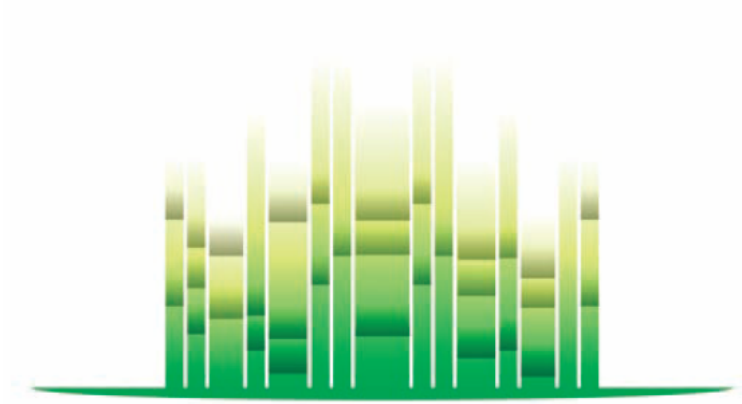
GUIDE TO THE EVALUATION OF COMMUTABILITY OF CONTROL MATERIALS



Ferruccio Ceriotti

Servizio di Medicina di Laboratorio, Ospedale San Raffaele, Milano

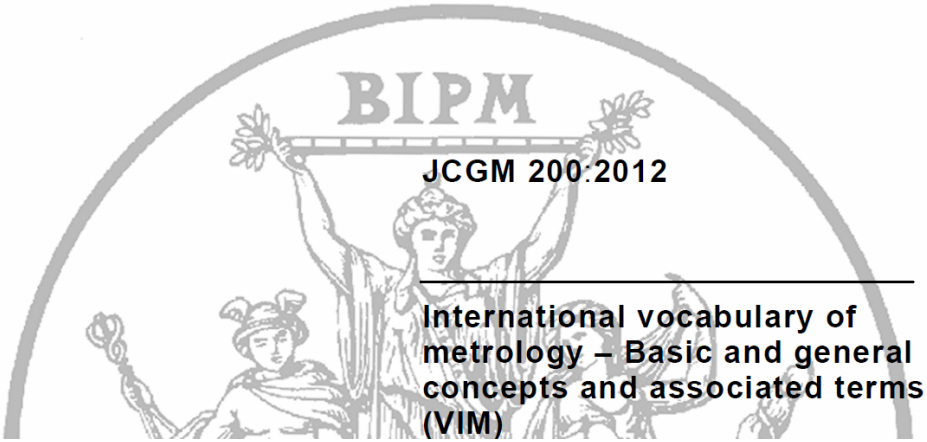
Preamble

- Most of the considerations and the data presented hereafter derive from the work of the

IFCC Working Group on Commutability (WG-C)

Coordinated by Greg Miller

The statistical work was performed by Göran Nilsson



5.15 commutability of a reference material

Property of a **reference material**, demonstrated by the closeness of agreement 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.

NOTE 1 The reference material in question is usually a **calibrator** and the other specified materials are usually routine samples.

NOTE 2 The measurement procedures referred to in the definition are the one preceding and the one following the reference material (calibrator) in question in a **calibration hierarchy** (see ISO 17511).

OSPEDALE SAN RAFFAELE

Commutability

- According to the definition, commutability concerns two measurement procedures. In laboratory medicine there are usually more than two measurement procedures for a measurand; consequently commutability needs to be assessed pair wise for all measurement procedures for which a RM is intended to be used.
- Commutability depends upon material – method interaction
- Non-commutability is an undesired byproduct of materials preparation combined with the specificity limitations of some clinical laboratory methods.

Consequences of non-commutability

Introduction of a BIAS

1. In the calibration process:
   - **break in the traceability chain**
2. In EQA schemes:
   - non reliable inter-method variability evaluation
   - **non feasibility of target value assignment based on Reference measurement procedure**

Assessment of commutability

Assessment of commutability requires the following steps:

- ✓Obtain representative clinical samples (CSs)
- ✓Obtain RM(s) to be evaluated
- ✓Measure the CSs and RM(s) with the measurement procedures for which the RM(s) are intended to be used
- ✓Determine the difference between the measurement procedure results for the RM(s) and the CSs
- ✓Determine if this difference is acceptable for the intended use of the RM(s)

1. Selecting clinical samples for inclusion in a commutability assessment

- The interval of quantities (e.g. concentrations) of the measurand in the CSs must include that of the RM(s).
- The interval of quantities must be within the dynamic range of both methods
- Clinical samples should be selected with consideration of measurement procedure selectivity limitations
  - a commutability assessment is not intended to evaluate the selectivity of measurement procedures for the measurand;
  - qualification of measurement procedures to be included in a commutability study should be done in advance.

