Commutability of control material: how should we examine it?

Christa Cobbaert, PhD, EurSpLM
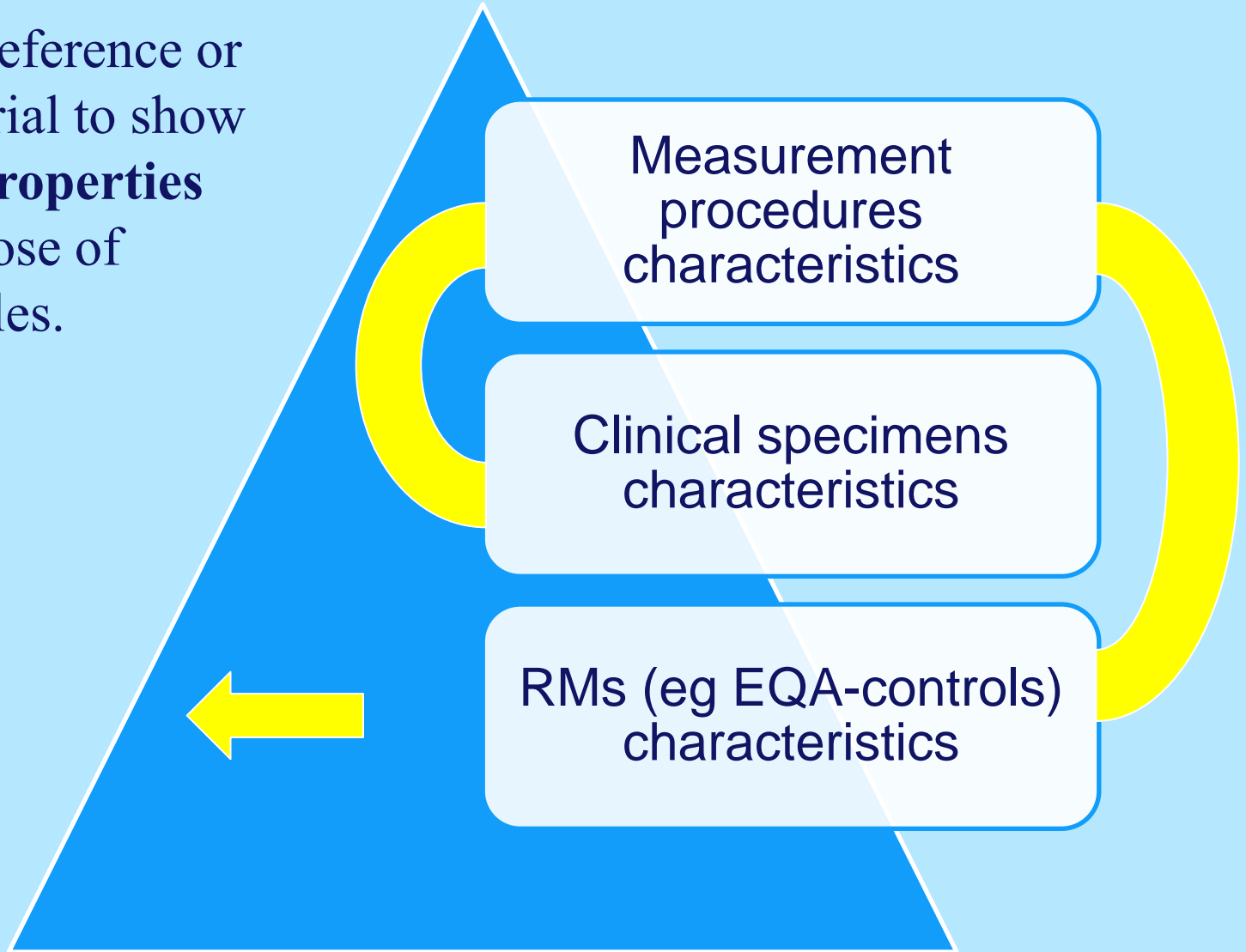on behalf of the SKML Chemistry Section
Chair Calibration 2.000
10 October 2013

Content

I. Introduction

Ability of a reference or control material to show **interassay properties similar** to those of human samples.

Similar interassay properties?

Measurement procedures characteristics

Clinical specimens characteristics

RMs (eg EQA-controls) characteristics

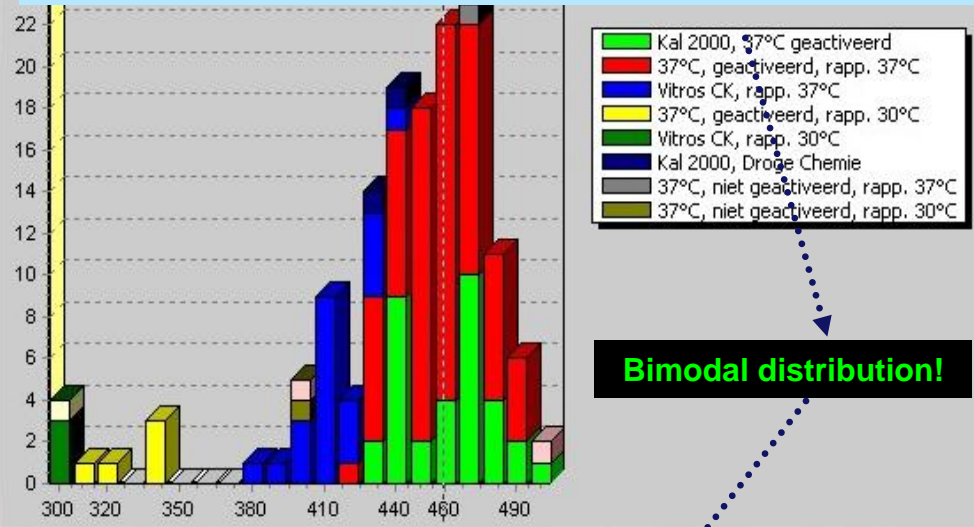Editorial Clin Chem, sept. 2013 < G Miller and G Myers:

"Commutability still matters"

1. Commutability of reference materials (RM) is an essential requirement to achieve comparability of patient results

2. It is required when a RM is used as calibrator **or as EQAS sample**.

Worldwide awareness of all stakeholders involved, among them:

- ✓ WHO recognizes importance of assessing commutability in WHO RMs

- ✓ JCTLM gives attention to commutability of RMs in Quality Manual of WG 1

- ✓ IFCC WG on Commutability established

- ✓ AACC International Consortium for Harmonization of Clinical Lab Results in place

- ✓ EQALM, EQAS organizations ……

Dutch EQA before 2005: processed controls

CK – all method groups, all labs

Legend:
- Kal 2000, 37°C geactiveerd
- 37°C, geactiveerd, rapp. 37°C
- Vitros CK, rapp. 37°C
- 37°C, geactiveerd, rapp. 30°C
- Vitros CK, rapp. 30°C
- Kal 2000, Droge Chemie
- 37°C, niet geactiveerd, rapp. 37°C
- 37°C, niet geactiveerd, rapp. 30°C

Bimodal distribution!

