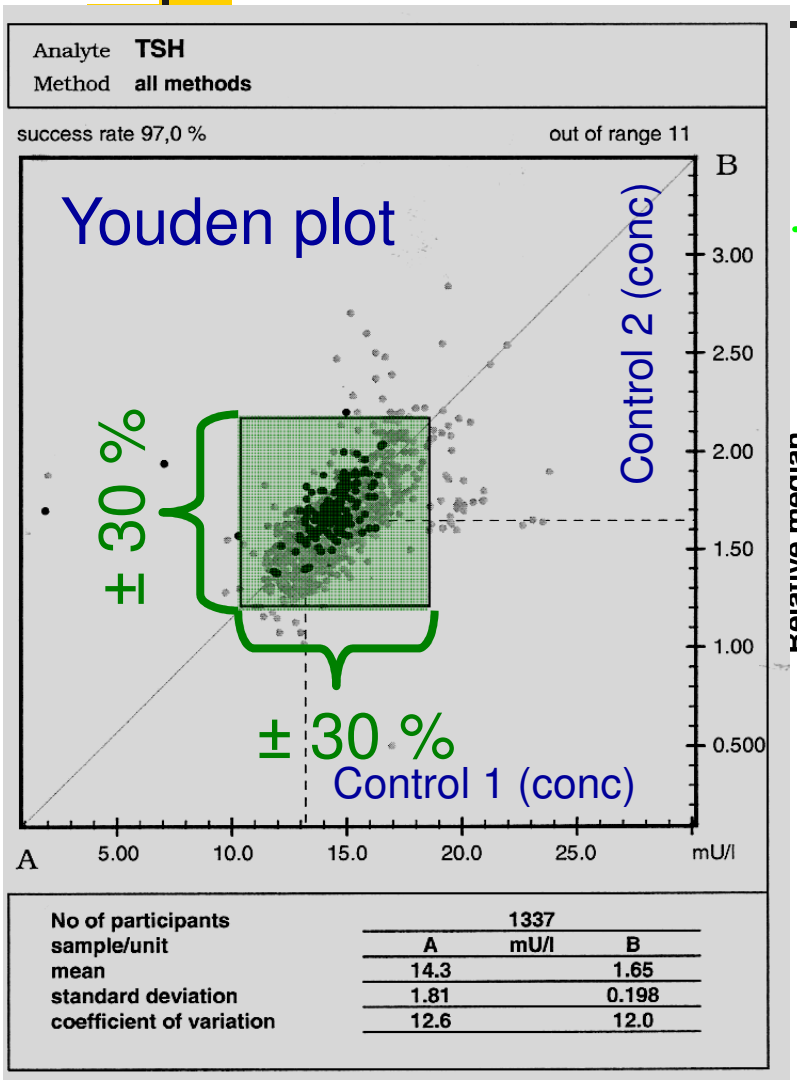
NORIP project on common reference intervals