1. Selecting clinical samples for inclusion in a commutability assessment (2)

- Individual CSs are preferred for a commutability assessment.
  - Pooling may affect commutability and represents a potential limitation, but volume limitations for difficult to obtain samples may require pooling;
  - If pools are used, a validation scheme to demonstrate commutability of a pool based on recovery of a value expected from the proportion of each donor sample used to prepare the pool should be performed (CLSI C37-A).
- Clinical samples must be collected, and aliquots prepared, stored and distributed such that no alteration of the measurand or matrix occurs.
  - Freezing or other storage may affect commutability and should be evaluated for suitability in a preliminary experiment.

2. Qualification of measurement procedures for inclusion in a commutability assessment

- It is desirable to include as many different analytical measurement principles as possible
- Measurement procedures must have acceptable performance characteristics to be included in a commutability assessment.
  - adequate precision, because an excessively large SD can inappropriately influence assessment of commutability
  - adequate selectivity for the measurand
  - residual around the relationship should be comparable with that of the other candidate measurement procedures.
  - linear relationship with other measurement procedures

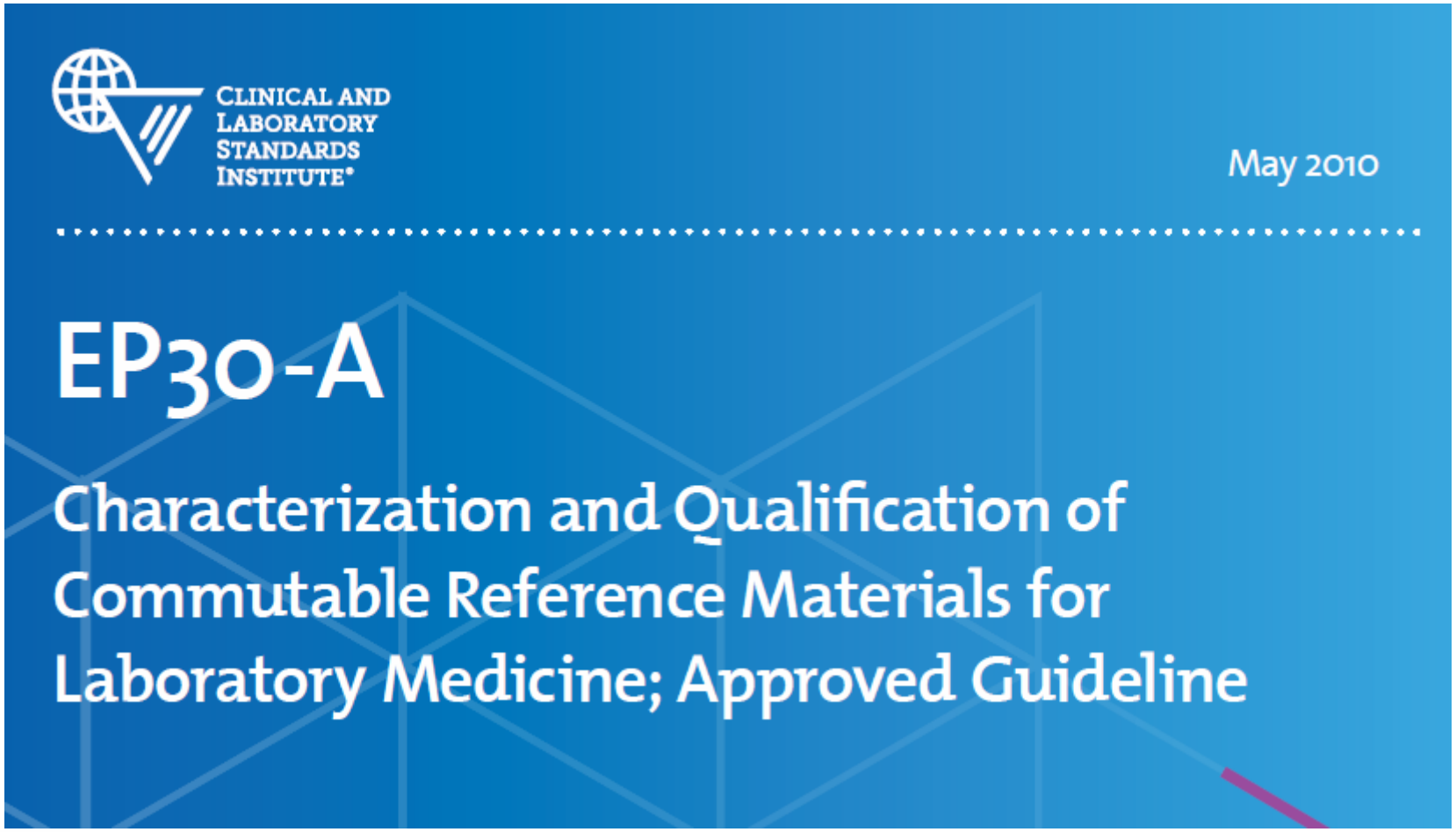
Possible approaches to assess commutability