Non-commutable

Characteristics of former EQA-materials

- Non-human matrix/additives

- Lyophilized

- Sucrose as a stabilizer

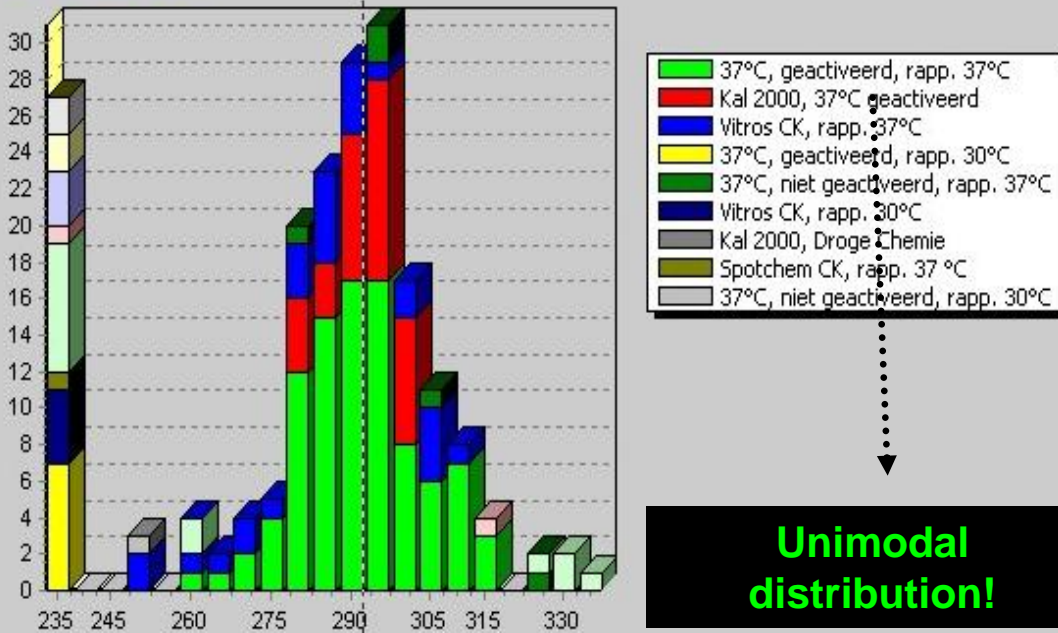- Matrix effects with dry chemistry assays

Implications/ intended use:

- Peer group comparisons

- Monitoring of intralab & interlab CVs

- No bias assessment

- No monitoring of IFCC std. efforts

LUMC

CK - all method groups, all labs

Potentially commutable

CK, Monster H

Legend:
- 37°C, geactiveerd, rapp. 37°C
- Kal 2000, 37°C geactiveerd
- Vitros CK, rapp. 37°C
- 37°C, geactiveerd, rapp. 30°C
- 37°C, niet geactiveerd, rapp. 37°C
- Vitros CK, rapp. 30°C
- Kal 2000, Droge Chemie
- Spotchem CK, rapp. 37 °C
- 37°C, niet geactiveerd, rapp. 30°C

**Unimodal distribution!**

Implications/ intended use:

- Trueness verification Monitoring of IFCC std.

- Monitoring of intralab & interlab CVs

Dutch EQA: <u>fresh frozen</u> controls for general clinical chemistry since 2005

1. Be clear about the measurand(s) intended to be measured

2. Human serum matrix

3. Not / minimally processed:
   - ✓ CLSI C37A single donor pools for lipids/apo's!
   - ✓ Regular pool procedure for chemistry with spiking; cave: human recombinant enzymes

4. Systematic concentration range (donor selection /spiking)

5. 24 interdependent samples per year (12 pairs; linear relation!)

6. Liquid frozen

7. Stored at - 70 ºC (enzymes!)

8. Value assigned with JCTLM-listed RMPs

Holy Grail
San Greal
Sang Real

LU MC

skml
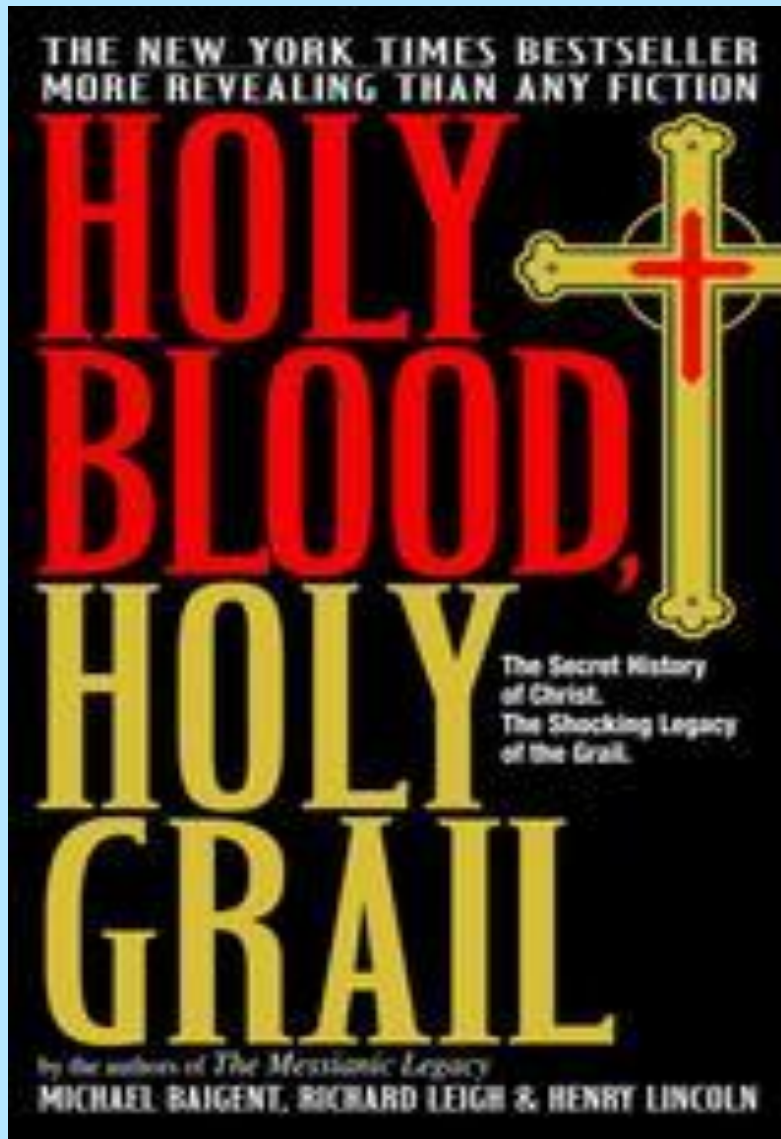
Calibration 2000

1. Commutable EQA materials

2. **Value assigned for trueness verification / temporary recalibration**

3. (Scoring system based on biological variation and clinical relevance)

Introduced in the Netherlands since 2005.

