Bias in clinical chemistry

Elvar Theodorsson
EFLM and Linköping University

LIU EXPANDING REALITY

Bias – a controversial subject

- Different perspectives
  - Reseachers
  - Users
  - Regulatory
  - Standardisation organisations
  - Metrologists
  - Industry
- Mutual respect and dialog is called for

LIU EXPANDING REALITY

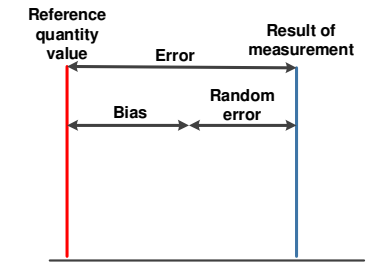
2

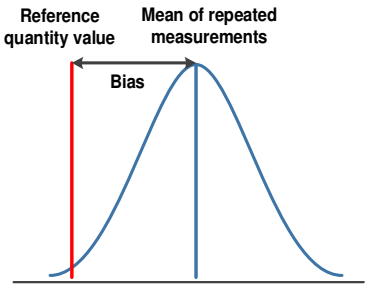
Topics for the present presentation

- Concepts and terms
- Automation has reduced repeatability- and day-to-day variation considerably.
- Bias has been reduced to a lesser extent than precision by reference measurement systems.
- Small and variable bias components will over time show random error properties and conventional random-error based methods for calculating measurement uncertainty can then be applied.
- Vital to minimize clinically important bias, especially bias within conglomerates of laboratories measuring samples from the same patients.
- Split sample/Mentor adept methods using patient samples are essential for this purpose