1. Regression approach

2. Based on the difference in bias between a RM and CSs

3. Effectiveness of a RM used as a calibrator to improve harmonization among results from different measurement procedures

Regression approach



OSPEDALE SAN RAFFAELE

Regression approach: statistical design

- A linear relationship between measurement procedure A and B is assumed.
- The relationship between the results obtained with these two measurement procedures on *n* native clinical samples can be expressed by the equation:

$$y = a + bx$$

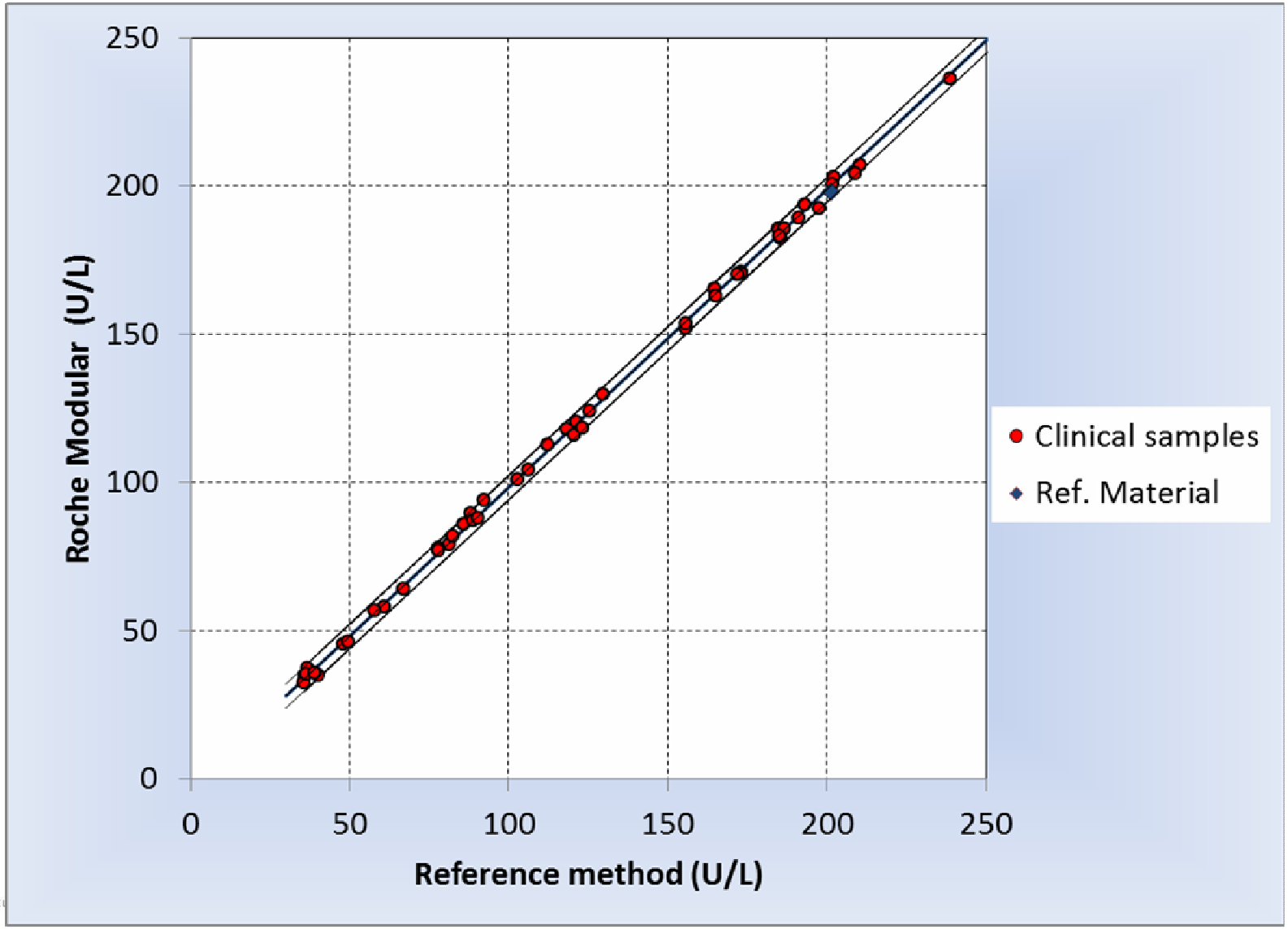
- A 95% prediction interval around the line described by the equation can be calculated.
- Commutability of a RM is assumed when its Y_R, X_R point falls inside the prediction interval