Clin Chim Acta 2012;414: 234-40

Daily language:

- Ability of a reference or control material to show **interassay properties similar** to those of human samples.

CLSI EP30-A (formerly C53-A) definition:

- The **<u>equivalence</u>** of the **mathematical relationship** among the results of different measurement procedures for a reference material and for representative samples of the type intended to be measured.

VIM (JCGM 200: 2012, 3rd edition) definition:

- Property of a reference material, demonstrated by the **<u>closeness of agreement</u>** between the **relation** among the measurement results for a stated quantity in this material, obtained according to two given measurement procedures, and the relation obtained among the measurement results for other specified materials.

Commutability definitions: implications for evaluation?

- According to EP30-A:

  - ✓ testing the hypothesis of equivalence?

- According to VIM:

  - ✓ point estimate of the relevant quantity with confidence interval?

  - ✓ i.e. an absolute or relative difference between the result of a RM and the average result for patient samples with routine methods

Under discussion in the IFCC WG on Commutability

III. CLSI EP30-A (formerly C53-A; since 2010)

C53-A
Vol. 30 No. 12
Replaces C53-P
Vol. 28 No. 26

Characterization and Qualification of Commutable Reference Materials for Laboratory Medicine; Approved Guideline

SCOPE:

This guideline provides recommendations for the characterization, assessment of commutability, and assignment of analyte concentration or activity values to reference materials (RMs) that are used for calibration and trueness assessment of *in vitro* diagnostic medical devices. This includes materials such as the following:

- Certified reference materials (CRMs)

- Materials without a formal certificate but with the characteristics of a CRM and attached information sufficient for use in instrument calibration or trueness control (eg, external quality assessment [EQA] or proficiency testing [PT] materials used to assess trueness)
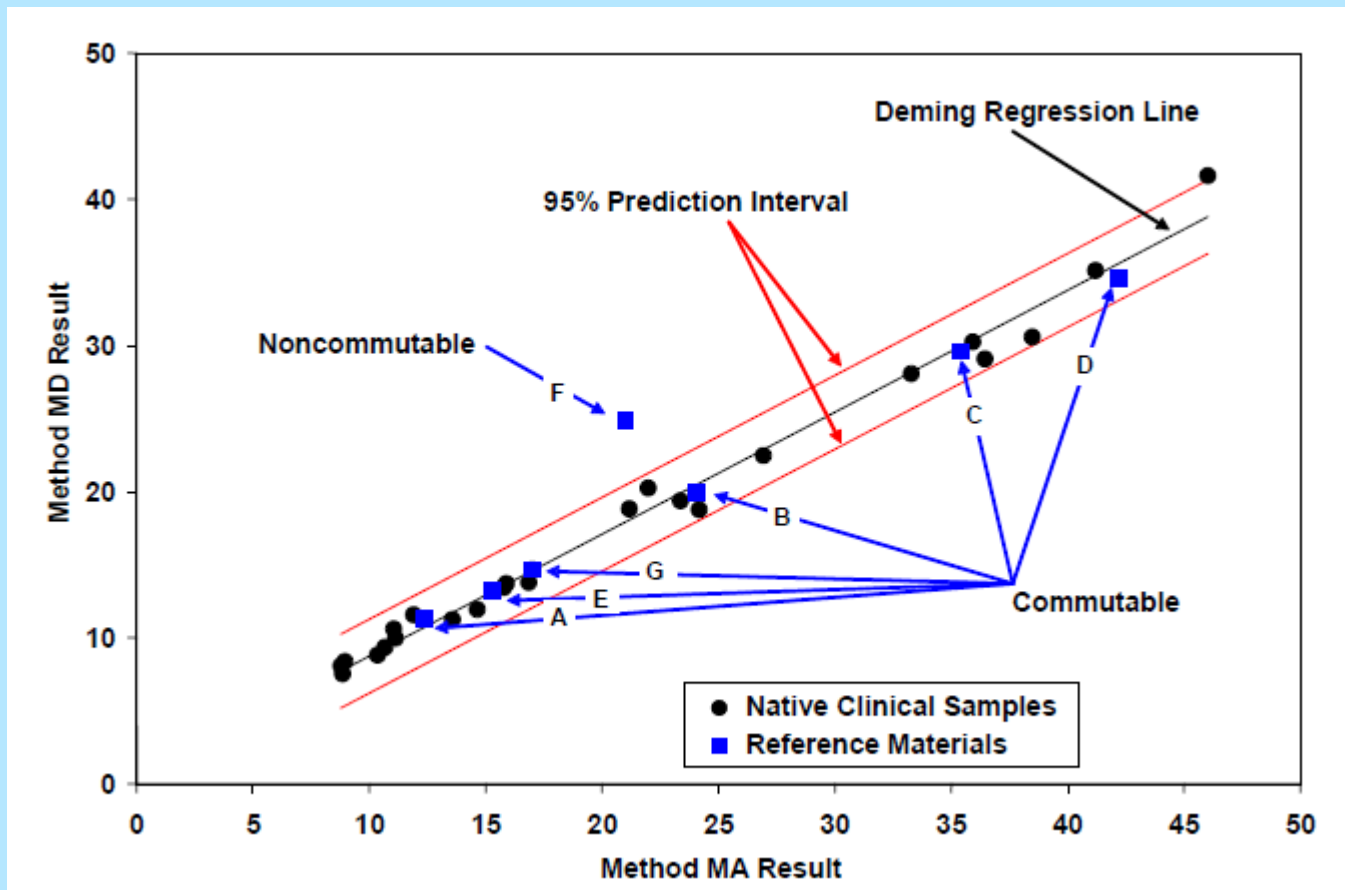
CLSI EP30-A – assessment of commutability

1. **Regression Approaches**

   - For 2-way comparison statistics, especially with RMPs

   - Regression analysis using the 95% prediction interval

   - Regression analysis using multiples of the standard error of regression (S_{y-x})

   - Objective: quantitative numeric values

2. **Multivariate Statistical Techniques**

   - If many patient and reference samples have to be examined with several measurement procedures;

   - Descriptive methods: PCA and correspondence analysis

   - Subjective

Use of the Regression Protocol and 95% Prediction Interval
to Evaluate Commutability of RMs between methods MA and MD.

