Performance criteria or ”quality specifications”

Gunnar Nordin

Dubrovnik Course in Zagreb 2015

EQUALIS

Do we need performance specifications?

EQUALIS

Do we need common performance specifications?

- Criteria for the use of a test in a specific clinical setting
- Criteria to share common reference interval and decision levels
- Criteria for acceptable performance in EQA

EQUALIS

The Stockholm consensus

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.

EQUALIS

EFLM
EUROPEAN FEDERATION OF CLINICAL CHEMISTRY AND LABORATORY MEDICINE

European Commission
Joint Research Centre
IRMM
Institute for Reference Materials and Measurements

CIRME

1st EFLM Strategic Conference

Defining analytical performance goals 15 years after the Stockholm Conference

8th CIRME International Scientific Meeting

Milan (IT)
24-25 November 2014

with the auspices of IFCC

EQUALIS

DE GRUYTER Clin Chem Lab Med 2015; 53(6): 833–835

Consensus Statement

Sverre Sandberg*, Callum G. Fraser, Andrea Rita Horvath, Rob Jansen, Graham Jones, Wytze Oosterhuis, Per Hyltoft Petersen, Heinz Schimmel, Ken Sikaris and Mauro Panteghini

Defining analytical performance specifications: Consensus Statement from the 1st Strategic Conference of the European Federation of Clinical Chemistry and Laboratory Medicine

The 5+ level hierarchy replaced by a 3 level hierarchy:

1. Evidence
2. Biology
3. Technique

EQUALIS

What type of evidence?

Clinical outcome

- Mortaility
- Time to treatment
- Financial benefits

Few studies

- Troponin test for diagnosis of acute coronary syndrome
- Rapid test for Strep A, to reduce the prescription of antibiotics
- Blood glucose for the monitoring of diabetes, and some more

EQUALIS

What type of evidence?

Clinical expectations are also evidence

The experienced clinician ”knows” the performance of a test.

Opinion Paper

Geir Thue and Sverre Sandberg*

Analytical performance specifications based on how clinicians use laboratory tests. Experiences from a post-analytical external quality assessment programme

EQUALIS

What biological variation?

Biological variation

- Between individual
- Within individual

EQUALIS

What biological variation

CIs and Power Calculations for Within-Person Biological Variation: Effect of Analytical Imprecision, Number of Replicates, Number of Samples, and Number of Individuals

Thomas Røraas,[1*] Per H. Petersen,[1] and Sverre Sandberg[2,3]

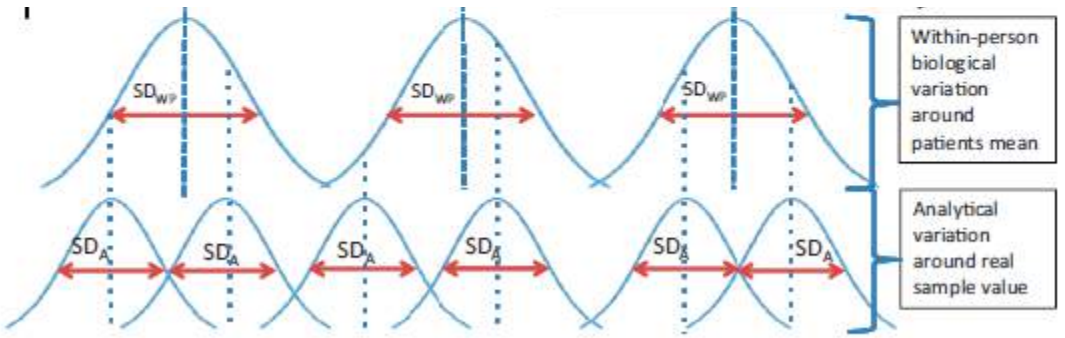


Fig. 1. The 2-level nested ANOVA model illustrating the relationship between the SD_{BP} and the SD_{WP} as well as the SD_A.

Röraas et al 2012

EQUALIS

Biological variation - terminology

Table 1. Terms and symbols most commonly used to define the component of BV applicable to individuals and groups.

Term	Frequency of term	Symbols	Frequency of symbol
Applicable to individuals			
Within-subject biological variation	18	CV_I	35
Intraindividual biological variation	11	CV_w	8
Intraindividual variation	11	CV_I	7
Intraindividual variation	10	CV_w	6
Within-subject variation	9	CV_w	6
Within-subject coefficient of variation	5	CV_I	6
Intraindividual variability	4	$CV_{within\text{-}subject}$	3
Within-person biological variation	4	$CV_{biological}$	3
Within-person variation	4	CV_b	3
Intraindividual biological variation	3		
Intraindividual CV	3		
Within-subject CV	3		
Within-subject biological variation	3		
Applicable to groups			
Between-subject biological variation	16	CV_G	29
Interindividual variation	14	CV_g	7
Between-subject variation	7	CV_g	5
Interindividual biological variation	5	CV_b	4
Interindividual CV	5	CV_b	2
Interindividual variability	5		
Between-person variation	4		
Between-subject coefficient of variation	4		
Interindividual variation	4		

Simundic et al, 2015

What biological variation?

Anna Carobene*

Reliability of biological variation data available in an online database: need for improvement

DE GRUYTER

Clin Chem Lab Med 2015; 53(6): 879–885

Opinion Paper

William A. Bartlett*, Federica Braga, Anna Carobene, Abdurrahman Coşkun, Richard Prusa, Pilar Fernandez-Calle, Thomas Røraas, Neils Jonker and Sverre Sandberg, on behalf of the Biological Variation Working Group, European Federation of Clinical Chemistry and Laboratory Medicine (EFLM)