LiU EXPANDING REALITY

3

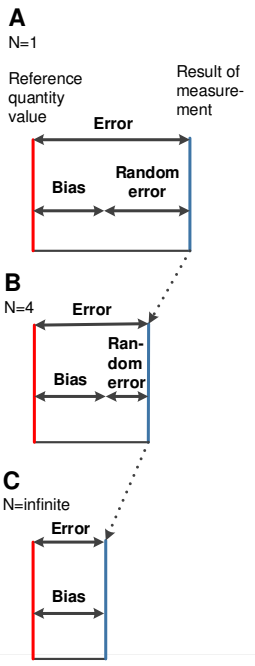




$$SEM = \frac{SD}{\sqrt{N}}$$

LiU EXPANDING REALITY

4



LiU EXPANDING REALITY

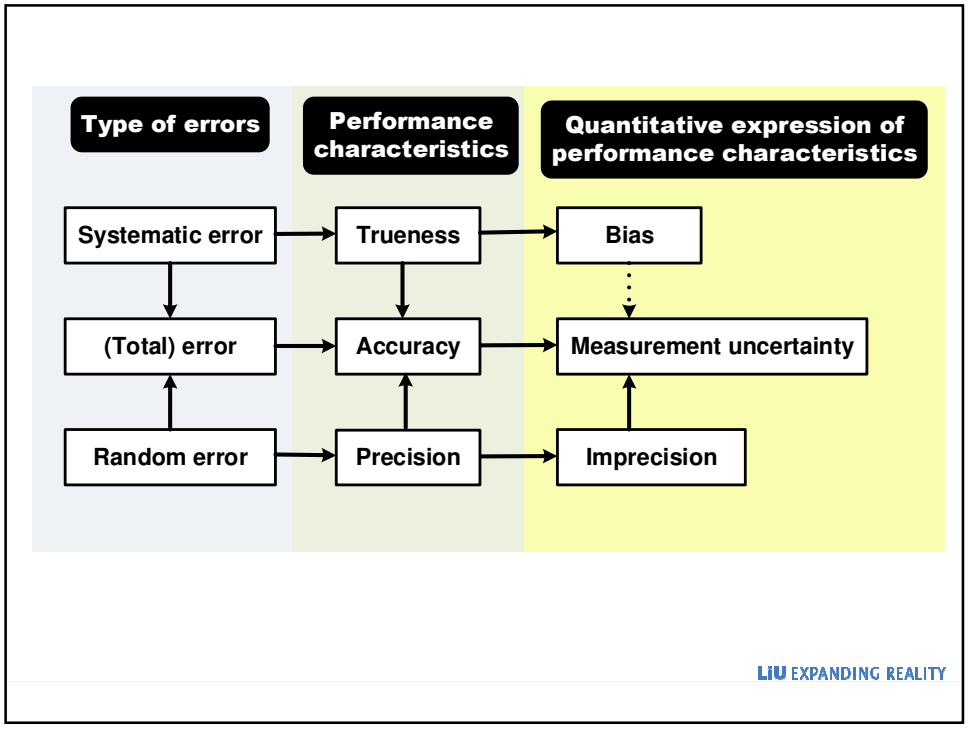
5

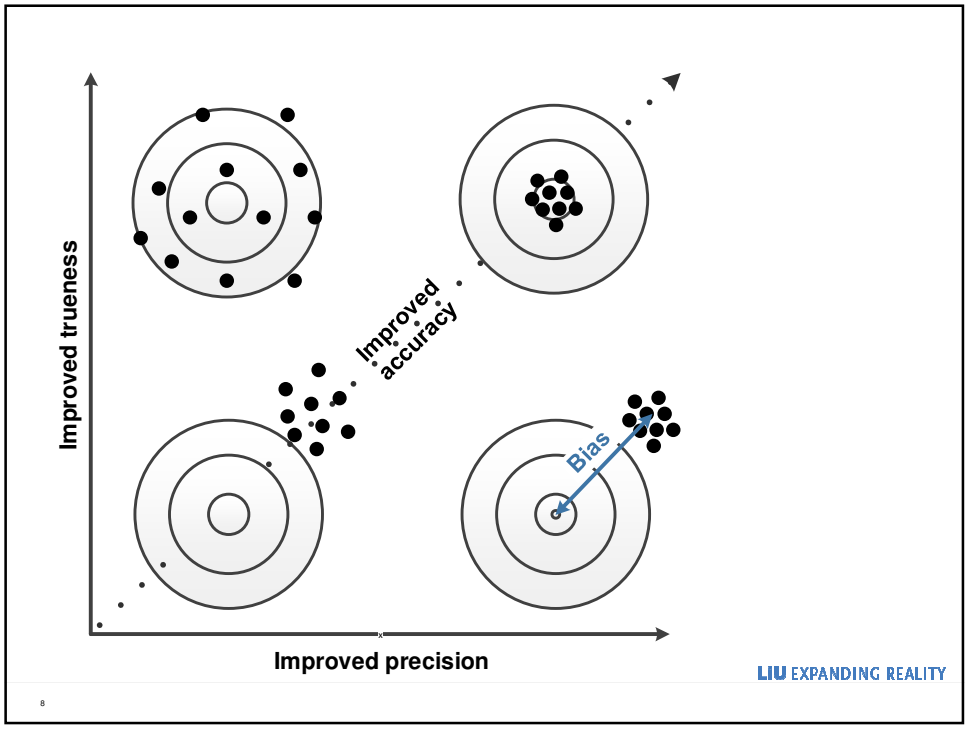
Bias

- **Trueness** is the “closeness of agreement between the average of an infinite number of replicate measured quantity values and a reference quantity value”. It is quantitatively expressed as ***bias***.

LiU EXPANDING REALITY

6


Type of errors
Performance characteristics
Quantitative expression of performance characteristics
Systematic error
Trueness
Bias
(Total) error
Accuracy
Measurement uncertainty
Random error
Precision
Imprecision
LiU EXPANDING REALITY

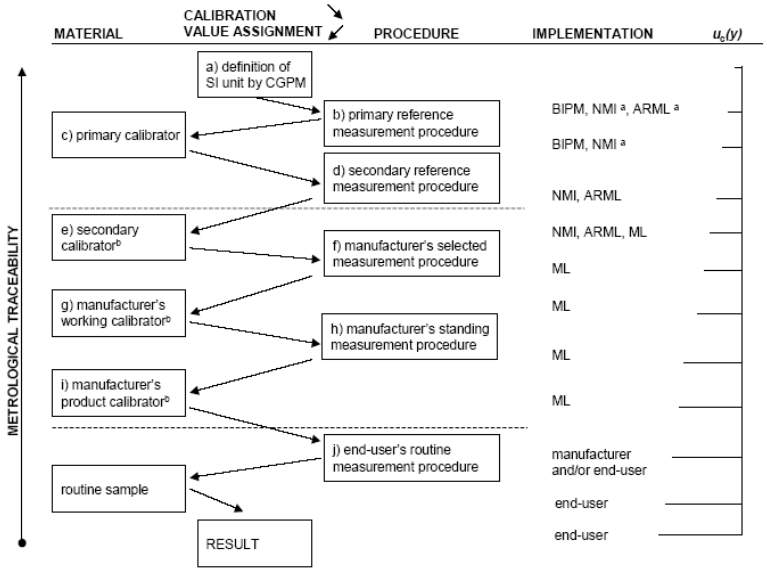

Improved trueness
Improved accuracy
Bias
Improved precision
LiU EXPANDING REALITY