CLSI EP30-A – assessment of commutability
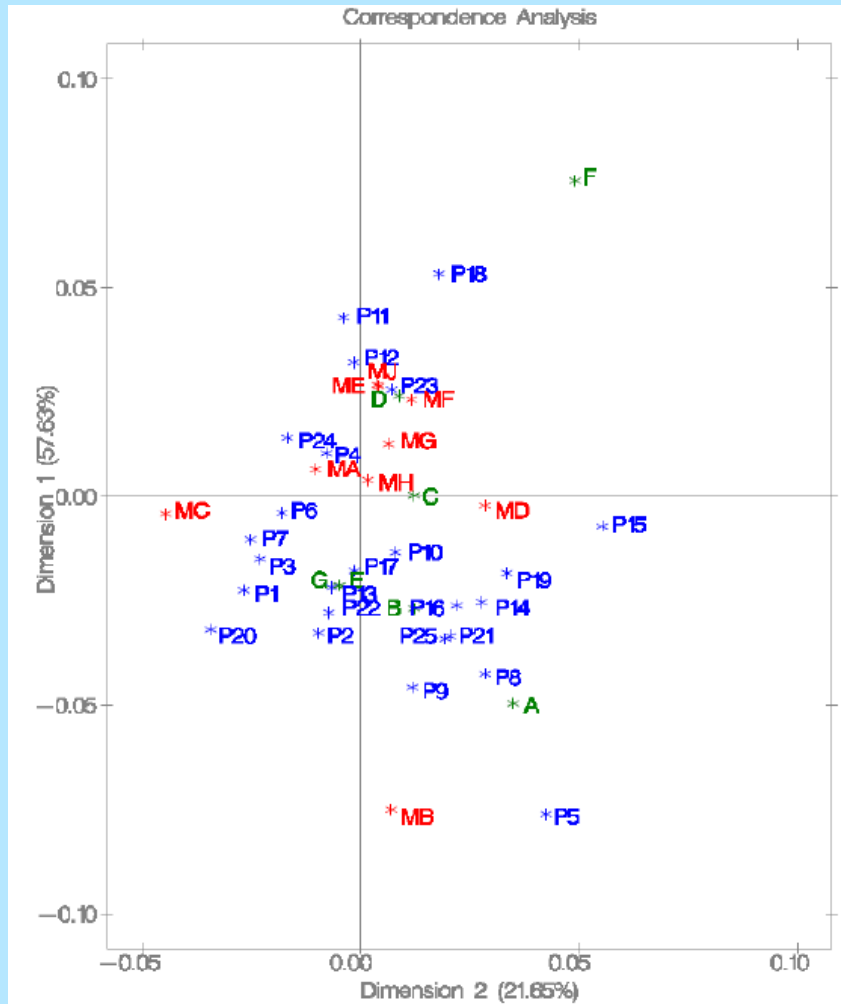
Projection of RMs near the center of the cluster defined by clinical samples represents a high degree of commutability.

Correspondence analysis of patient samples (P1-P25), reference materials (A-G) and measurement procedures (MA-MJ).

Criteria for acceptance of equivalence

- EP30-A: **equivalence** of the mathematical relationship of a RM to that of native samples is accepted when RM results are
    - within a region representing a probability to include 95% of patient results,
    - within $\pm$ 2 times S_{y-x} for the normalized residual procedure.

- **Acceptance depends on the estimate of dispersion in the numeric relationship observed for the native clinical samples**. Cave: influenced by number and type of clinical specimens; selectivity of routine methods; # within –run replicates!

- **The reliability of the conclusion depends on the robustness of the estimate** of the mathematical relationships among the measurement procedures for the native samples and for the RMs.

- *But….. shouldn't acceptance criteria be depended on the <u>intended use</u> of the commutable RM!?*

  - less stringent criteria for trueness controls (EQA) than for calibrators (IVD)?

  - linked to clinical requirements/needs ?

Commutability assessment: role of EQA organizers!?

- Evaluation of commutability:
    - shared responsibility of
        - Providers of RMs
        - Providers of clinical lab measurement procedures

- (Periodic) verification of commutability:
    - responsibility of EQA organizers

    Example: publication on the importance of commutability for ceruloplasmin std, from Zegers I et al., 2013

Results: example commutability assessment

> The recovery of ceruloplasmin in ERM-DA470 is reasonable with both the DAKO Hitachi and Beckman Immage methods (certified concentration 205 $\pm$ 11 mg/L)

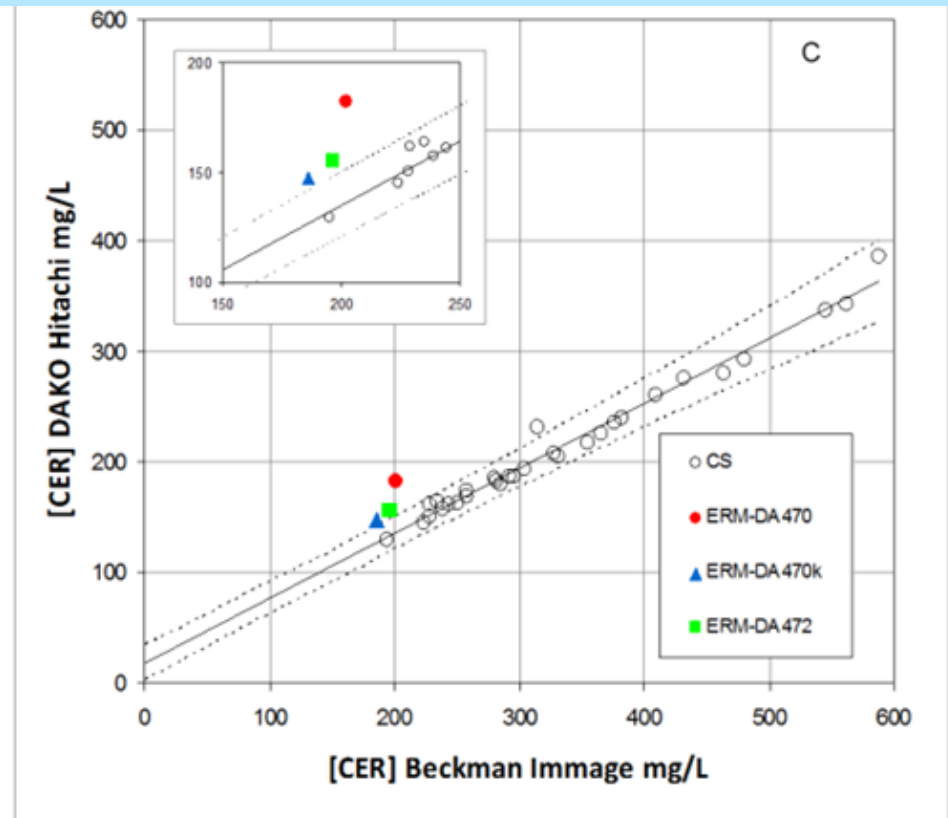> ERM-DA470 is not commutable for this combination of methods

Figure 3. Example of a pair-wise commutability assessment for ceruloplasmin (CER) measurements in sera from 30 healthy individuals (CS), ERM-DA470, ERM-DA470k, and ERM-DA472. The plotted values are the means per sample. The continuous black line is the results of a Deming regression, the dotted black lines correspond to the limits of the prediction interval at the 95% confidence level.