A checklist for critical appraisal of studies of biological variation

EQUALIS

Length of study	3.4 (C)	Length of the study periods should be clearly identified.
Sampling	3.5 (C)	Sampling protocols (e.g., subject preparation, sampling conditions) that minimise pre-analytical variation should be adequately described to enable transportability of the data [25]. Numbers of samples taken should be sufficient to deliver the required power to the study [25, 26].
Samples	3.6 (C)	Recorded details should include the beginning and end date of the study and timings of sampling. Sampling conditions and sample type should be described in detail. Pre-analytical storage conditions of samples should be described.
Conditions for analysis of samples	3.7 (C)	A description of conditions under which the samples were analysed. Analytical protocols should be designed to minimise sources of analytical variation (Optimal Conditions Precision) [24].
Data analysis	4	Data analysis techniques should be described. The power of the study to identify indices of biological variation should be calculated and presented[b] [26].
Outlier analysis	4.1 (C)	Outliers should be excluded from the final analysis of the data. Test for outliers should be applied to all levels of data (between replicate analysis, between samples within subject, between subjects) [25]. The numbers of outliers and reasons for their exclusion must be given.
Heterogeneity of variance	4.2 (C)	Subjects with outlying within subject variance should be rejected from calculations used to determine an estimate of common true variance. The numbers of outliers and reasons for their exclusion must be given[b].
Statistical methods described and appropriate	4.3 (C)	Statistical methods used should be appropriately identified, fit for purpose and referenced. Data that do not conform to a normal distribution should be appropriately transformed [25].
Results	5	Unified terminology [13] should be used and appropriately defined metadata clearly presented to enable understanding and transportation of the data through time and across health care systems.
Terminology	5.1 (D)	Terms and symbols should be used to describe biological variation should conform standards identified by Simundic et al. [13].
Results clearly presented and managed	5.2 (D)	Biological variation data, with derived indices, should be tabulated in a format that enables extraction of the key data unambiguously associated with a minimum data set to enable transportability of the data. Power of the study and confidence limits around estimates of biological variation should be presented [26]. The results section should clearly identify the results of outlier analysis undertaken and confirm homogeneity of the data sets. If data are stratified the variables used to enable this should be clearly characterised.

Bartlett et al, 2015

What biological variation?

Biological variation

- Between individual
- Within individual

Anything else?

- Matrix effects
- Sample specific error components

EQUALIS

What biological variation?

- If the within-subject variation is very heterogenic, a common figure that is representative for a group of individuals can not be found.
- Performance specification can not be based on data unless representative for the target patient population.
- The matrix effect, or sample specific error component, must be considered.

EQUALIS

What technique?

”State of the art” quality?

The imprecision should be less than half of the within-subject biological variation (CV_{bw}) or the total imprecision should be less than the imprecision that can be achieved by the better 50% of laboratories in the external quality schemes.

Quality specifications as stated by an EQA oragnisation

The best possible quality?

EQUALIS

Task and finish group "Allocation of laboratory tests to different models for performance specifications" (TFG-DM).

TFG - DM

- *Terms of Reference:* To **allocate different tests to different models** recognized in the Strategic Conference Consensus Statement and **to give an overview and a reason for why tests are allocated to the different models**.
- *Deliverable:* To **produce a list of laboratory tests** allocated to the different performance specifications (**starting with the most common**) to be put on the EFLM website. **To publish a paper describing the rationale behind listing** the different tests in the different model groups.

TFG – DM composition

Chair **Ferruccio Ceriotti -** Italy

Members

George Klee - USA
Pilar Fernández-Calle - Spain
Gunnar Nordin - Sweden
Mauro Panteghini - Italy
Sverre Sandberg - Norway
Thomas Streichert - Germany
Joan-Lluis Vives Corrons - Spain

1) Possible criteria

1. The measurand has a central role in diagnosis and monitoring of a specific disease $\Rightarrow$ outcome criteria;
2. The measurand has a high homeostatic control $\Rightarrow$ BV criteria;
3. Neither central diagnostic role nor sufficient homeostatic control $\Rightarrow$ state-of-the-art.

One or several different performance specifications for each test?

Different specification for POCT and hospital use of a test?

Different specification due to the intended use of a test?

- Screening versus confirmation tests?
- Use for monitoring, diagnosis or something else?

Yes, we need one performance specification for each intended use!

EQUALIS

Equalis performance specifications (”quality goals”)

Different specifications for different materials

Tabell 1. Allmän klinisk kemi. Kvalitetsmålen är uppställda [illegible] expertgrupp för Allmän klinisk kemi.

Storhet	Maximal avvikelse (±%)		Maximal avvikelse i absoluta tal vid referensintervallets övre gräns	Kommentar
	Modifierat serum	Naturligt serum		
P—ALAT	12	12+	0,13 µkat/L	
P—Albumin	5	5	2,4 g/L	
P—ALP	12	12	0,22 µkat/L	
P—Amylas	12	12	0,24 µkat/L	
P—Pankreasamylas	12	12	-∓	
P—ASAT	12	12	0,09 µkat/L	
P—Bilirubin	12	12	3,0 µmol/L	
P—Bilirubin, konjugerat	12	12	-∓	
P—Calcium	3	3	0,08 mmol/L	
P—CK	12	12	0,81 µkat/L	
P—Fosfat	6	6	0,10 mmol/L	
P—Glukos	10	10	-	
P—GT	12	12	0,23 µkat/L	
P—HDL-kolesterol	10	10	0,27 mmol/L	
P—Järn	12	12	4,1 µmol/L	
P—Kalium	4	4	0,18 mmol/L	
P—Klorid	2	2	-∓	
P—Kolesterol	5	5	0,39 mmol/L	
P—Kreatinin	8	8	8,4 µmol/L	
P—Laktat	12	12	-∓	
P—LD	12	12	0,50 µkat/L	
P—LDL-Kolesterol	12	12	0,64 mmol/L	

EQUALIS

Equalis performance specifications (”quality goals”)

Different specifications for individual results and groups of results (methods or ”conglomerates”)

Tabell 2. Hematologi. Kvalitetsmålet uppställt av Equ[illegible]xpertgrupp för hematologi.

Storhet	Maximal avvikelse för enskilt resultat från målvärde (±%)	Maximal avvikelse för hel metodgrupp (metodbias) (±%)	Kommentar
B—Hemoglobin	5	2	
B—Leukocyter	15	6	
B—EVF	5	2	
B—Trombocyter	16	6	
B—Lymfocyter	16		
B—Granulocyter	23		
B—MCV	3	1	
B—Erytrocyter	5	2	

EQUALIS

Different quality specifications due to different ways to calculate them?