Standards, reference measurement systems and organizations

- *The Joint Committee for Traceability in Laboratory Medicine* (JCTLM) was established in 2002 in response to the implementation of the European Community Directive 98/79/EC on in vitro medical devices
  - JCTLM publishes list of higher order reference materials, reference methods and reference laboratories
- *International Consortium for Harmonization of Clinical Laboratory Results* (ICHCLR)
  - AACC
- The International Federation of Clinical chemistry (IFCC)
- The European Federation of Clinical Chemistry and Laboratory Medicine (EFLM)

LiU EXPANDING REALITY

9



LiU EXPANDING REALITY

Causes of bias 1(3)

- Bias **when taking samples**, e.g. when samples are sometimes taken when the patient has been walking around and sometimes when he/she has been lying down. When the regulatory systems of the body adapt to gravity, the blood plasma volume is reduced in the order of 10% from a lying to a standing position thus increasing the concentration of macromolecules and cells in the blood of the patient.
- **Instability of the sample during transport or storage**, e.g. during transport in extremes of heat and cold and mechanical effects on cells and blood gases when transporting samples through pneumatic tubes in hospital transport systems.
- **Uncorrected loss of measurand at extraction** e.g. when preparing samples for measurement using high-performance liquid chromatography or mass-spectrometry.

LiU EXPANDING REALITY

11

Causes of bias 2(3)

- **Errors when the calibrator is prepared**, including errors in volume measurements or in weighing of calibrators in the laboratory
- Using sample **matrix** which differs from the matrix in the samples e.g. using de-fatted and lyophilized stable materials for internal quality control or proficiency testing programs.
- **Interferences** in the samples, e.g. the color of hemoglobin and bilirubin in hemolytic and icteric samples or the presence of high concentrations of proteins or lipids in the sample (myeloma or hyperlipidemia)
- The presence of molecules which **specifically interfere** with the reagents used in the measurement process, e.g. heterophilic antibodies (e.g. human antibodies against mouse IgG)

LiU EXPANDING REALITY

12

Causes of bias 3(3)

- **Specificity** for different epitopes in macromolecules of antibodies used in immunochemical measurement methods e.g. when measuring macromolecules including prostate-specific antigen, troponins and protein- or peptide hormones.

LiU EXPANDING REALITY

13

Clinically important or clinically unimportant bias

- Measurements in clinical chemistry are used for 1) **diagnosing** diseases or for 2) monitoring the effects of treatment weather a bias is clinically important depends on whether the method is used for diagnosing or for monitoring treatment effects
- A **clinically important bias** is a bias which is likely (with a predefined probability – commonly $p<0.05$) to influence the clinical decision between health and disease when studied in the context of all t
he other uncertainty components involved, including biological variation. A clinically unimportant bias is a bias which does not fulfill this criterion.

LiU EXPANDING REALITY

14

The within and between individuals components of biological variation of hemoglobin A1C and of Alanine aminotransferase

System	Component	Within-individual biological variation	Between-individuals biological variation
Blood	Hemoglobin A1C (HbA1C)	1.9%	5.7%
Serum	Alanine Aminotransferase (ALAT)	19.4%	41.6%

LiU EXPANDING REALITY

15

HbA1c and ASAT

- The between – individuals biological variation of HbA1C (5.7%) is much smaller than for ALAT (41.6%)
- A possible bias in the measurement of the concentrations of HbA1C is much more likely to influence clinical decisions in diagnosing diabetes mellitus than a possible bias in the measurement of ALAT when diagnosing e.g. liver conditions
- The large (41.6%) biological variation of ALAT is likely to be the major uncertainty component when the concentrations/activity of ALAT is used for diagnosis.
- A bias of e.g. 2% when measuring the concentrations/activity of ALAT is therefore usually clinically unimportant.
- A bias of e.g. 2% when measuring the concentrations of HbA1C is important

LiU EXPANDING REALITY

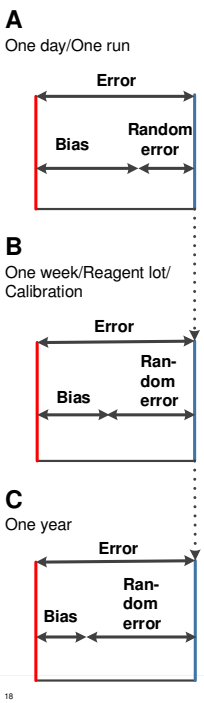
16

Whether a bias between measurement systems in clinical chemistry for a certain component is clinically important or unimportant is a question of

1. Knowledge about the medical risk that a certain concentration or change in concentrations implies
2. Whether the measurement is used for diagnosis or for monitoring of the effects of treatment
3. Knowledge about the biological variation of the component.