IV. Dutch experiences with commutability assessment of EQA-controls

Para-meters	Pools used as basis for EQA-controls	Clinical specimens	Representative routine labs / methods/ mfrs	Other RMs?	Acceptance criterion
Lipids and apo's; N = 6	C37-A: •Native (3) •Lyo (3) •Fro (3)	Fresh (12)	42 lab pairs	-	≤ 3 SD_{SA} state-of-the-art within-lab SD
General serum chemistry analytes; N = 18	Regular preparation procedure + spiking + •Lyo (2) •Fro (9)	fresh frozen (75)	6 labs	-	≤ 3 S_{y-x} normalized residuals
Serum enzymes; N = 7	Regular preparation procedure + spiking (12)	fresh frozen (40)	4 labs	IRMM Asahi-Kasei	95-105% recovery of ALTM

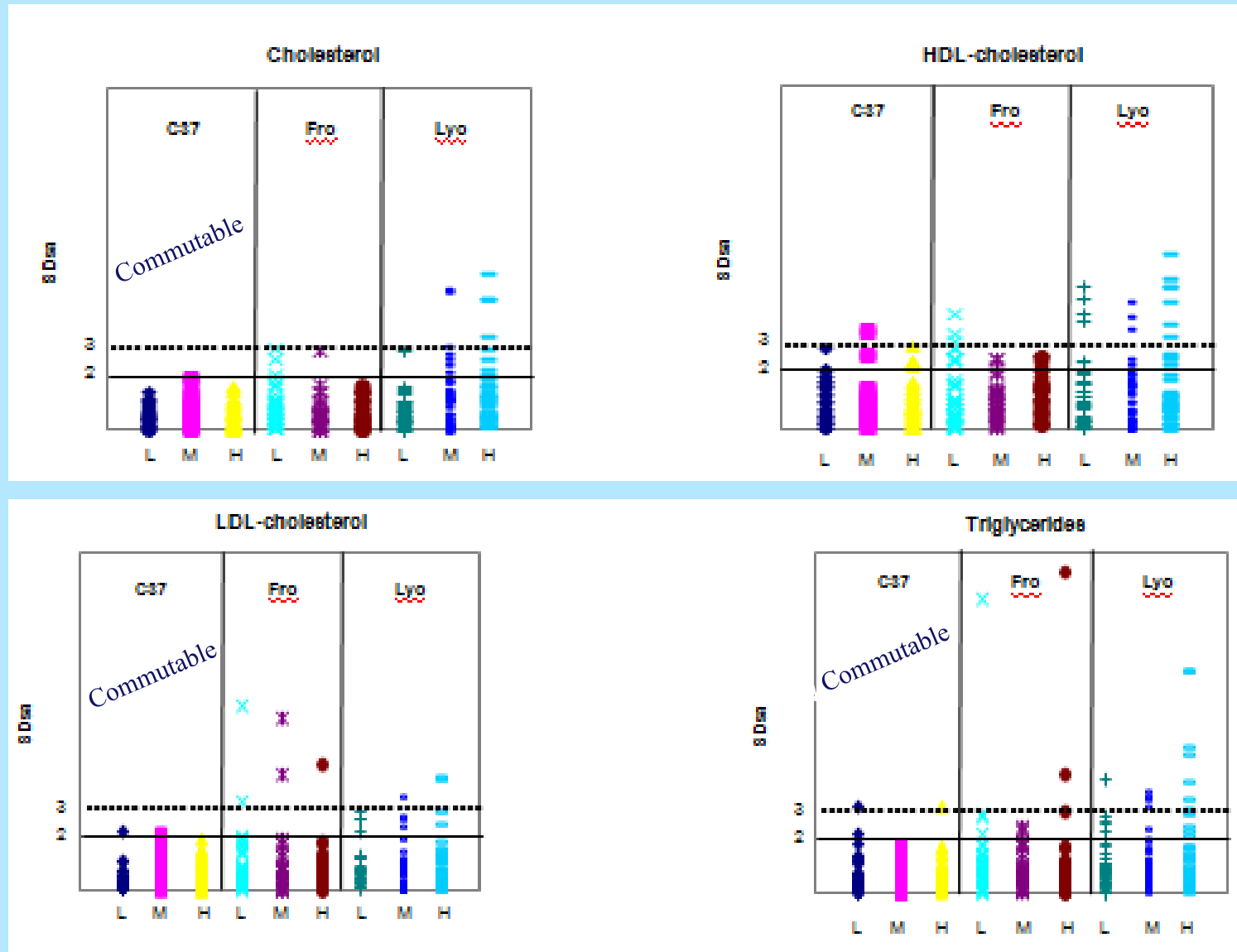**Twin-study design** for serum lipids/apolipoproteins:

- 84 labs => 42 lab couples

- 12 fresh patient sera

- 9 CRMs: CLSI C37-A based single donorpools (3 native; 3 lyo; 3 frozen)

 - Native specimens = no spiking

  - Low, medium, high HDLc

P/B regression analysis:

- Perpendicular distances of CRMs to patient regression line are expressed in multiples of *state-of-the-art* within-laboratory SD

