Recommendations for assessing commutability part 3: based on the calibration effectiveness of a reference material

Supplemental Data

Table S1. Example set of data from seven measurement procedures (MPs) across 40 clinical samples (CS ID 1 to 40) plus each clinical sample's trimmed mean.

CS ID	MP1	MP2	MP3	MP4	MP5	MP6	MP7	CS Trimmed Mean
1	31	30	36	49	23	24	26	29.4
2	34	39	43	54	23	27	31	34.8
3	38	42	46	58	25	31	33	38.0
4	40	47	49	65	29	34	39	41.8
5	44	46	49	63	29	34	38	42.2
6	46	51	60	71	34	35	41	46.6
7	49	48	82	67	31	34	40	47.6
8	49	52	60	72	34	37	43	48.2
9	47	56	62	74	33	39	44	49.6
10	48	55	60	75	36	40	46	49.8
11	51	57	60	78	36	44	45	51.4
12	58	57	62	79	36	42	47	53.2
13	57	64	60	82	36	43	51	55.0
14	57	47	67	95	49	58	35	55.6

CS ID	MP1	MP2	MP3	MP4	MP5	MP6	MP7	CS Trimmed Mean
15	55	66	68	84	39	46	50	57.0
16	57	63	76	85	41	45	52	58.6
17	63	65	70	87	41	48	52	59.6
18	62	66	67	91	45	51	56	60.4
19	59	67	74	89	44	50	57	61.4
20	70	69	77	92	45	49	59	64.8
21	65	73	86	96	47	53	59	67.2
22	69	74	79	100	50	54	62	67.6
23	70	76	79	98	47	53	61	67.8
24	65	78	82	102	49	53	66	68.8
25	69	80	90	105	50	60	66	73.0
26	73	74	94	106	55	63	65	73.8
27	73	86	89	111	54	53	71	74.6
28	80	79	97	108	44	57	67	76.0
29	72	84	97	106	50	60	68	76.2
30	81	88	97	115	60	65	68	79.8
31	78	95	92	118	56	63	73	80.2
32	83	87	91	117	58	67	74	80.4
33	83	95	108	121	59	70	77	86.6
34	85	98	108	123	60	69	74	86.8
35	87	99	119	127	66	76	78	91.8
36	94	117	115	145	70	85	81	98.4

CS ID	MP1	MP2	MP3	MP4	MP5	MP6	MP7	CS Trimmed Mean
37	101	107	119	136	69	85	82	98.8
38	104	112	131	147	76	85	84	103.2
39	109	121	131	156	81	93	87	108.2
40	119	130	138	160	83	96	88	114.2

Table S2. Example set of data from the seven measurement procedures (MPs) after recalibration using reference material with two levels plus the initial clinical sample (CS) trimmed mean from Table S1.

CS ID	MP1C	MP2C	MP3C	MP4C	MP5C	MP6C	MP7C	Initial Trimmed Mean
1	30.8	27.7	30.0	31.9	33.9	23.8	28.1	29.4
2	33.8	35.9	35.8	35.4	33.9	26.8	33.8	34.8
3	37.8	38.6	38.3	38.1	36.8	30.7	36.0	38.0
4	39.8	43.1	40.8	43.0	42.4	33.6	42.9	41.8
5	43.8	42.2	40.8	41.6	42.4	33.6	41.8	42.2
6	45.8	46.7	49.8	47.2	49.4	34.5	45.2	46.6
7	48.8	44.0	67.8	44.4	45.2	33.6	44.1	47.6
8	48.8	47.6	49.8	47.9	49.4	36.4	47.6	48.2
9	46.8	51.2	51.5	49.3	48.0	38.4	48.7	49.6
10	47.8	50.3	49.8	50.0	52.2	39.3	51.1	49.8
11	50.8	52.1	49.8	52.2	52.2	43.1	49.9	51.4
12	57.9	52.1	51.5	52.9	52.2	41.2	52.3	53.2
13	56.9	58.4	49.8	55.1	52.2	42.2	57.0	55.0
14	56.9	43.1	55.5	64.6	70.0	56.4	38.3	55.6
15	54.9	60.2	56.4	56.5	56.4	45.0	55.8	57.0
16	56.9	57.5	62.9	57.2	59.1	44.1	58.2	58.6
17	63.0	59.3	58.0	58.7	59.1	46.9	58.2	59.6
18	61.9	60.2	55.5	61.6	64.6	49.8	63.1	60.4
19	58.9	61.0	61.3	60.2	63.2	48.8	64.3	61.4
20	70.0	62.8	63.7	62.4	64.6	47.9	66.7	64.8