	Equalis	Labquality	RCPAQAP	Rili-bäk
P-Albumin	5 %	5 %	6 %	20 %
P-Calcium	3 %	3 %	4 %	10 %
P-Phosphate	6 %	6 %	8 %	16 %
P-Chloride	2 %	2 %	3 %	8 %
P-Creatinine	8 %	8 %	8 %	20 %
P-Cholesterol	5 %	5 %	6 %	13 %
B-Haemoglobin	5 %	5 %		6 %
B-Leukocytes	15 %	10 %		18 %
HbA1c	7 % at 48 mmol/mol	8 %	8 %	18 %
P-CRP	10 % (hosp) 15 % (POCT)	15 %		20 %
Erc-MCV	3 %	5 %		

EQUALIS

A complex matter

For each measurand:

Quality specification x number of intended uses x number of EQA-materials x number of calculation models x etc x ….

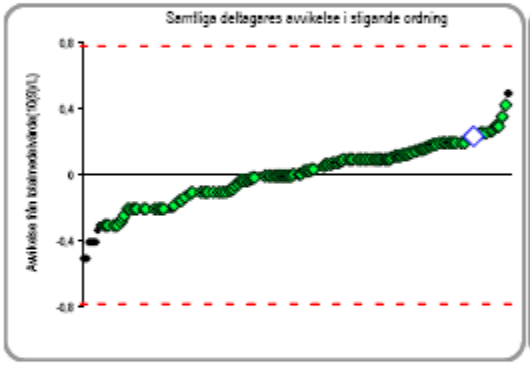
We need a simple model….

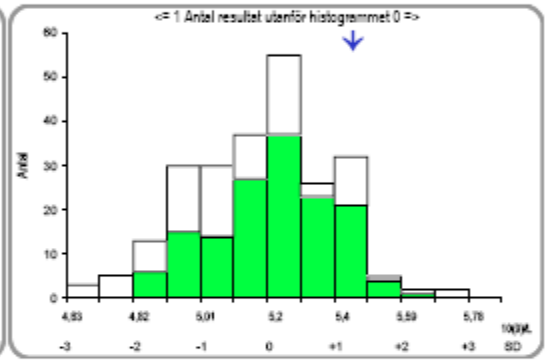
EQUALIS

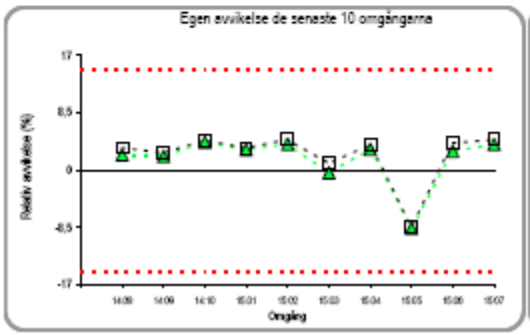
B-Leukocyter (10(9)/L)

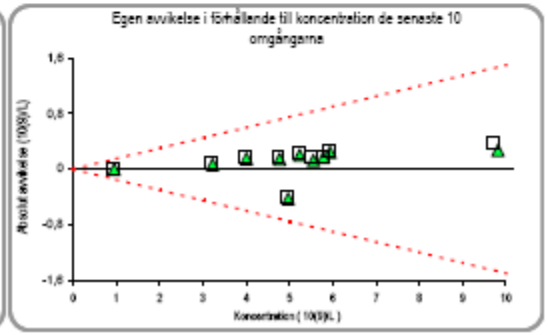
Kvalitetsmål (%): +/- 15
Egen rapportgrupp: Sysmex

		Egen rapportgrupp (148)		Samtliga (241)	
Eget resultat:	5,44	Medelvärde:	5,24	Medelvärde:	5,20
		SD:	0,17	SD:	0,19
Åsatt värde:	-	CV%:	3,2	CV%:	3,7
		Egen avvikelse		Egen avvikelse	
		Absolut (enheter):	+0,2	Absolut (enheter):	+0,24
		Relativ (%):	+3,8	Relativ (%):	+4,5
		Antal SD	+1,20	Antal SD	+1,23
		Medelavvikelse%:	+1,58	Medelavvikelse%:	+2,29
		(senaste 10 omgångarna)		(senaste 10 omgångarna)	







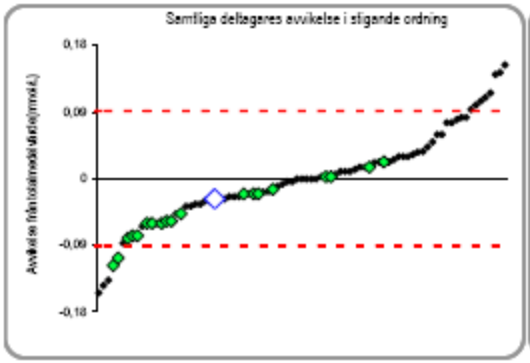


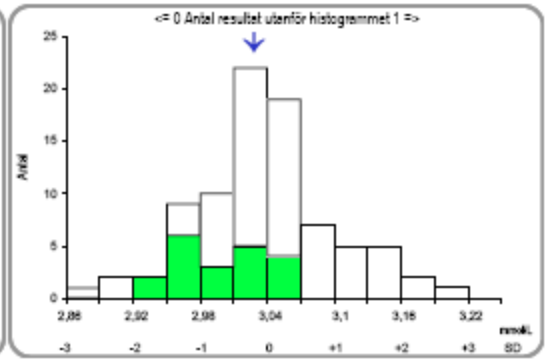
EQUALIS

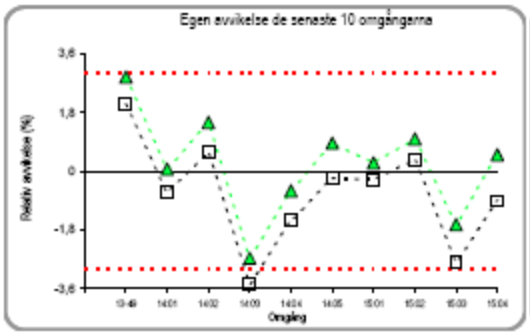
P-Calcium (mmol/L)

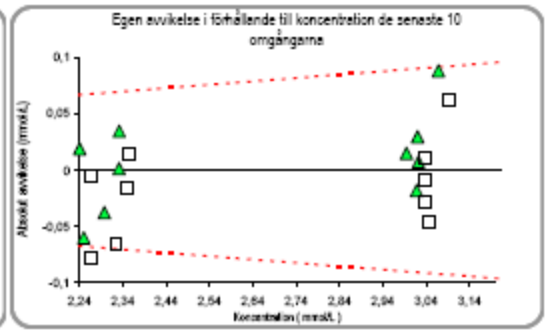
Kvalitetsmål (%): +/- 3
Egen rapportgrupp: Siemens