- **$\underline{3}$ SD$_{SA}$ acceptance criterion**

Commutability assessment using a Twin-study design - serum lipids

Commutability assessment using a Twin-study design - serum apolipoproteins
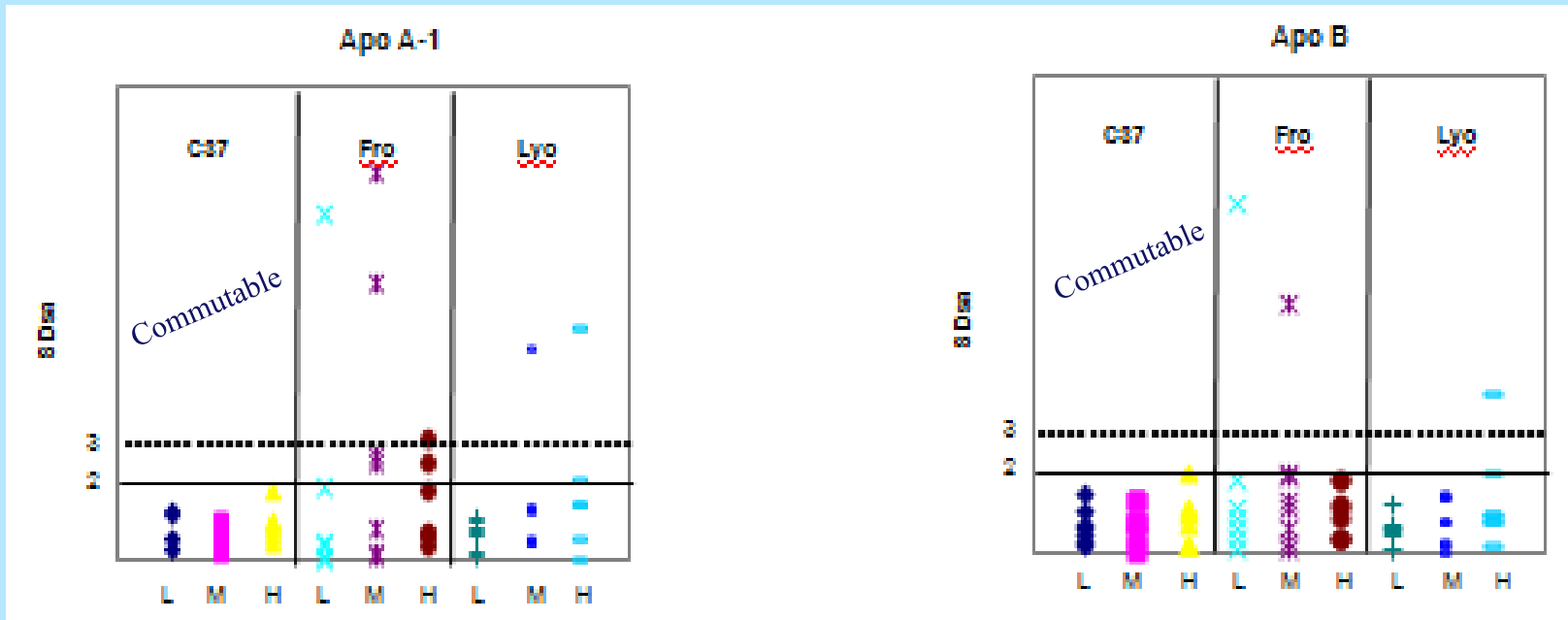
Conclusions

1. Only CLSI C37-A prepared pools are commutable with all tested assays for lipids and apolipoproteins, except for HDLc → introduced since 2005.
2. The twin-study design involving 84 labs is logistically very demanding.

**Twin-study** focusing on 18 general clinical chemistry analytes:

- 6 labs - major manufacturers/ methods

- 75 *frozen* patient sera (selected!)

- 13 CRMs; regularly pooled, processed, spiked multi-parameter sera

    - 1-11: fresh frozen

    - 12-13: lyophilized

Linear regression analysis, lab-to-lab
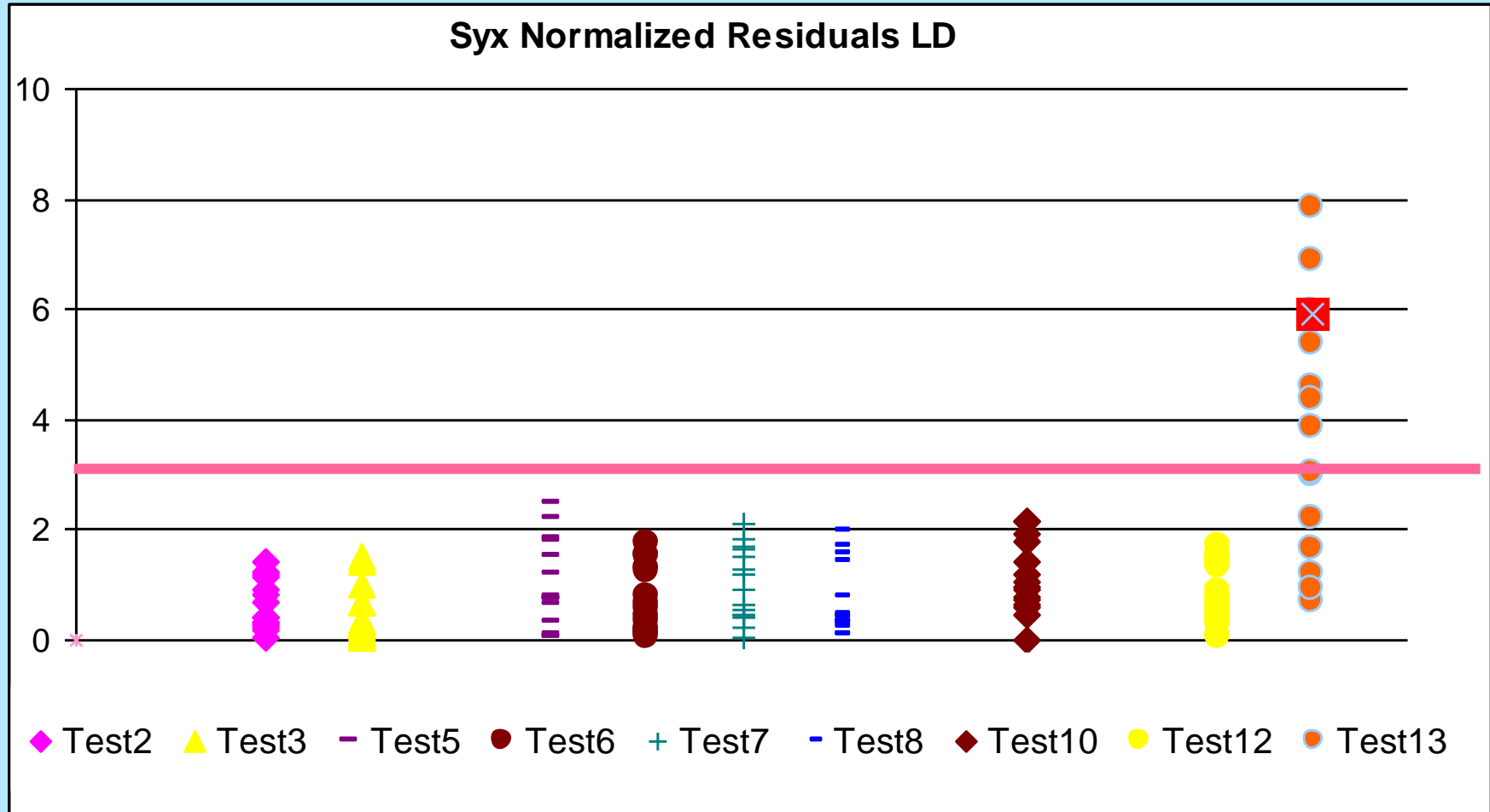
- 3 S_{y_x} acceptance criterion