LiU EXPANDING REALITY

17



Variable bias components become random errors over time

- Provided clinically important and large bias components are reduced or eliminated, small bias components, e.g. caused by changes in reagent lots and re-calibration of measurement methods, will behave as random errors and routine methods for calculating measurement uncertainty based on random components can be used.

LiU EXPANDING REALITY

18

Measurement bias can be estimated using one or more of the following principles:

- Comparing the concentration found by laboratory's own methods with the stated concentration of a suitable certified reference material.
- Comparing the concentrations obtained by laboratory's own method in natural samples with the concentrations measured by a reference method in the same sample.
- Participating in proficiency testing schemes. The majority of these programs use consensus concentrations in modified control samples, but some use comparison with reference methods. Evidently the latter are preferable.
- Measuring the recovery of the measurand in spiked natural samples

LiU EXPANDING REALITY

19

In addition separate investigation of possible bias can be performed:

- By comparing the **serial dilution** of a natural sample or that of a spiked natural sample with the serial dilution of the calibration curve**.**
- Studying possible **interferences**, that is **selectivity**. Selectivity varies amongst different measurement methods and fields of study. In clinical chemistry the interferences by bilirubin, hemoglobin, lipids, proteins and drugs are most frequently occurring. Selectivity is "property of a measuring system, used with a specified measurement procedure, whereby it provides measured quantity values for one or more measurands such that the values of each measurand are independent of other measurands or other quantities in the phenomenon, body, or substance being investigated"

LiU EXPANDING REALITY

20

Types of bias

Type

- Constant
- Concentration-dependent

Expression

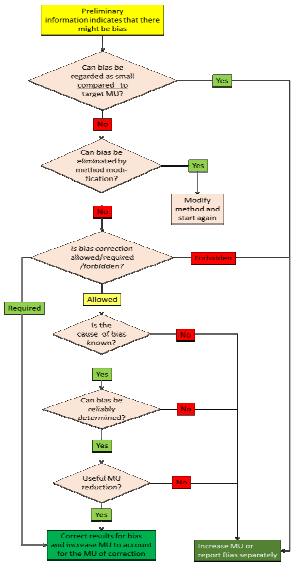
- Abslute $Bias = \bar{x} - y_o$
- Relative $Bias(relative) = \frac{\bar{x} - y_o}{y_o}$
  - Percentage $Bias(\%) = \frac{\bar{x} - y_o}{y_o} \; x \; 100$

LiU EXPANDING REALITY

21

To correct for bias or not, and by whom?

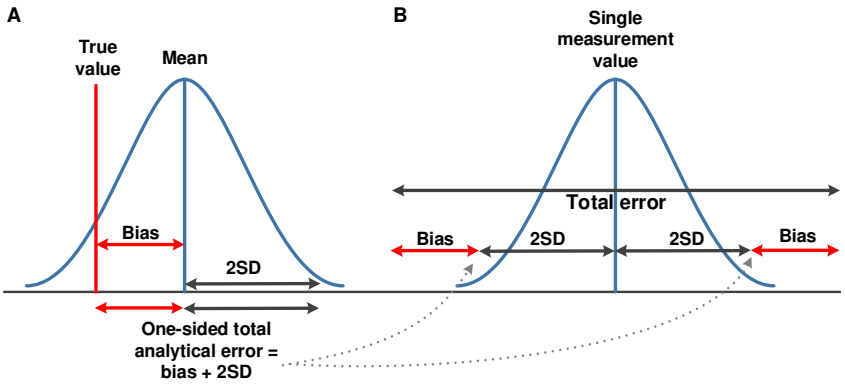
- There is no point in trying to eliminate or correct a small and clinically unimportant bias, since both elimination and correction need resources and may increase the measurement uncertainty.