		Egen rapportgrupp (20)		Samtliga (86)	
Eget resultat:	3,01	Medelvärde:	2,99	Medelvärde:	3,04
		SD:	0,04	SD:	0,06
Åsatt värde:	-	CV%:	1,5	CV%:	2,0
		Egen avvikelse		Egen avvikelse	
		Absolut (enheter):	+0,02	Absolut (enheter):	-0,03
		Relativ (%):	+0,5	Relativ (%):	-0,9
		Antal SD	+0,35	Antal SD	-0,46
		Medelavvikelse%:	+0,22	Medelavvikelse%:	-0,68
		(senaste 10 omgångarna)		(senaste 10 omgångarna)	









EQUALIS

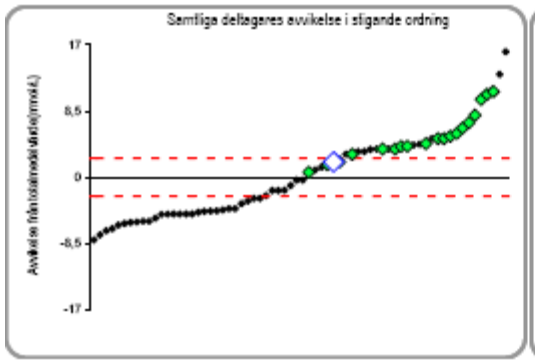
P-Klorid (mmol/L)

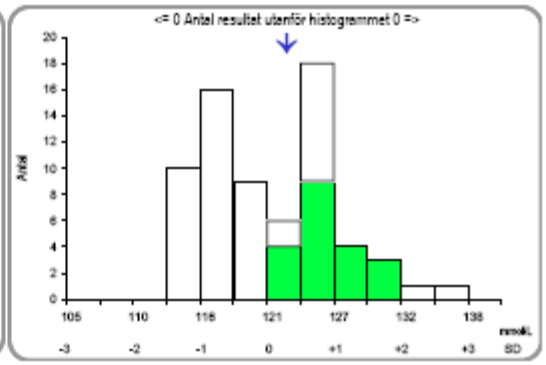
Kvalitetsmål (%): +/- 2
Egen rapportgrupp: Siemens

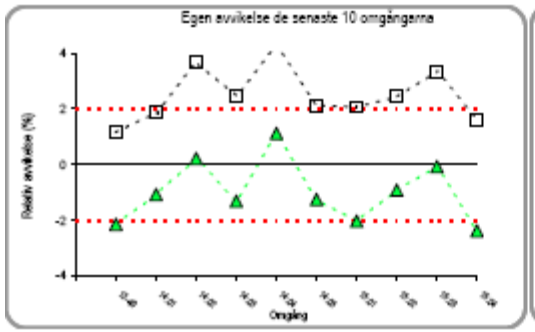
Eget resultat:	123
Åsatt värde:	-

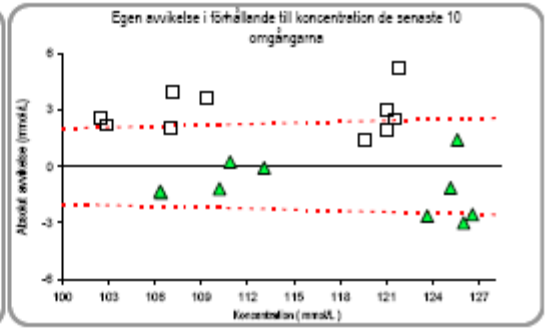
Egen rapportgrupp (20)	
Medelvärde:	126
SD:	3,02
CV%:	2,4
Egen avvikelse	
Absolut (enheter):	-2,99
Relativ (%):	-2,4
Antal SD	-0,99
Medelavvikelse%:	-0,98
(senaste 10 omgångarna)	

Samtliga (68)	
Medelvärde:	121
SD:	5,50
CV%:	4,5
Egen avvikelse	
Absolut (enheter):	+1,96
Relativ (%):	+1,6
Antal SD	+0,36
Medelavvikelse%:	+2,5
(senaste 10 omgångarna)	









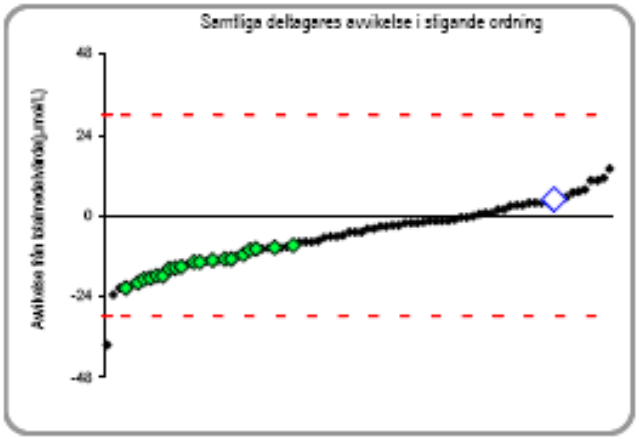
EQUALIS

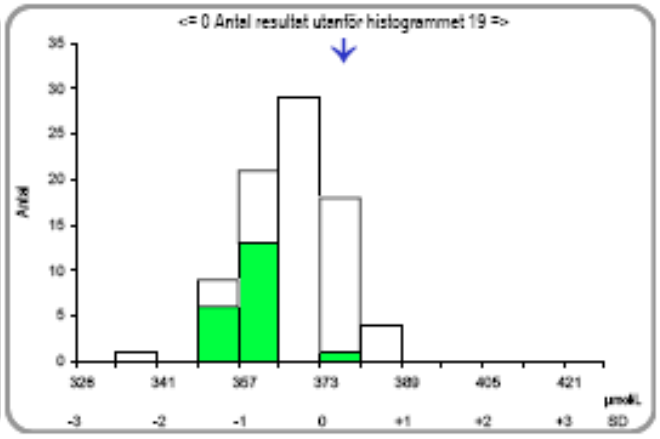
P-Kreatinin (µmol/L)