Regression approach: experimental design

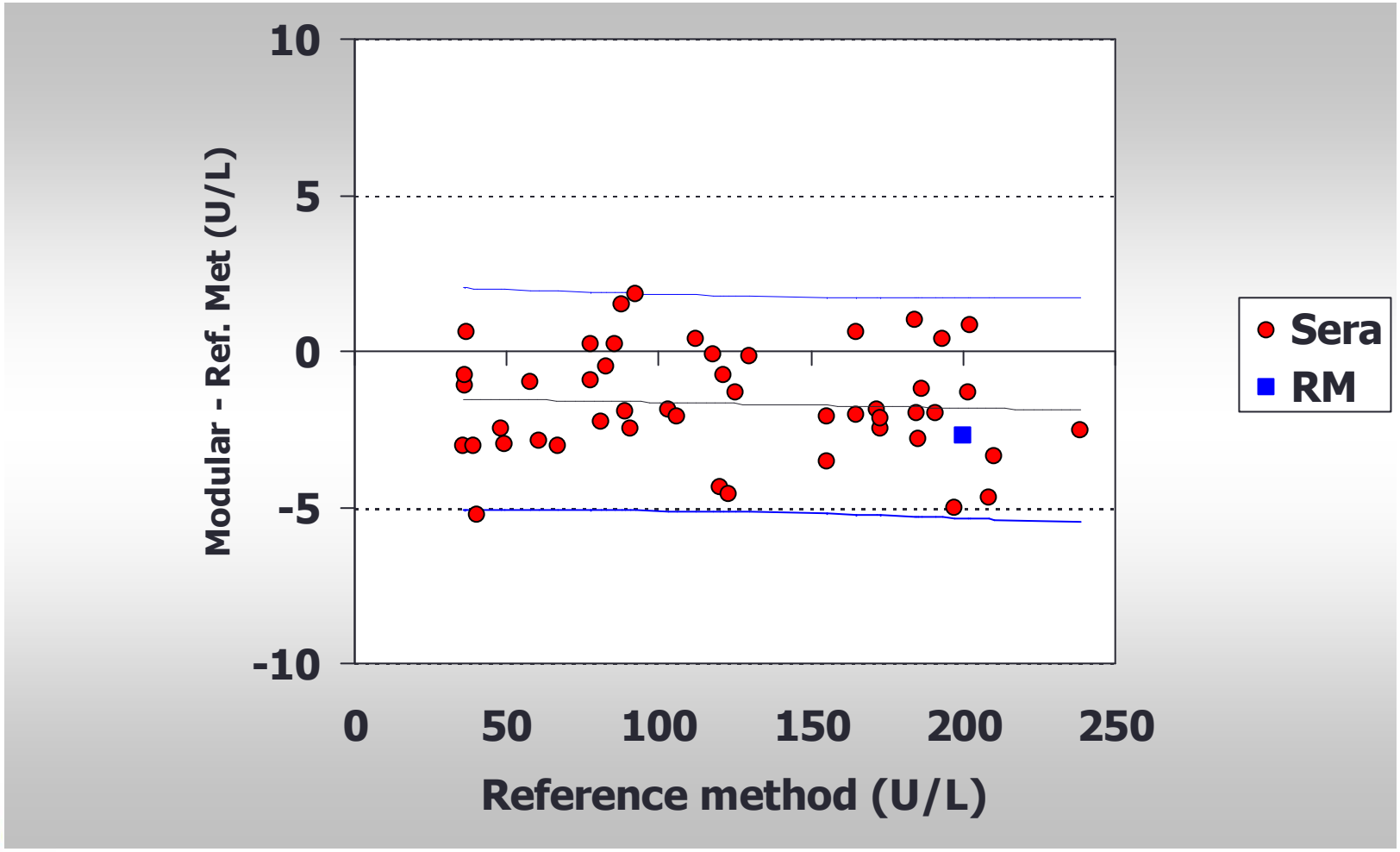
- Select at least 20 clinical samples (CSs),
- Analyze them in triplicate in a single run with both methods together with the RM also in triplicate.
- Calculate the regression equation according to Deming and the 95% prediction interval using only the CSs data.
- Plot the data on a graph and verify if the RM falls inside the prediction interval