LiU EXPANDING REALITY

22

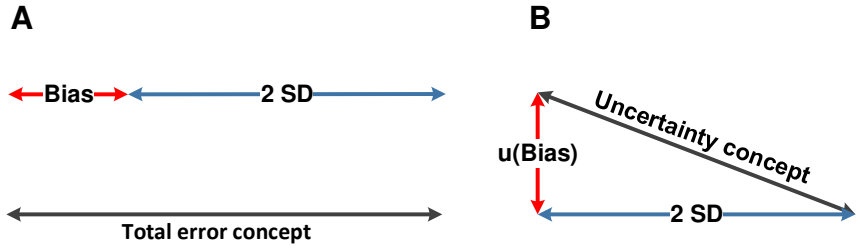
Total error approach



LiU EXPANDING REALITY

23

Total error and uncertainty approaches



LiU EXPANDING REALITY

24

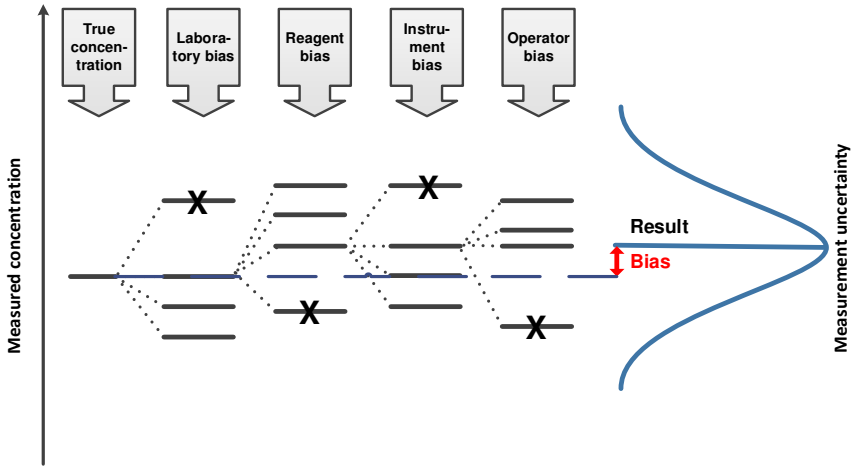
RiliBÄK- approach (Richtlinien der Bundesärztekammer)

$$\Delta_{max} = \sqrt{k^2 * s^2 + Bias^2}$$

- Δ_{max}=Maximum allowable error when measuring a control sample
- s = standard deviation
- k = a statistical coverage factor which depends on the purpose
- Bias = mean concentration measured in the control samples - target value of the control sample provided by its manufacturer

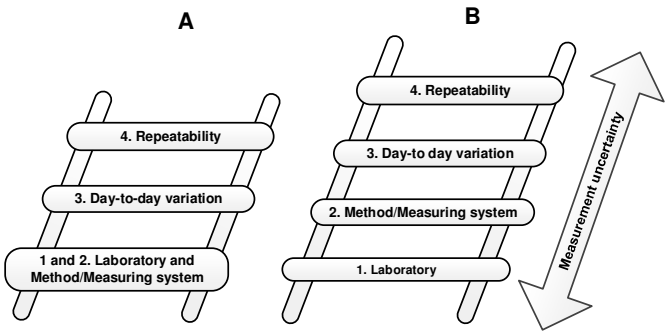
LiU EXPANDING REALITY

25



LiU EXPANDING REALITY

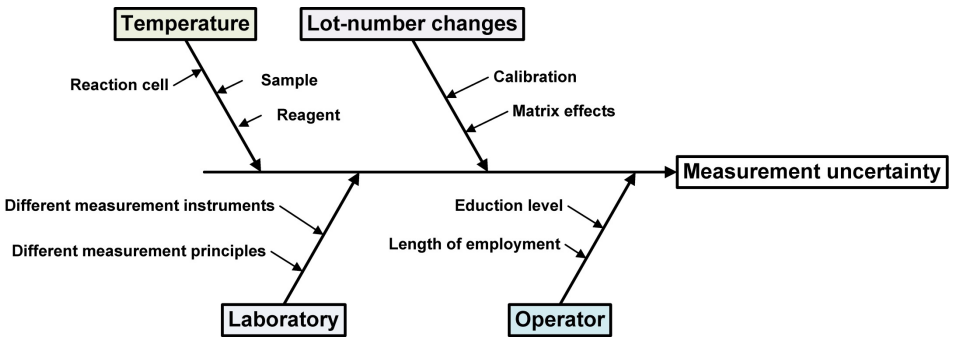
26



Step 1	The laboratory bias – a bias for an individual laboratory. The "laboratory" can be a single laboratory or a laboratory organisation e.g. all laboratories within a community that a patient is using
Step 2	The method and/or measurement system bias
Step 3	The day-to-day variation – a combination of random error and short-term bias owing to, among other factors, time effects, change of reagents etc.
Step 4	The repeatability – the random error occurring between replicate determinations performed within a short period of time.