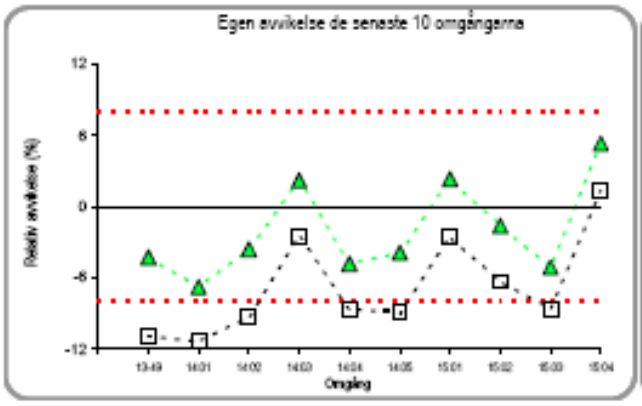
Kvalitetsmål (%): +/- 8

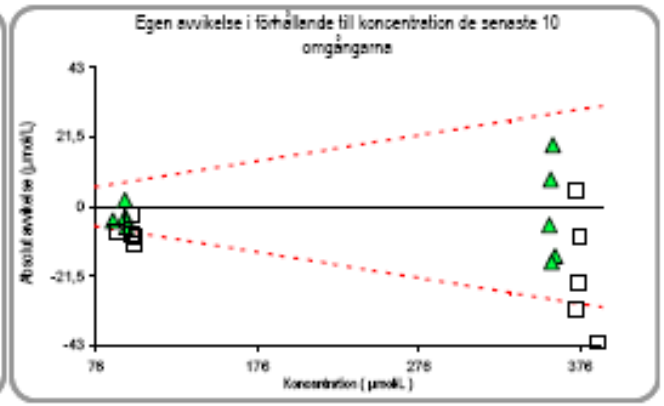
Egen rapportgrupp: Siemens enzymatisk

		Egen rapportgrupp (20)		Samtliga (101)	
Eget resultat:	378	Medelvärde:	359	Medelvärde:	373
		SD:	4,43	SD:	15,87
Åsatt värde:	-	CV%:	1,2	CV%:	4,3
		Egen avvikelse		Egen avvikelse	
		Absolut (enheter):	+19	Absolut (enheter):	+4,8
		Relativ (%):	+5,3	Relativ (%):	+1,3
		Antal SD	+4,30	Antal SD	+0,30
		Medelavvikelse%:	-2,01	Medelavvikelse%:	-6,76
		(senaste 10 omgångarna)		(senaste 10 omgångarna)	









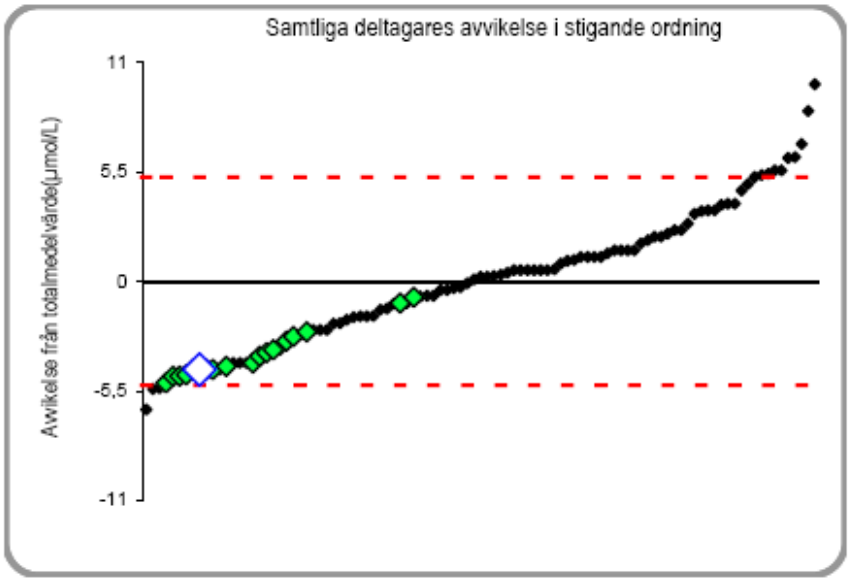
EQUALIS

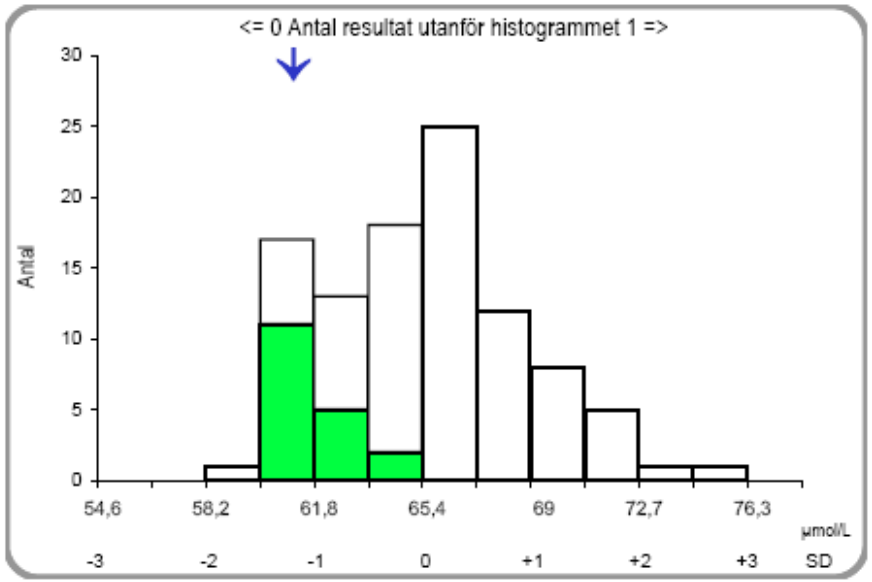
P-Kreatinin (µmol/L)