- [alternatively, instead of an x – y plot a bias plot can be used]

AST: method comparison (x-y plot)



AST: method comparison (bias plot)



OSPEDALE SAN RAFFAELE

Regression approach

- Limits:
  - ✓ using a prediction interval for assessment of commutability is a test of the hypothesis that the RM can be considered to belong to the same population as the CSs.
  - ✓ Not rejecting a hypothesis does not prove that it is true
  - ✓ The prediction interval approach does not quantify how closely the RM agrees with the average relationship for the CSs at the concentration of interest.

Possible approaches to assess commutability

1. Regression approach

2. **Based on the difference in bias between a RM and CSs**

3. Effectiveness of a RM used as a calibrator to improve harmonization among results from different measurement procedures

Difference in bias between a RM and CSs: statistical design

- Commutability of the RM concerns how close the systematic difference between the measurement procedures (the bias) for the RM is to the average bias for the CSs, at the concentration of the RM.
- The difference between the bias for the RM and the average bias for the CSs (D_{RM}) expresses the closeness of agreement between the bias for the RM and the bias for the CSs.
- For the assessment of commutability we need to specify a maximum value of $|D_{RM}|$ for the RM to be considered as commutable (commutability criterion = **C**).

Difference in bias between a RM and CSs: statistical design

1. The uncertainty interval $D_{RM} \pm U(D_{RM})$ is within $0 \pm C$: the RM is commutable.
2. The uncertainty interval $D_{RM} \pm U(D_{RM})$ is outside $0 \pm C$: the RM is non-commutable.
3. The uncertainty interval $D_{RM} \pm U(D_{RM})$ and $0 \pm C$ are overlapping: the result is inconclusive.

An inconclusive result can be caused by:

- A too stringent commutability criterion.
- A too large uncertainty. When $U(D_{RM}) > C$ it will obviously not be possible to verify commutability. The experimental design of a commutability assessment must give a sufficiently small value of $U(D_{RM})$.

Difference in bias between a RM and CSs: Experimental design

- *n* CSs, “evenly” distributed in the interval that includes the concentration of the RM
- CSs and the RM are measured in three replicates, in one run, with both of the measurement procedures.
- The RM is measured in triplicates in *p* different positions in the run.