LiU EXPANDING REALITY

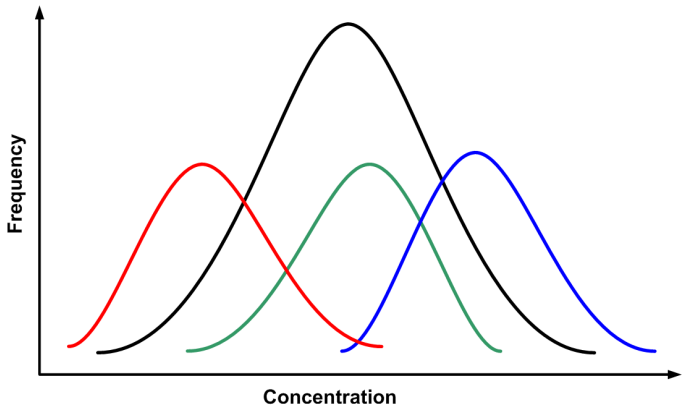
27



LiU EXPANDING REALITY

28

Bias – common in immunochemical measurement methods



Different colors depict different measurement methods

LiU EXPANDING REALITY

Differences in the epitopes that the antibodies react with

- Proteins are complex macromolecules containing several **epitopes**
- **Chance** determines which epitopes induce the production of antibodies
- The **specificity of the epitopes** determines the concentration measured
- International **calibrators** usually constitute a **mixture** of different epitopes

LiU EXPANDING REALITY

Matrix effects

- Effects on the final analytical results on **all other factors/substances** in the sample and in the sample container except those you intend to measure, e.g.
  - Sample container
  - Anticoagulants
  - Plasma proteins
  - Lipids

Control materials

- **Modified** to increase their stability during storage, e.g. by delipidation, addition of analytes and lyophilization– causes for matrix effects
- **Matrix effects** result in lack of modified control materials with addition of analytes to result in identical or comparable concentrations using all available techniques

Commutability

- To what extent reference materials, calibrators and control materials show matrix properties similar to those of fresh natural samples.
- Fresh natural patient samples represent the ultimately commutable materials for comparing measurement methods in clinical chemistry
- Natural patient samples are widely used in the industry to make sure that commercially available measurement methods measure the same concentrations in natural patient samples as reference methods, thereby making sure there is an unbroken traceability chain from reference materials to the routinely used measurement procedures

LiU EXPANDING REALITY

33

Secondary adjustment

- Secondary adjustment is usually performed usually by linear regression of the results from a properly calibrated adept method in order to eliminate its possible bias from the mentor method.
- Demings orthogonal regression

LiU EXPANDING REALITY

34

Patient samples are commutable

- The control materials are **modified** and the concentrations of the analytes in them adjusted by **addition**. Different instruments and methods may react differently to the consequent matrix effects
- Methods used for analysing patient samples should ideally not differ since normal patient sample is the sample matrix the methods were/are optimized for
- The most important issue is that the measurement instrument should report the **correct/optimally fit for purpose** results for patient samples.
- **Patient samples are optimal** for monitoring the quality of the analytical results for instruments and methods and for monitoring overall measurement uncertainty

Mentor measurement system

- A mentor measurement system in a conglomerate of laboratories is taken to be devoid of bias.

36

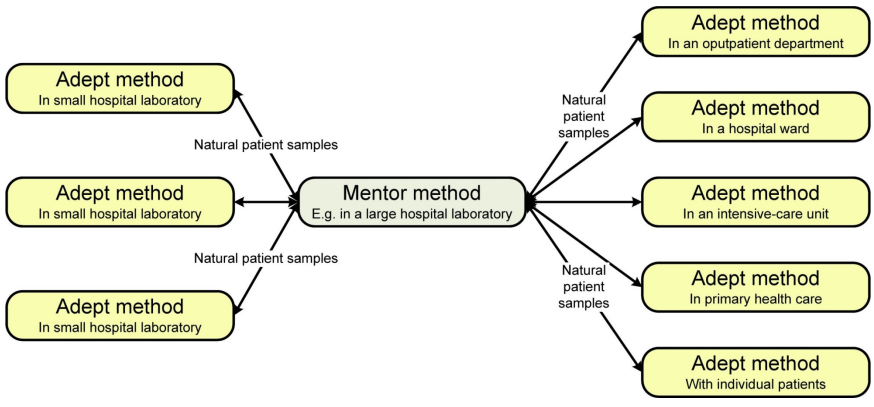
Split sample/mentor method

- A fresh natural sample measured using two measurement systems for the purpose of comparison, calibration or quality control.