Kvalitetsmål (%): +/- 8

Egen rapportgrupp: Siemens enzymatisk

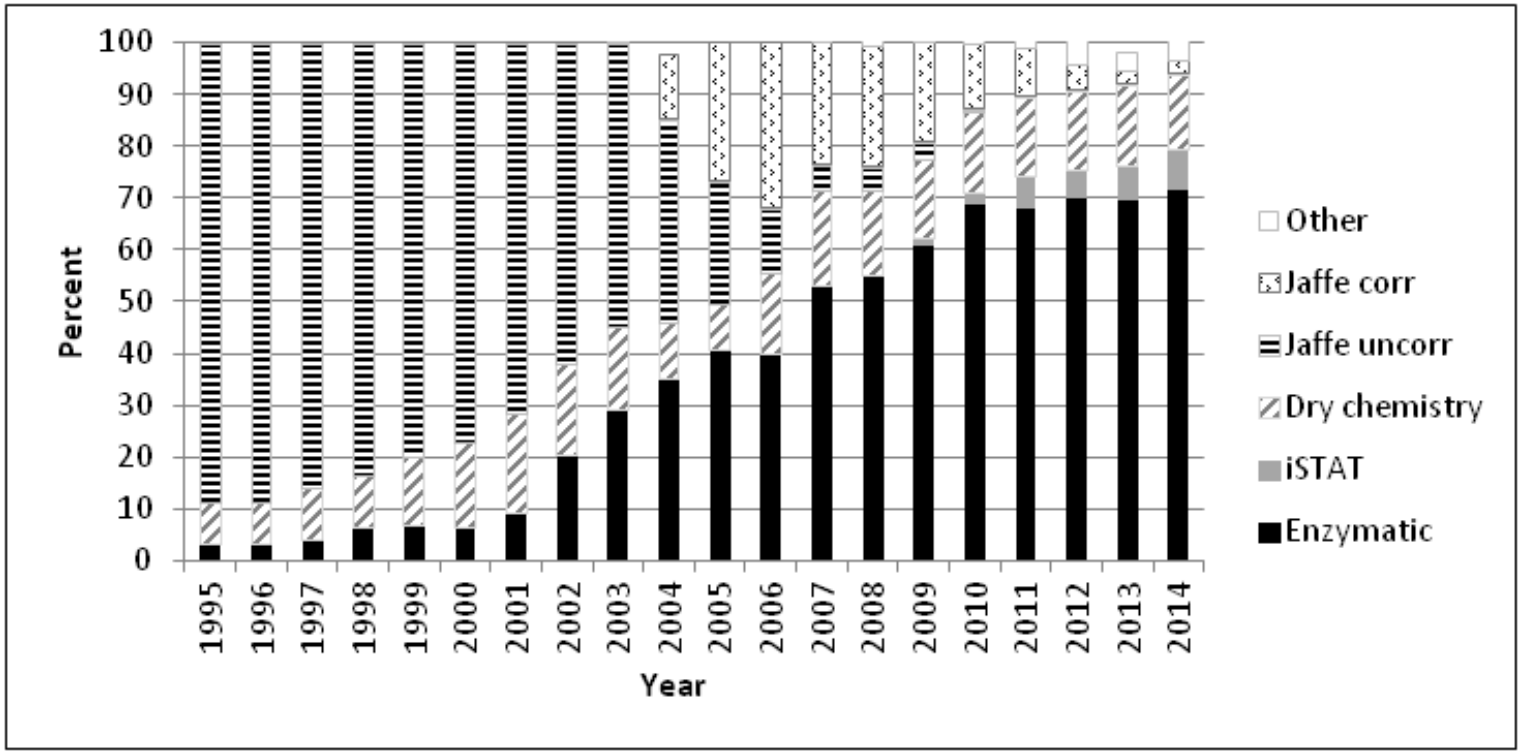
		Egen rapportgrupp (18)		Samtliga (102)	
Eget resultat:	**61**	Medelvärde:	61,6	Medelvärde:	65,4
		SD:	1,05	SD:	3,62
Åsatt värde:	-	CV%:	1,7	CV%:	5,5
		Egen avvikelse		Egen avvikelse	
		Absolut (enheter):	-0,6	Absolut (enheter):	-4,43
		Relativ (%):	-1,0	Relativ (%):	-6,8
		Antal SD	-0,57	Antal SD	-1,22
		Medelavvikelse%:	-1,83	Medelavvikelse%:	-6,97
		(senaste 10 omgångarna)		(senaste 10 omgångarna)	





