Taylor&Francis
healthsciences

Scand J Clin Lab Invest 2004; 64: 399–406

Effect of analytical quality on establishing common reference intervals and their use

P. RUSTAD* & P. HYLTOFT PETERSEN†

*Fürst Medical Laboratory, Oslo, Norway; †Odense University Hospital, Odense, Denmark

Rustad P, Hyltoft Petersen P. Effect of analytical quality on establishing common reference intervals and their use. Scand J Clin Lab Invest 2004; 64: 399–406.

In the Nordic Reference Interval Project (NORIP), reference intervals were established for 25 common clinical biochemical quantities. In the project, samples from more than 3000 reference individuals collected in the 102 participating laboratories from all five Nordic countries were analysed locally. In order to maintain a high level of analytical quality and to document this quality, a common calibrator/reference preparation (CAL) and a number of control samples were analysed together with the reference samples. All these materials were serum pools of unprocessed serum from many donors in order to obtain commutable materials. The CAL was used to harmonize the many different analytical procedures and calibrations by simple recalibration by the factor T/M where T is the target value based on reference methods and M is the mean of 10 replicate measurements of CAL in each laboratory. The analytical quality specifications (analytical goals) were based on specifications created directly for the purpose of sharing common reference intervals and only the bias criteria were used because bias is the dominating problem in transfer of reference intervals. These specifications were different for the evaluation of reference values to create common reference intervals and for the laboratories to use these common reference intervals (when established). An interesting outcome was that it was only for the biologically well-regulated quantities serum-sodium and serum-calcium that the selection of the best laboratories gave considerably narrower reference intervals.

Key words: Quality goal; reference limit; reference value

Pål Rustad, Fürst Medical Laboratory, Søren Bulls vei 25, NO-1051 Oslo, Norway. E-mail. prustad@furst.no

INTRODUCTION

In the recommendations of the International Federation of Clinical Chemistry (IFCC) for establishing reference intervals in the late 1980s [1–3], the focus was on each single laboratory establishing its own reference intervals. Since then, the concept of sharing common reference intervals across analytical measurement procedures and calibrations has been investigated in the Nordic countries [4, 5] and within a country and same type of equipment and calibration from the same producer [6, 7]. The theoretical basis for evaluation of the analytical quality

DOI 10.1080/00365510410006315

needed to share common reference intervals (besides all the biological prerequisites) was established in 1988 [8] for Gaussian and in 1989 [9] for log-Gaussian distributions. This is based on what is considered to be an acceptable fraction of reference values that will fall outside the reference limits if a laboratory intends to use the common reference interval. This concept was later improved by defining three levels of analytical quality (optimum, desirable and minimum quality) for acceptable bias [10].

In the Nordic Reference Interval Project 102 laboratories were invited to collect blood containing heparin and no additives, to produce serum and plasma samples from at least 25 reference individuals evenly distributed for gender and age, and at least 50% fasting. The samples were stored at $-80^{\circ}C$ until analysis. The laboratories received 5 reference materials CAL, X, P, HIGH and LOW (called controls) on dry ice to be measured together with the reference samples on 25 of the most frequently used quantities in clinical biochemistry. The data were submitted to the project group for calculation of reference intervals.

As the concept of the project was to use a variety of routine methods for analysis, it was important to have highly commutable controls with reliable target values, i.e. certified values, if possible. The "control" CAL, a serum pool from men not processed other than sterile filtering before freezing at $-80^{\circ}C$, was used as a project "calibrator" for the non-enzymes, i.e. all reference values were multiplied by the factor T/M where T is the target value of CAL and M is the mean value of 10 replicates of CAL measured together with the reference samples. This study was conducted in an early stage of the project. Therefore also the enzymes were treated as they were non-enzymes, i.e. reference values were corrected with the factor T/M as mentioned above.

After correcting reference samples according to this procedure, the rest of the control material was used to evaluate quality of the routine methods. Based on the analytical quality specifications for optimum bias, analytical results not fulfilling the goals were discarded and the reference intervals obtained with the best results were compared with the bulk of the calculated reference intervals.

Thus, the aim of this study was to examine the effect of removing inferior-quality series before calculating reference intervals.