Commutability assessment of 13 CRMs using a Twin-study design - serum LD
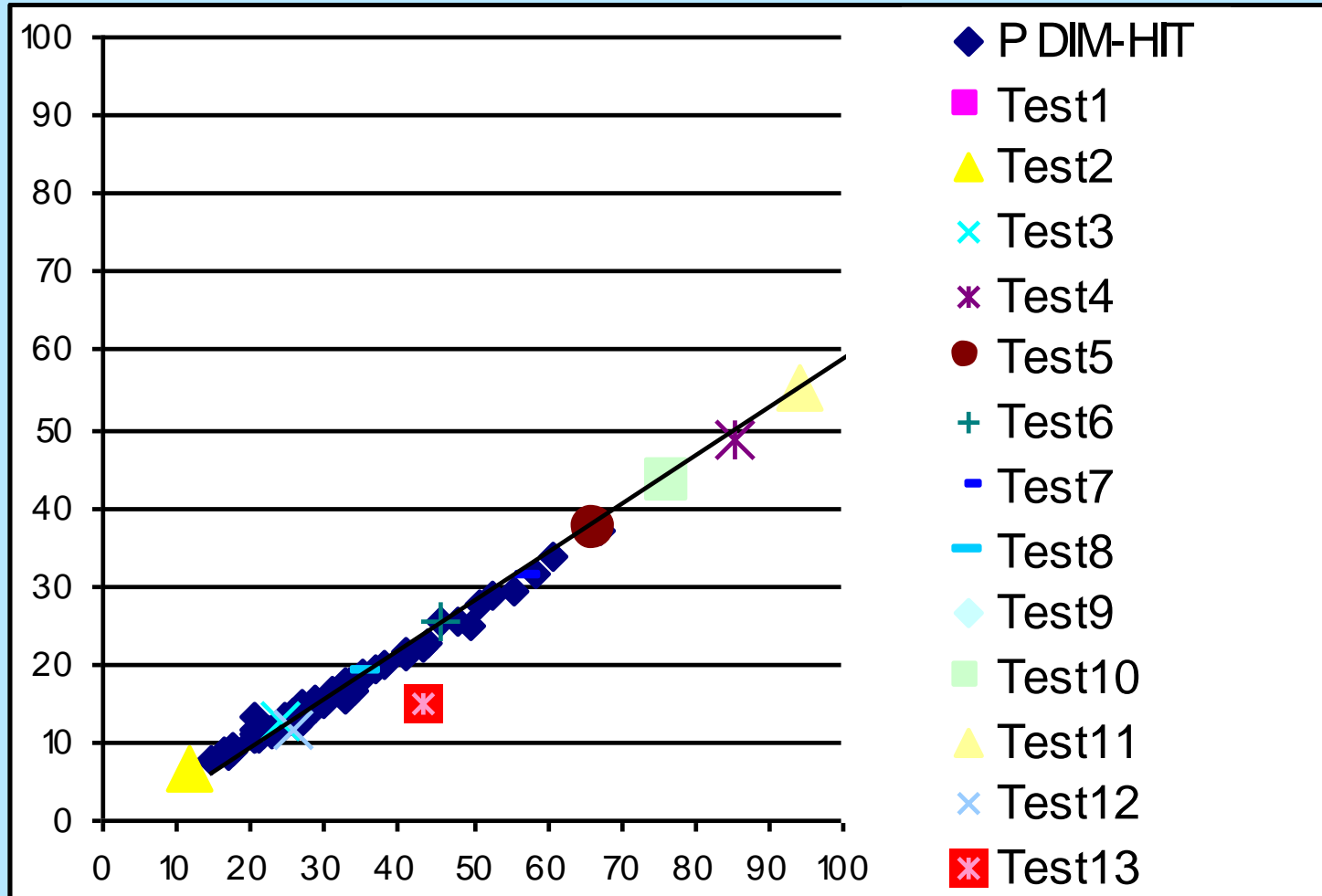

Commutability evaluation for LD using a TWIN STUDY DESIGN. **Between lab couples using different analytical systems, behaviour of test EQA-samples is compared to behaviour of patient samples and to the patient regression line. In this figure Dimension and Hitachi 917 labs are compared. NTKCL, 2008; 33: 154-7.**

Commutability assessment of 13 CRMs using a Twin-study design - serum LD

Syx Normalized Residuals LD

Legend: Test2, Test3, Test5, Test6, Test7, Test8, Test10, Test12, Test13

Syx normalized residuals of tested EQA-samples for all twin pair labs involved. NTKCL, 2008; 33: 154-7

Commutability assessment of 13 CRMs using a Twin-study design - serum ALAT
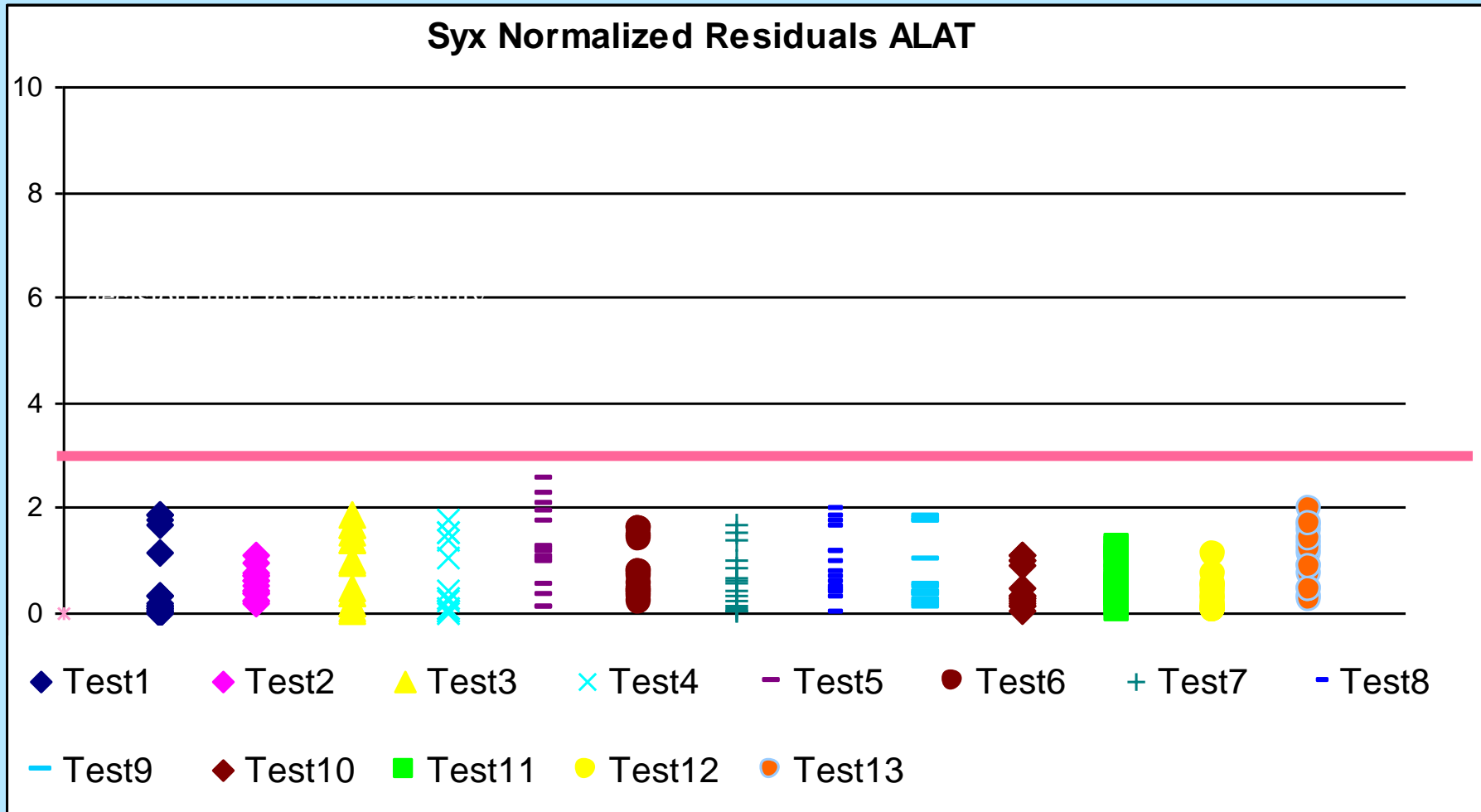
Syx normalized residuals of tested EQA-samples for all lab couples involved.
NTKCL, 2008; 33: 154-7.