LiU EXPANDING REALITY

37

The mentor principle



LiU EXPANDING REALITY

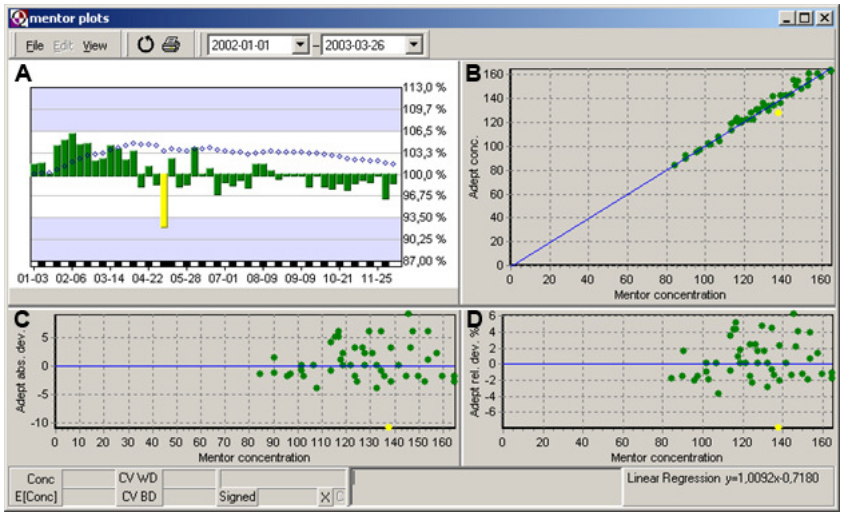
38

Norming the results

Express each of the adept values as a percent of the corresponding mentor value.

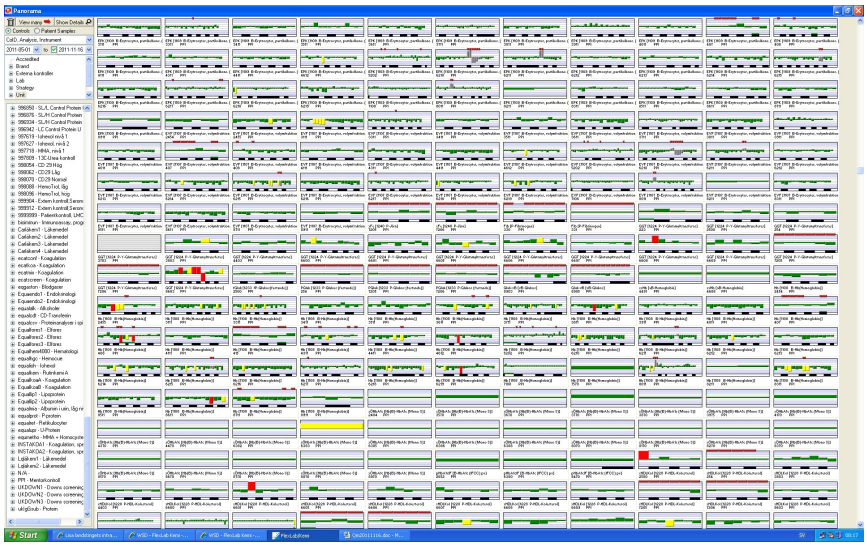
"The results of the adept method in this case is about 1% lower than the measurements performed on the mentor instrument. This bias varies with a standard deviation of 1,24%

Tidsstämpel	Instrument	Adept	Mentor	Normerat värde
2002-07-01 12:00	925	163,0	165,0	98,79%
2002-07-09 09:40	925	96,0	97,6	98,36%
2002-07-15 07:30	925	101,0	102,0	99,02%
2002-07-24 10:00	925	94,0	96,0	97,92%
2002-07-29 09:40	925	130,0	128,0	101,56%
2002-08-09 10:00	925	133,0	131,0	101,53%
2002-08-15 09:29	925	155,0	154,0	100,65%
2002-08-21 10:09	925	134,0	135,0	99,26%
2002-08-30 10:30	925	119,0	119,0	100,00%
2002-09-02 12:49	925	102,0	102,0	100,00%
2002-09-09 11:10	925	122,0	122,0	100,00%
2002-09-16 07:59	925	150,0	153,0	98,04%
2002-09-23 10:50	925	128,0	128,0	100,00%
2002-10-02 09:00	925	83,0	84,6	98,11%
2002-10-08 10:00	925	136,0	139,0	97,84%
2002-10-15 09:35	925	143,0	145,0	98,62%
2002-10-21 10:02	925	143,0	145,0	98,62%
2002-10-28 10:30	925	122,0	125,0	97,60%
2002-11-04 11:39	925	134,0	136,0	98,53%
2002-11-12 14:35	925	113,0	114,0	99,12%
2002-11-19 08:50	925	158,0	160,0	98,75%
2002-11-25 10:20	925	142,0	142,0	100,00%
2002-12-02 10:50	925	104,0	108,0	96,30%
2002-12-09 11:10	925	148,0	150,0	98,67%
			Medelvärde	**99,05%**
			SD	**1,24%**