The example in the figure includes 50 CSs and 5 RMs, each one measured 5 times in triplicate

Calculations

- Bias RM (B_{RM}) = yRM – xRM
- $u(B_{RM}) = \sqrt{\frac{SD^2_{mean(x)} + SD^2_{mean(y)}}{p}}$

where p = number of means (suggested = 5)

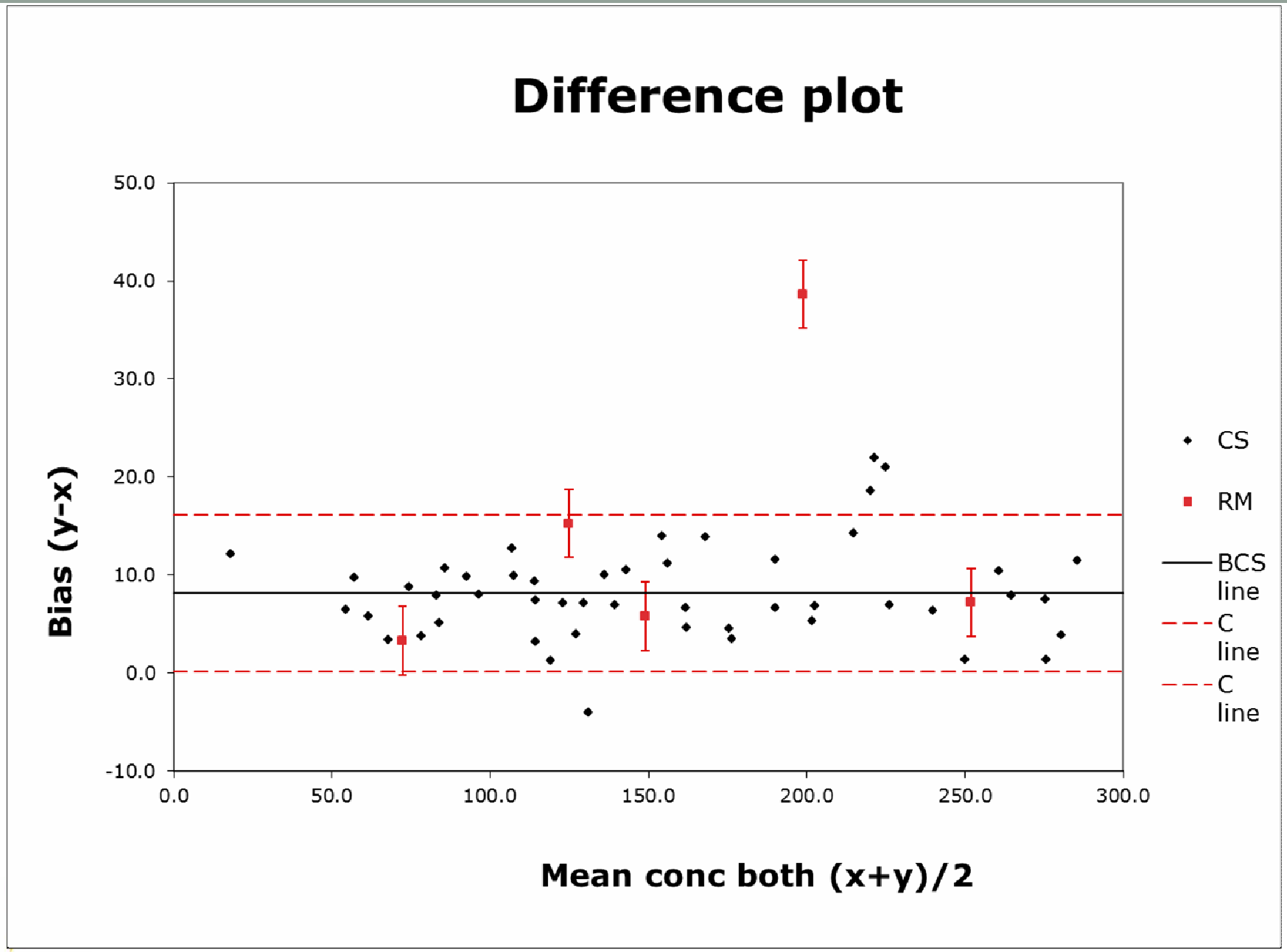
- Bias Clinical samples (B_{CS}) = mean of the biases (y-x)
- $u(B_{CS}) = \frac{SD_{Bias}}{\sqrt{n}}$

where n = number of CSs (suggested = 40)

$$D_{RM} = B_{RM} - B_{CS}$$

$$u(D_{RM}) = \sqrt{\frac{SD^2_{mean(x)} + SD^2_{mean(y)}}{p} + \frac{SD^2_{Bias}}{n}}$$