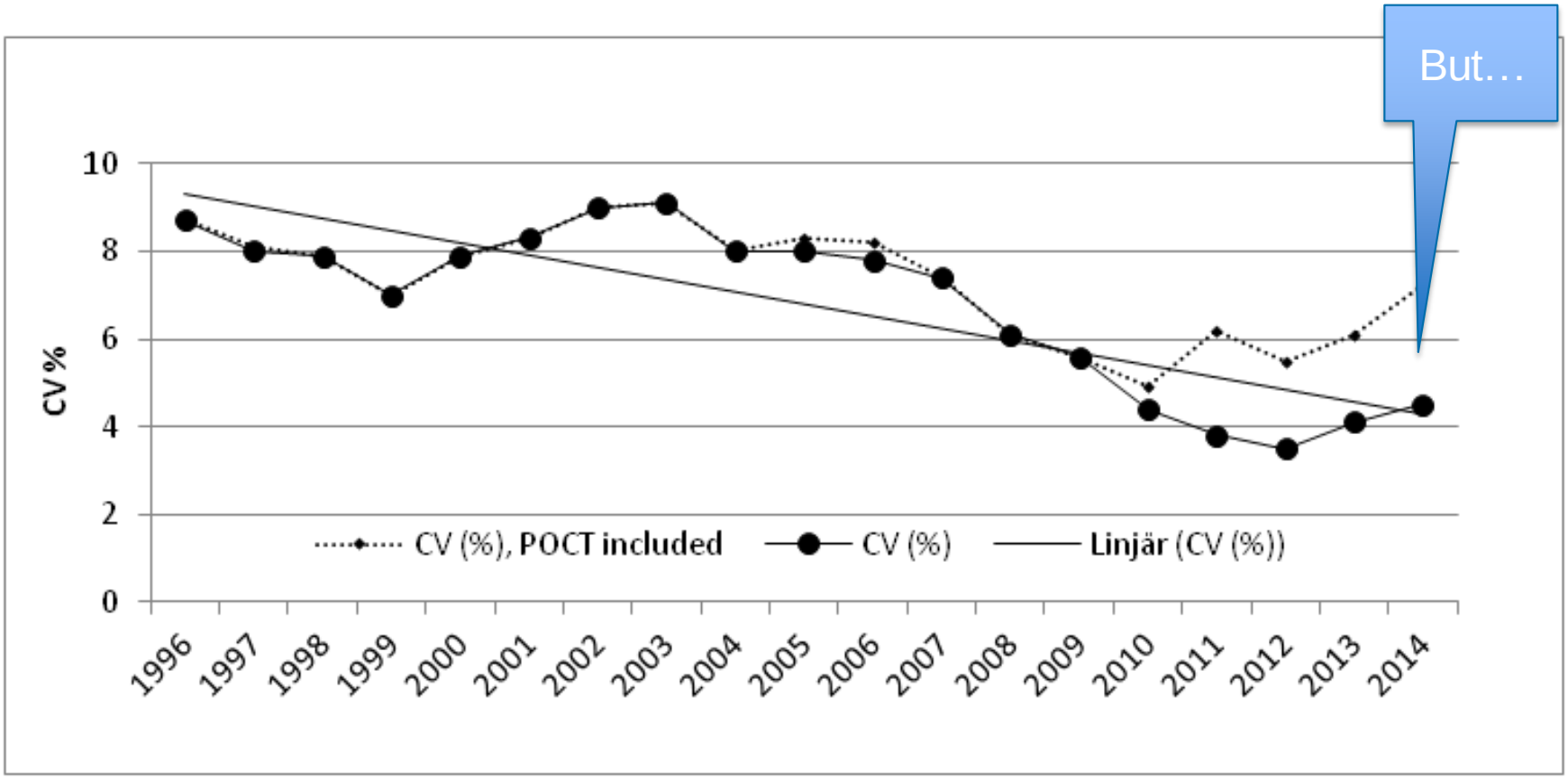
EQUALIS

A transition from Jaffe to enzymatic methods in Sweden



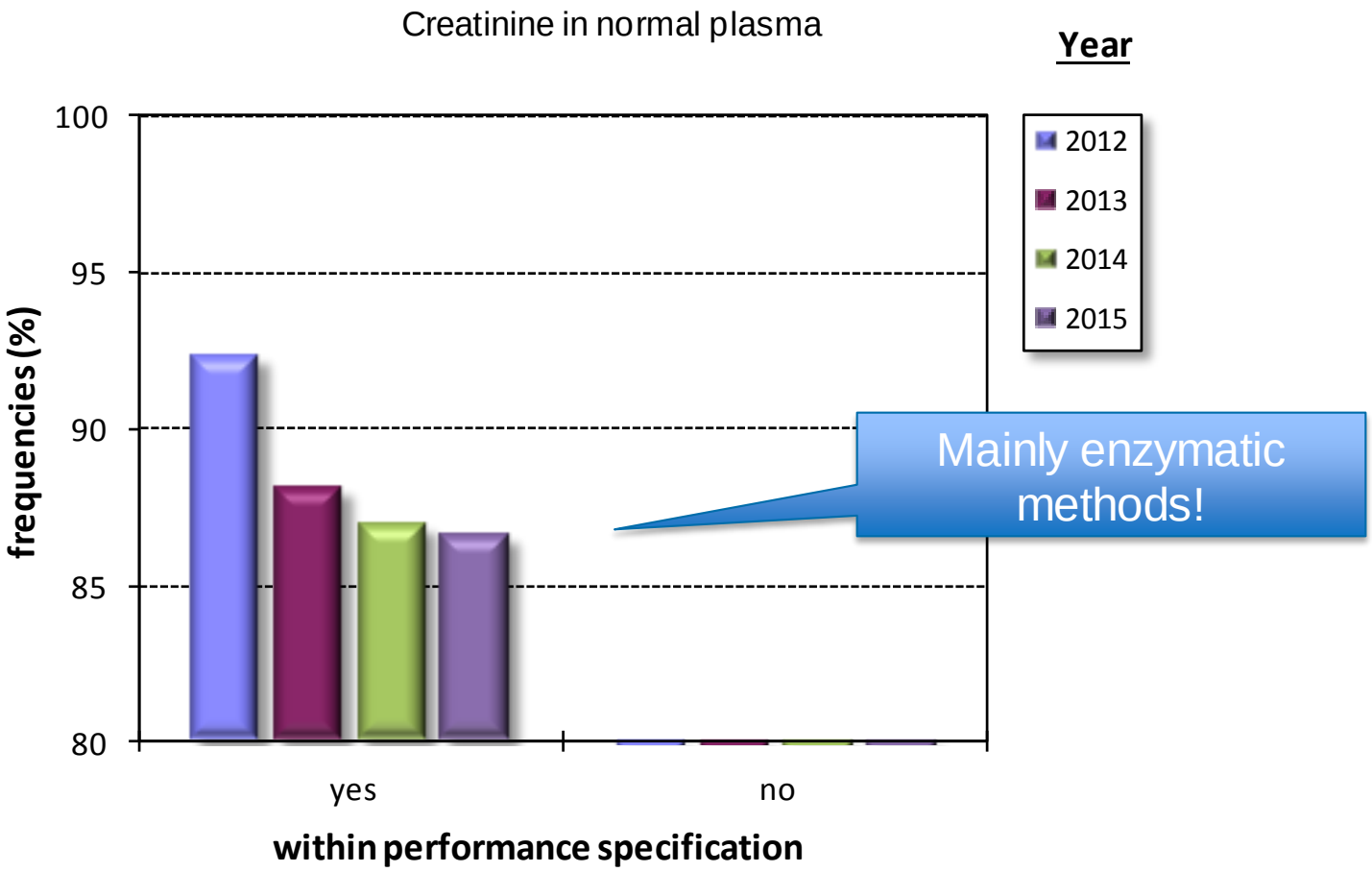
EQUALIS

Improved performance over years



EQUALIS

The last 4 years: the fraction of creatinine results within performance specification (+/- 8 %) has declined