LIU EXPANDING REALITY

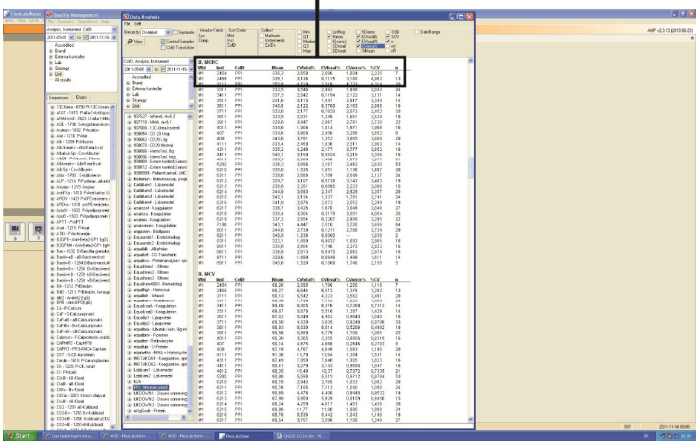
40



LiU EXPANDING REALITY

B, MCHC

Mtd	Inst	CoID	Mean	CVtotal%	CVtreat%	CVerror%	%CV	n
M1	2454	PPI	336,3	2,658	2,086	1,804	2,325	7
M1	2455	PPI	335,1	3,126	0,7115	3,180	4,963	13
M1	3111	PPI	350,8	4,719	2,319	4,222	4,214	20
M1	3311	PPI	332,5	3,546	2,992	1,946	2,042	24



LiU EXPANDING REALITY

42

Osbstacles to mentor-adept methods and to secondary adjustments

- Regulatory organizations including the EU (IVD) and the FDA
- Accreditation authorities
- Risks isolating the adept laboratories from the community of laboratories participating in regular external quality control/proficiency testing schemes

LiU EXPANDING REALITY

43

Accred Qual Assur (1999) 4:401–405

GENERAL PAPER

René Dybkaer

From total allowable error via metrological traceability to uncertainty of measurement of the unbiased result

Received: 5 March 1999
Accepted: 21 April 1999

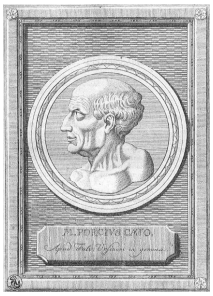
Presented at: 4th Conference on Quality [R]evolution in Clinical Laboratories, Antwerp, Belgium 29–30 October 1998

R. Dybkaer
Copenhagen Hospital Corporation, Department of Standardization in Laboratory Medicine, H:S Kommunehospitalet, Øster Farimagsgade 5, DK-1399 Copenhagen K, Denmark
Tel.: +45-33-38-37 85/86
Fax: +45-33-38-37-89

Abstract The concept of "total allowable error", investigated by Westgard and co-workers over a quarter of a century for use in laboratory medicine, comprises bias as well as random elements. Yet, to minimize diagnostic misclassifications, it is necessary to have spatio-temporal comparability of results. This requires trueness obtained through metrological traceability based on a calibration hierarchy. Hereby, the result is associated with a final uncertainty of measurement purged of known biases of procedure and laboratory. The sources of bias are discussed and the importance of commutability of calibrators and analytical specificity of the measurement procedure is stressed. The practicability of traceability to various levels and the advantages of the GUM approach for estimating uncertainty are shown.

Key words Metrological traceability · Total allowable error · Trueness · Unbiased result · Uncertainty of measurement

TY



"Ceterum censeo Carthaginem esse delendam" = Furthermore, I consider that Carthage must be destroyed

Marcus Porcius Cato = Cato the Elder (234-149 BC)

Ceterum censeo BIAS esse delendam!

LiU EXPANDING REALITY

45

Linköping University

expanding reality

www.liu.se