MATERIALS AND METHODS

Analytical quality specifications (goals)

Based on the biological coefficient of variation for the population (CVb), the quality specifications for relative bias (B) have been proposed [10] as:

$$\text{Optimum}: \quad |B| < 0.125 \cdot CVb$$

$$\text{Desirable}: \quad |B| < 0.250 \cdot CVb$$

$$\text{Minimum}: \quad |B| < 0.375 \cdot CVb$$

As this evaluation was made during an early stage of the project, the CVb was estimated from reference intervals used in Malmö and Odense, according to $(\ln H - \ln L)/4$, where H and L are the upper and lower reference limits, respectively. This is an approximation assuming all distributions were log-normal. An optimal goal for bias would then be $|B| < 0.125 \cdot (\ln H - \ln L)/4$. This optimum quality for maximum bias was chosen in order to obtain the best possible reference intervals.

As all non-enzyme reference values were corrected, the values to be tested against quality goals are the corrected means of controls (or C/CAL, where C is control X or P) and HIGH/LOW for each laboratory. Therefore, in this part of the study there is no need for target values of CAL or to know the units of raw data, as only quotients between different control values are considered.

Analytical uncertainty concerning the analytical bias was added to the above-mentioned quality goal. The uncertainty of the fraction was added as $k \cdot CVa \cdot (i^{-1} + j^{-1})^{1/2}$, where CVa is median within series analytical coefficient of variation for CAL based on all series, i and j are the numbers of replicates of control C and CAL in the quotient for the laboratory (generally 3 and 10, respectively) for P/CAL and X/CAL and k is a coverage factor (2 for 95% confidence). The relative bias goal used in this

study could therefore be expressed as

$$|B_{Goal}| = 0.125 \cdot (\ln H - \ln L)/4 + 2 \cdot CVa \cdot \left(i^{-1} + j^{-1}\right)^{1/2}$$
$$\approx 0.125 \cdot CVb + 1.5 \cdot CVa \text{ used in the calculations.}$$

To estimate bias, a median of quotients is used as the target because the median is not influenced by extreme values.

For each laboratory B can be expressed:

for X as B(X) = X/CAL − median(X/CAL)
for P as B(P) = P/CAL − median(P/CAL)
for HIGH/LOW as B(HIGH/LOW) = HIGH/LOW − median(HIGH/LOW)

For HIGH/LOW, separate target values are estimated for wet and dry chemistry (Ortho, Vitros), as interferences for diluted samples are well known for dry chemistry.

All quotients X/CAL, P/CAL, HIGH/LOW should lie within quality goals to be included in the database to calculate the reference intervals.

RESULTS

The results are presented in two tables. The premises for calculation, i.e. target values, goals and number of series deleted after applying quality goals to each series, are presented in Table I.

In Table II we present the reference values for all and for each gender before and after removing series not fulfilling quality goals.

Pancreatic amylase is removed as no laboratories were excluded.

In Figure 1, the change in reference ranges (difference between upper and lower reference limits) for each component is quantified. The

TABLE I. Premises for exclusion of laboratories not fulfilling bias quality goals.

Component	Target High/Low	Target High/Low Ortho	Target P/CAL	Target X/CAL	Goal, opt.	CVa	Goal, opt. +1.5CVa	Sum	0	1	2	3
Albumin	1.931	2.4062	0.971	1.020	0.81%	1.42%	2.9%	110	5	7	47	51
ALP	1.961	1.7107	0.927	1.097	5.47%	1.44%	7.6%	96		3	4	89
ALT	2.135	1.4034	0.695	1.368	5.03%	4.53%	11.8%	107	1	11	19	76
AMY	2.015	2.0880	1.094	1.032	5.03%	2.25%	8.4%	74	2	1	18	53
AST	1.931	2.2105	0.761	1.075	5.03%	2.59%	8.9%	96	1	4	19	72
Bilirubin	2.424	2.3542	0.844	1.043	7.20%	4.74%	14.3%	110		5	26	79
Calcium	1.493	1.4590	1.013	1.022	0.52%	1.08%	2.1%	107	3	13	25	66
Carbamide	1.989	2.2130	0.916	1.020	4.70%	1.97%	7.7%	101		1	5	95
Cholesterol	1.998	2.0816	1.019	1.064	2.09%	1.37%	4.1%	109		3	8	98
CK	2.038	1.5917	0.667	1.120	5.03%	1.46%	7.2%	101	2	1	9	89
Creatininium	1.946	2.1718	0.987	1.048	1.60%	2.00%	4.6%	108	4	8	43	53
Glucose	2.000	2.1358	0.892	0.988	1.43%	1.16%	3.2%	97	1	5	15	76
GT	1.973	1.6095	0.658	0.984	8.66%	1.54%	11.0%	96	1	3	8	84
HDL-chol	1.940	2.5714	1.217	1.043	2.58%	1.67%	5.1%	104	1	3	18	82
Iron	1.997	2.5462	0.930	0.946	3.09%	1.51%	5.4%	93		3	9	81
LD	2.072	2.1932	1.009	1.113	3.76%	2.12%	6.9%	89	2	8	23	56
Magnesium	1.966	2.0617	1.003	1.013	1.41%	1.65%	3.9%	84	2	6	25	51
Phosphate	1.978	2.0340	1.056	1.013	1.93%	1.34%	3.9%	102	2	3	21	76
Potassium	2.000	1.9859	1.035	1.000	1.20%	0.85%	2.5%	108		6	29	73
Protein	1.962	1.9890	1.026	1.025	0.90%	1.26%	2.8%	80	1	2	18	59
Sodium	1.379	1.3882	1.029	1.025	0.22%	0.53%	1.0%	108	2	11	33	62
TIBC	2.037		1.115	1.013	1.47%	1.95%	4.4%	34		3	10	21
Triglyceride	1.981	2.0664	1.079	0.986	4.70%	1.34%	6.7%	106		3	10	93
Urate	2.010	2.0901	0.856	1.064	3.07%	1.22%	4.9%	105	1	2	12	90