OSPEDALE SAN RAFFAELE


Difference plot
Bias (y-x)
50.0
40.0
30.0
20.0
10.0
0.0
-10.0
0.0
50.0
100.0
150.0
200.0
250.0
300.0
Mean conc both (x+y)/2
CS
RM
BCS line
C line
C line

Possible approaches to assess commutability

1. Regression approach

2. Based on the difference in bias between a RM and CSs

3. **Effectiveness of a RM used as a calibrator to improve harmonization among results from different measurement procedures**

Effectiveness of a RM used as a calibrator to improve harmonization among results from different measurement procedures: statistical design

- If a reference material is commutable, its use as a calibrator will allow different methods to produce comparable results on clinical samples.
- The commutability is assumed if the MP to MP clinical sample result variability is reduced to a specified limit (e.g. CV≤3%). In addition there may be requirements for EQA/PT variability (e.g. CV≤5%).
- It is necessary to identify if an MP or individual samples need to be excluded from the set of clinical samples and MPs to be used in the RM commutability assessment.

Effectiveness of a RM used as a calibrator to improve harmonization among results from different measurement procedures: experimental design

- Use the RM as calibrator in a reference measurement procedure and in all the measurement procedures in which is requested to test its commutability.
- Measure 40 CSs in triplicate
- Calculate the average of the inter method CV estimates, “Inter Measurement Procedure CV” (IMPCV), and compare to the predetermined variability requirements.
- Inspect the data to verify if some sample or some MP has peculiar behavior and eventually discard them and recalculate the CV

Example

Nr.	Target	M6	M1	M7	M2	M4	M3	M5	Mean	SD	CV
1	30	27,4	27	29,0	30,5	33,6	30,6	29	29,6	2,2	7,52
2	34	27,4	35	34,8	34,6	33,6	34,7	34	33,4	2,7	8,08
3	37	32,6	38	36,8	37,6	36,6	39,8	37	36,9	2,2	5,97
4	41	36,0	42	39,7	41,7	41,7	42,9	42	40,9	2,4	5,75
5	43	36,0	43	39,7	42,7	42,7	43,9	43	41,6	2,8	6,72
6	44	37,7	44	41,6	44,7	44,7	43,9	44	43,0	2,5	5,92
7	47	39,4	46	48,4	47,8	48,8	44,9	46	45,9	3,2	6,95
8	48	42,0	47	48,4	48,8	48,8	47,0	48	47,1	2,4	5,08
9	49	41,1	51	50,3	49,8	47,8	50,0	49	48,4	3,4	6,98
10	50	42,9	50	48,4	50,8	51,9	51,1	51	49,4	3,1	6,27
11	52	44,6	52	48,4	52,9	52,9	56,2	50	51,0	3,7	7,34
12	52	46,3	52	50,3	53,9	52,9	53,1	52	51,5	2,6	
13	55	48,9	58	48,4	55,9	52,9	55,1	57	53,7	3,8	
14	57	48,9	60	55,2	58,0	55,9	59,2	56	56,2	3,7	
15	58	48,9	59	56,1	60,0	59,0	61,3	58	57,5	4,1	7,17
16	58	49,7	57	61,0	59,0	60,0	57,2	58	57,4	3,7	6,42
17	61	51,4	43	54,2	66,1	71,2	74,6	39	57,1	13,9	24,27
38	105	89,1	102	105,5	106,8	110,8	108,3	93	102,2	8,1	7,97
39	112	97,7	110	105,5	113,9	116,9	118,5	97	108,5	8,8	8,07
40	123	103,7	124	111,3	123,1	124,1	122,6	101			8,78
	Mean 1-40	58,0	67,4	65,3	68,8	69,1	69,0	65,4			7,3