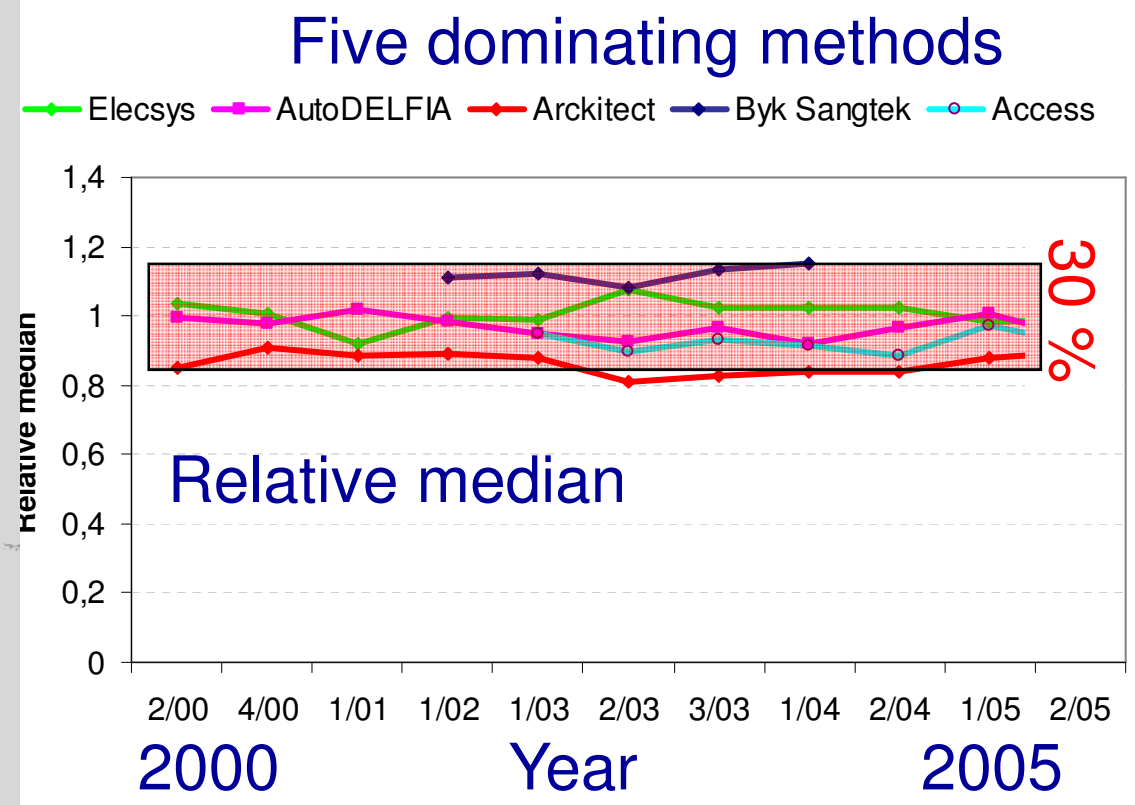
Creatinine



External control of TSH

Deutsche Gesellschaft für Klinische Chemie





Petersen et al. Symposium abstracts – IFCC – WorldLab Fortaleza, Clin Chem Lab Med 2008;46 special suppl:S148

What was not achieved in Stockholm?

There was no discussion about matrix-effects and consequently no specifications for allowable matrix

There was no discussion about measurements on ordinal scale

There was no conclusion about absolute and relative quality:

- Deviation from a 'true' value
- Deviation from the method mean

There was no agreement on which level of quality should be achieved

There was no agreement on consequences of poor quality

There was no agreement on the relation between clinical/biological specifications and specifications for EQAS and PT

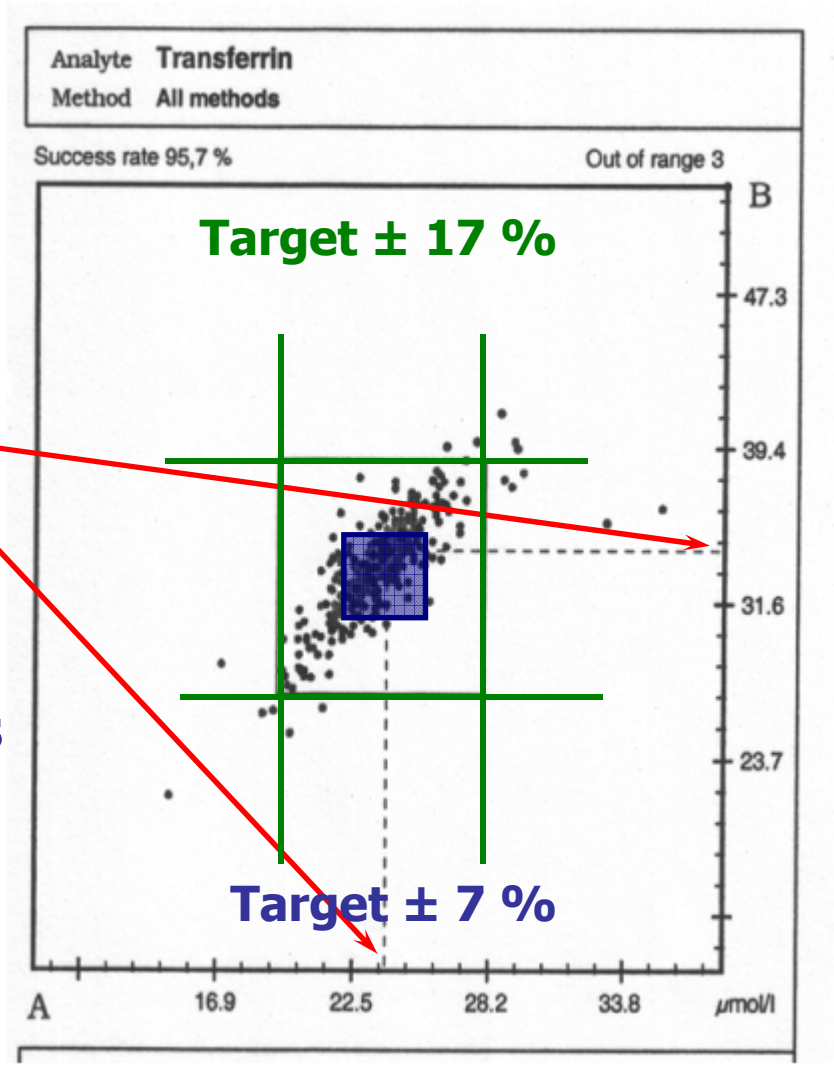
Transferrin: External Quality Assessment