Commutability assessment of 13 CRMs and 18 tests using a Twin-study design – aggregated data

Fresh frozen CRMs 1-11

Lyophilized CRMs 12-13
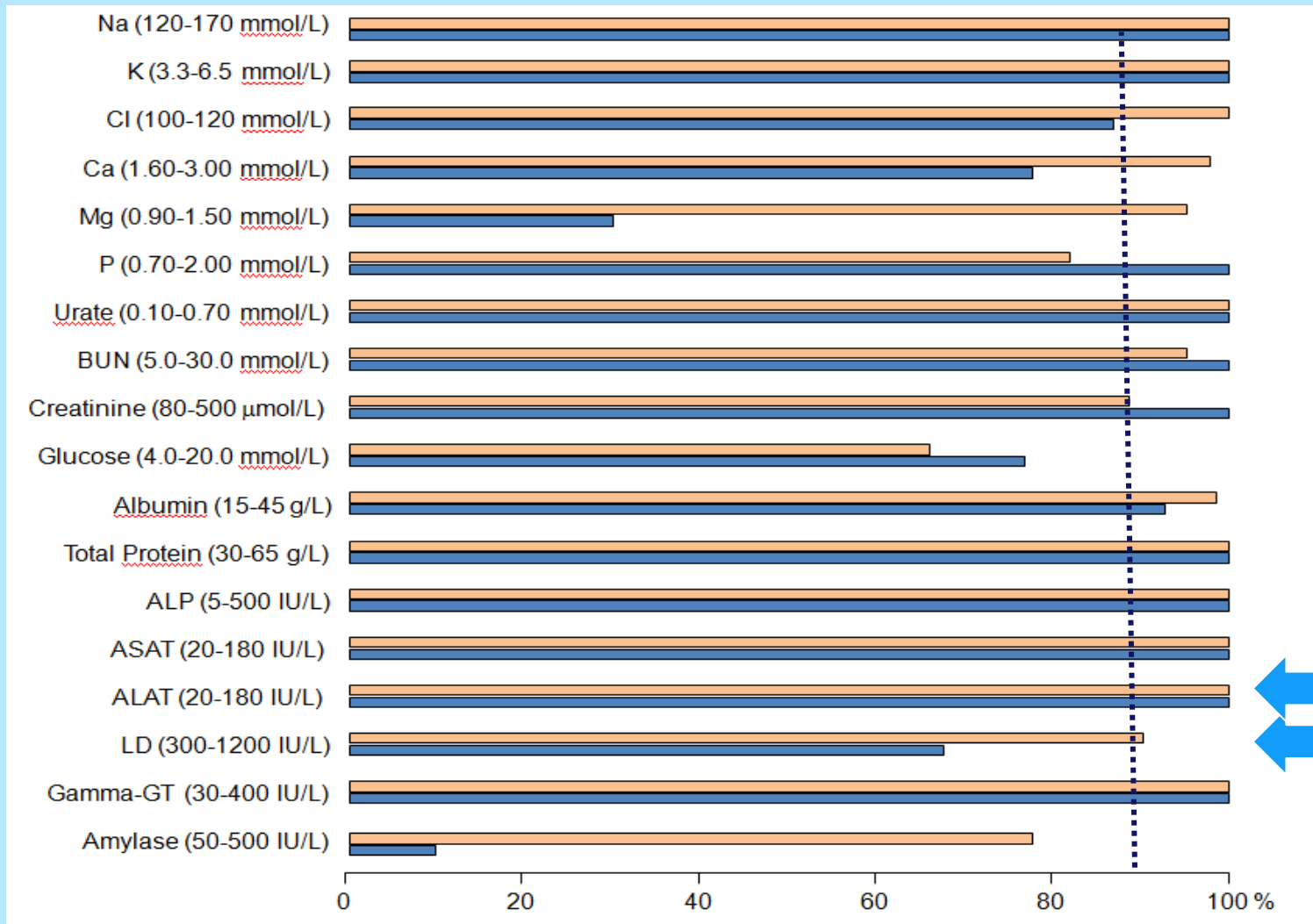
Percentage normalized residuals of fresh frozen and lyophilized CRMs within the 3 Syx interval for 18 clinical chemistry analytes.

- Fresh frozen EQA-controls:
  - 100% commutable with the tested routine methods for half (9 out of 18) of the parameters;
  - > 90% commutable for 15 out of 18 parameters.

- Lyophilized EQA-controls:
  - 100% commutable with the tested routine methods for 11 out of 18 parameters.
  - > 90% commutable for 12 out of 18 parameters

→ Fresh frozen is preferred above lyophilized EQA.

- **Discussion points**:

  - Is this degree of commutability **good enough for trueness verification** in EQA-surveys?
    - Multiparameter, multilevel RMs are mandatory, and hence some minor processing.
    - Shouldn't we knock out unselective methods?
  - How do we make a **sustainable approach**, encompassing periodic reevaluation in case of new batches RM and/or new/additional parameters?

Overall conclusions
on commutability assessment of EQA-control material

- **How to develop EQA-controls that are expected to be commutable?**
  Challenging considering the multiparameter/multilevel needs.
    - C37-A pool procedure and donor selection for lipids/apo's;
    - proprietary procedure with spiking for general clinical chemistry tests.

- **Twin-study (lab-lab or lab-ALTM) design & regression statistics:**
    - Acceptance criteria: less strict than EP30-A ($\leq 3SD_{SA}$, $\leq 3S_{y-x}$) and/or pragmatic (95-105% recovery) because of the intended use
    - Logistically manageable and affordable: 84 labs → 4-6 labs

- **Feasiblity of integrated commutability and trueness verification studies** by combined analyses of old & new EQA-control batches in time.

• Learning curve

• Under debate:
  • type and number of healthy subjects and/or patient samples needed
  • analytical specificity of routine methods
  • number of replicates (within-run)?
  • sound statistical approach related to intended use, i.e. with acceptance criteria for commutability of EQA-controls

  • ….