Higher CV

IMPCV

Example

Negative bias

Variablity of the residuals

Nr.	Target	M6	M1	M7	M2	M4	M3	M5
1	30	-3	-3	-1	1	4	1	-1
2	34	-7	1	1	1	0	1	0
3	37	-4	1	0	1	0	3	0
4	41	-5	1	-1	1	1	2	1
5	43	-7	0	-3	0	0	1	0
6	44	-6	0	-2	1	1	0	0
7	47	-8	-1	1	1	2	-2	-1
8	48	-6	-1	0	1	1	-1	0
9	49	-8	2	1	1	-1	1	0
10	50	-7	0	-2	1	2	1	1
11	52	-7	0	-4	1	1	4	-2
12	52	-6	0	-2	2	1	1	0
13	55	-6	3	-7	1	-2	0	2
14	57	-8	3	-2	1	-1	2	-1
15	58	-9	1	-2	2	1	3	0
16	58	-8	-1	3	1	2	-1	0
17	61	-10	-18	-7	5	10	14	-22
18	62	-11	-1	-2	0	2	1	1
19	62	-9	-2	-8	1	3	3	0
20	64	-9	-1	-2	0	1	-2	2
21	66	-9	0	4	1	2	1	-1

OSPEDALE SAN RAFFAELE

Example

Nr.	Target	M6	M1	M7	M2	M4	M3	M5	Mean	SD	CV
1	30		27		30,5	33,6	30,6	29	30,1	2,4	8,00
2	34		35		34,6	33,6	34,7	34	34,4	0,6	1,70
3	37		38		37,6	36,6	39,8	37	37,8	1,2	3,30
4	41		42		41,7	41,7	42,9	42	42,1	0,5	1,17
5	43		43		42,7	42,7	43,9	43	43,1	0,5	1,15
6	44		44		44,7	44,7	43,9	44	44,3	0,4	0,96
7	47		46		47,8	48,8	44,9	46	46,7	1,6	3,34
8	48		47		48,8	48,8	47,0	48	47,9	0,9	1,91
9	49		51		49,8	47,8	50,0	49	49,5	1,2	2,43
10	50		50		50,8	51,9	51,1	51	51,0	0,7	1,30
11	52		52		52,9	52,9	56,2	50	52,8	2,2	4,22
12	52		52		53,9	52,9	53,1	52	52,8	0,8	1,52
13	55		58		55,9	52,9	55,1	57	55,8	2,0	3,50
14	57		60		58,0	55,9	59,2	56	57,8	1,8	3,19
15	58		59		60,0	59,0	61,3	58	59,5	1,2	2,09
16	58		57		59,0	60,0	57,2	58	58,2	1,3	2,16
39	112		110		113,9	116,9	118,5		114,8	3,7	3,26
40	123		124		123,1	124,1	122,6		123,4	0,7	0,60
	Mean 1-40	#####	67,5	#####	68,9	69,0	68,9	62,6	68,5	1,8	2,6

OSPEDALE SAN RAFFAELE

Limits

- Better applicable to RM intended to be used as calibrators
- Assumes a calibration function with no intercept
- Requires multiple methods (but this is usually the case)
- Statistical requirements not yet defined in detail, especially regarding how to perform calibration.

The “ratio” method to confirm commutability of a new batch of commutable RM

- **Rationale**: if a first batch of RM is assessed and demonstrated to be commutable, the following batches, provided that are manufactured according to the same process, do not need a new extensive evaluation of commutability, but commutability can be confirmed with a simpler method.
- **Experimental design**: the new and old RM batches are measured with N replicates with the relevant measurement procedures (MP) including a reference MP (if available) in the same analytical run for each MP.
- The ratio (mean from new batch)/(mean from old batch) is calculated for each MP. When ratios for the respective MPs fall within the pre-established range, it is concluded that the new batch of RM has the same commutability properties as the old batch.

OSPEDALE SAN RAFFAELE

Conclusions

- All the presented approaches require a large number of measurements to reach an high statistical power (low uncertainty), so their applicability to large commutability studies is questionable.
- The use of simplified designs based on a small number of pools in place of individual CSs and has still to be evaluated and validated.
- The “ratio method” is still under statistical evaluation, but appears a valid solution for commutability confirmation and maybe not only for confirmation.


Thanks for your attention!