Deutche Gesellschaft
für Klinische Chemie

Youden plot

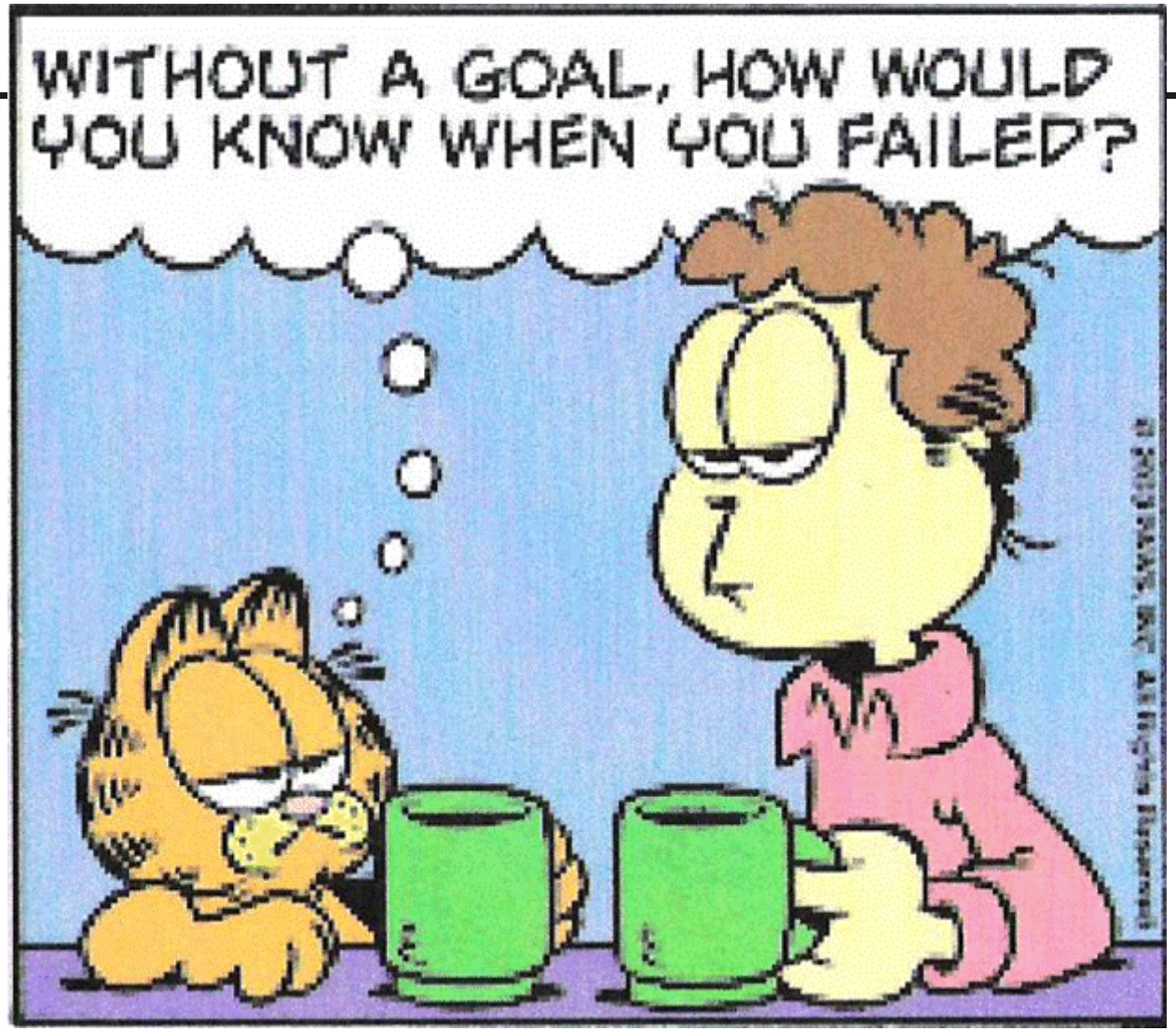
Transferrin results for two controls

Analytical quality specifications according to the EGE-Lab criteria: ± 7 %




I THINK SETTING GOALS
IS VERY IMPORTANT


GOOD IDEA


WITHOUT A GOAL, HOW WOULD YOU KNOW WHEN YOU FAILED?

What was not achieved in Stockholm?

There was no discussion about matrix-effects and consequently no specifications for allowable matrix

There was no discussion about measurements on ordinal scale

There was no conclusion about absolute and relative quality:

- Deviation from a 'true' value
- Deviation from the method mean

There was no agreement on which level of quality should be achieved

There was no agreement on consequences of poor quality

There was no agreement on the relation between clinical/biological specifications and specifications for EQAS and PT