EQUALIS

The example glucose

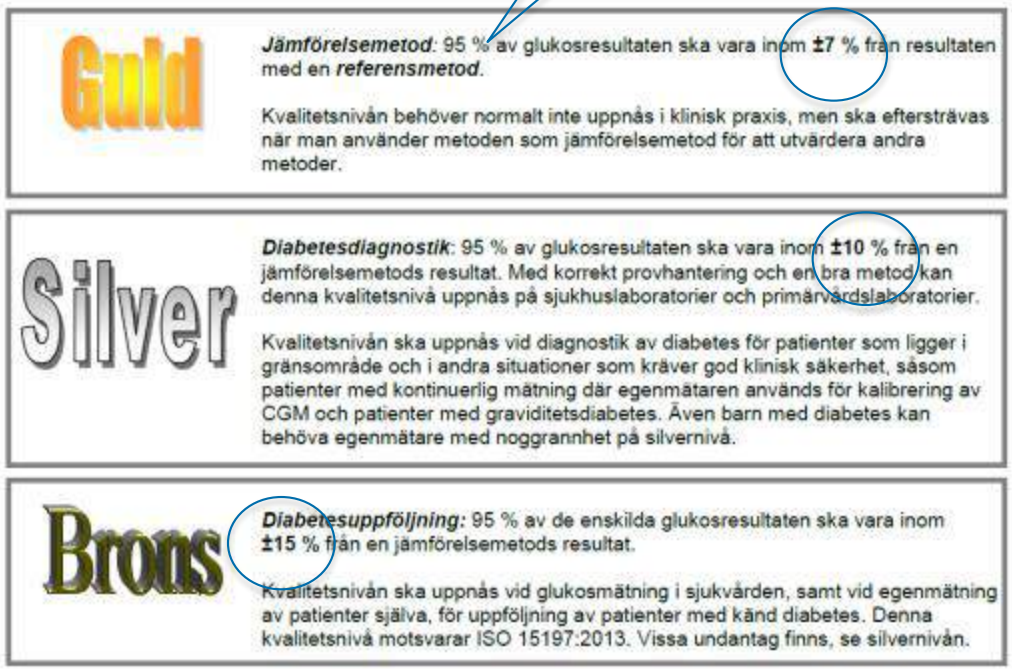
A JCTML recognized RMP!

Performance specifications has been agreed with the professional societes in Sweden for three different intended uses:

Comparison method:

Diagnostic method:

Monitoring method:



EQUALIS

The example glucose

Definition of the measurand.

1) The component or analyte is easy to defind: glucose
2) The measurand: the concentration of glucose in the patients blood plasma

Practical advice to laboratories and manufacturers on how to verify their methods

1) Step 1: Verify that the selected comparison method fulfills 'gold criteria': TEa 7%, including preanalytical errors and sample specific errors.

2) Step 2: Compare your working method with the comparison method.

   95% of the results should be within +/- 10% of the comparison method (silver criteria) or +/- 15 % (bronze criteria)

EQUALIS

The example glucose

The total allowable error should also consider **preanalytical errors** and the '**matrix effect**' ('sample specific errors')

5.3. Resultatets osäkerhet

Mätrutinens totalfel kan beräknas med följande formel för både venösa och kapillära prover:

$$Totalfelet = \pm 1{,}64 \times \sqrt{imprecision^2 + matriseffekt^2 + preanalytiskt\ slumpfel^2} + |bias|$$

Totalfelet får alltså vara högst ±10 % om kvalitetsmålet ska anses vara uppfyllt. Observera dock att matriseffekt samt preanalytiskt fel är svårkvantifierade.

EQUALIS

The example glucose

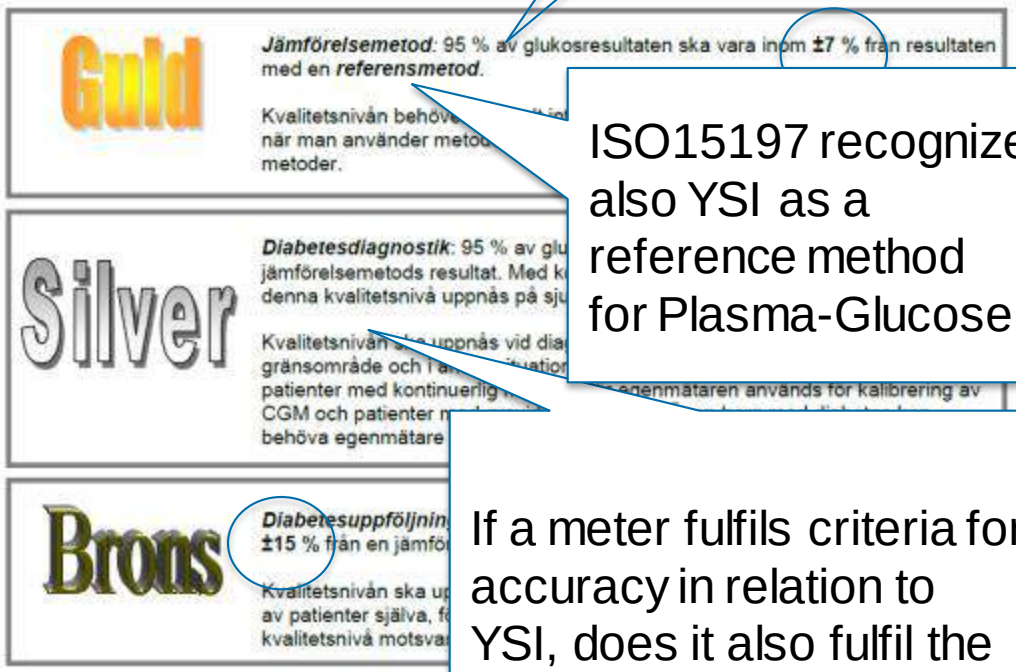
A JCTML recognized RMP!

Performance specifications has been agreed with the professional societes in Sweden for three different intended uses:

Comparison method:

Diagnostic method:

Monitoring method:



ISO15197 recognize also YSI as a reference method for Plasma-Glucose.

If a meter fulfils criteria for accuracy in relation to YSI, does it also fulfil the Silver criteria?

EQUALIS

The summary

Quality specifications might be useful

Different specifications according to the intended use, sample material used, etc, make the situation very complicated.

A simplified model must be used to reduce the number of possible specifications to a number of needed specifications

A hard work to underpin specifications

EQUALIS

EQUALIS

Take a home message

Performance specifications for test results are needed, simply in order to evaluate if results from a test method fulfils them or not.

Different performance specifications might be needed according to the intended use of a test and according to which method that is being used to specify the quality of the test results..

The number of possible performance specifications need to be restricted to a number of needed performance specifications.

Quality specifications should be based on one of the three models; clinical evidence, biological variation and state-of-the-art. The EFLM Task and finish group "Allocation of laboratory tests to different models for performance specifications" (TFG-DM) will discuss how the three models should be implemented for different measurands and various intended uses of the test results.

EQUALIS