21	65.0	66.4	71.0	65.3	67.3	51.7	66.7	67.2
22	69.0	67.3	65.3	68.3	71.4	52.6	70.4	67.6
23	70.0	69.0	65.3	66.8	67.3	51.7	69.2	67.8
24	65.0	70.8	67.8	69.8	70.0	51.7	75.4	68.8
25	69.0	72.6	74.3	72.1	71.4	58.2	75.4	73.0
26	73.1	67.3	77.5	72.9	78.1	61.0	74.1	73.8
27	73.1	77.8	73.4	76.7	76.7	51.7	81.7	74.6
28	80.2	71.7	79.9	74.4	63.2	55.4	76.6	76.0
29	72.1	76.1	79.9	72.9	71.4	58.2	77.9	76.2
30	81.2	79.6	79.9	79.8	84.7	62.9	77.9	79.8
31	78.2	85.7	75.9	82.1	79.4	61.0	84.3	80.2
32	83.3	78.7	75.1	81.3	82.1	64.7	85.6	80.4
33	83.3	85.7	88.8	84.5	83.4	67.5	89.4	86.6
34	85.3	88.3	88.8	86.1	84.7	66.6	85.6	86.8
35	87.3	89.2	97.6	89.2	92.6	73.0	90.7	91.8
36	94.5	104.8	94.4	104.0	97.8	81.2	94.7	98.4
37	101.7	96.1	97.6	96.5	96.5	81.2	96.0	98.8
38	104.8	100.5	107.2	105.7	105.5	81.2	98.7	103.2
39	109.9	108.2	107.2	113.3	111.8	88.4	102.7	108.2
40	120.2	115.9	112.8	116.8	114.4	91.1	104.0	114.2

Four-parameter logistic curve (4PLC) model for calibration

The considerations used for the simple case of a linear regression also apply when using a more complex model such as 4PLC:

$$y = d + \frac{a-d}{1+(x_s/c)^b} + \varepsilon \quad (1)$$

where y is the signal, x_s is concentration value, ε is the random variation of the signal, a is the minimum signal value at zero concentration, d is the maximum signal value at $+\infty$ concentration, c is the inflection concentration point of the logistic curve and b is the slope of signal versus concentration at concentration c. Accounting for the sources of systematic error and solving for x_i, the following equation would give a correct CS value:

$$x_i = \frac{c}{g_i \cdot h_i} \cdot \left(\frac{a-d}{y_i \cdot f_i - d} - 1\right)^{1/b} + e_i \quad (2)$$

where e_i is the random variation of concentration values. This random error is comprised of the same components mentioned in the main text.

For such a calibration curve, the function f could be much more complex than for a linear regression fit. The theoretical basis for using such a calibration model has been described based on the kinetics of immunoassays (1). Using this model, the function f deviates from unity when kinetic theory does not match the assay's biochemical behavior.

Beyond this limitation however, the other systematic error terms are all in the concentration realm (after calibration) so their implications would be the same as in the linear example. Therefore in using the calibration curve to estimate concentration the following equation would be used, thus ignoring the systematic error terms as above:

$$\hat{x}_i = c \cdot \left(\frac{a-d}{y_i - d} - 1\right)^{1/b} + e_i \quad (3)$$

The random error of concentration results e_i at any point on the curve (x_i, y_i) is, as above for the linear regression example, the signal random error ε_i at that same point multiplied by the calibration slope (dx/dy) at that point. Solving for the derivative of the equation above and assuming a symmetric distribution

$$[within\ calibration:]e_i = \left|\frac{\varepsilon_i c(d-a)((a-y_i)/(y_i-d))^{1/b}}{b(a-y_i)(y_i-d)}\right| \quad (4)$$

This equation typically results in a non-constant random error over the concentration range that is best characterized by a precision profile[1].

Computation of the Quasi-range

The standard deviation is a measure of dispersion that is highly influenced by outlying data points. Given that specific specimens may demonstrate such outlying behavior for selected MPs, a more robust measure such as the quasi-range

$$W_{(i)} = X_{(n+1-i)} - X_{(i)}$$

may be more appropriate as a measure of dispersion where X is the ordered set of n results for an MP. For data sets with $n \geq 32$ the best quasi-range is $W_{(3)}$ (2). In the above example (where n = 40), the results from MP1 display a relatively normal distribution. Therefore, the ratio of SD/W for MP1 is used to scale $W_{(3)}$ for all the other MP results.

References for Supplemental Data:

1. Rodbard D, Feldman Y. Kinetics of two-site immunoradiometric (‘sandwich’) assays – 1: Mathematical models for simulation, optimization and curve fitting. Int J Molec Immun 1978;15:71-6.

2. Cadwell JH. The distribution of quasi-ranges in samples from a normal population. Ann Math Statist 1953;24:603-13.