ALP = alcalic phosphatase; ALT = alanine transaminase; Amy = amylase; AST = aspartate transaminase; LD = lactate dehydrogenase; GT = γ = glutamyltransferase; CK = creatine kinase; TIBC = total iron-binding capacity.

Explanations to columns:

Target values for High/Low for wet and dry chemistry (Ortho, Vitros), for P/CAL and X/CAL. "Goal, opt" is optimal bias goal as 0.125(lnH − lnL)/4 where H is upper and L lower reference limit (from Odense/Malmö). CVa is median within series analytical variation from NORIP. " Goal, opt + 1.5CVa " is the goal to compare each laboratory quotient.

Of the last 5 columns the first is "Sum" with number of analytical series evaluated, "0" is number of laboratories that does not fulfil quality criteria for any of the quotients X/CAL, P/CAL and HIGH/LOW etc. Last column show number of laboratories fulfilling the quality goal that all quotients should lie within bias limits.

TABLE II. Percentiles 50 (median), 2.5 and 97.5 for both (ALL) and each gender before and after removing laboratories not fulfilling defined quality goals.

	Male				Female				ALL				
Component	50	2.5	97.5	N	50	2.5	97.5	N	50	2.5	97.5	N	Rest
Albumin	41.9	36.2	47.4	1466	40.4	34.8	46.1	1623	41.0	35.2	46.8	3090	100%
	0.4	0.6	0.5	678	0.4	0.8	0.2	788	0.5	1.2	0.3	1466	47%
ALP	65	40	114	1265	58	33	107	1408	62	35	111	2674	100%
	0	0	−1	1145	0	0	2	1291	0	0	0	2437	91%
ALT	22	9	65	1413	16	7	42	1565	19	7	58	2980	100%
	1	1	−1	981	0	1	−2	1101	0	1	−4	2082	70%
Amylase	56	26	108	1056	57	24	104	1155	56	25	106	2212	100%
	0	1	1	728	0	3	−2	804	−1	3	0	1532	69%
AST	24	15	46	1292	21	13	37	1417	23	14	43	2711	100%
	0	0	−3	919	0	0	1	1032	0	0	−2	1952	72%
Bilirubin	12	6	29	1455	9	5	23	1618	10	5	25	3075	100%
	0	0	−4	1004	0	0	−1	1131	0	0	−1	2137	69%
Calcium	2.35	2.16	2.54	1421	2.32	2.14	2.53	1570	2.34	2.15	2.54	2993	100%
	0	0.03	−0.01	850	0.01	0.02	0.00	973	0.00	0.02	−0.01	1824	61%
Carbamide	5.4	3.4	8.2	1322	4.5	2.7	7.4	1486	4.9	2.9	7.9	2810	100%
	−0.1	−0.1	0.0	1257	0.0	0.0	0.0	1415	0.0	0.0	0.0	2674	95%
Cholesterol	5.2	3.2	7.3	1447	5.2	3.4	7.6	1607	5.2	3.3	7.4	3056	100%
	0.0	0.1	0.0	1278	0.0	0.0	0.1	1430	0.0	0.0	0.1	2710	89%
CK	121	50	480	1316	83	36	274	1450	98	40	391	2768	100%
	0	−1	1	1146	−2	0	2	1284	0	−1	7	2432	88%
Creatininium	79	64	101	1410	66	52	85	1555	72	53	97	2967	100%
	0	0	0	688	−1	0	−2	780	0	1	0	1470	50%
Glucose	5.0	4.0	6.6	1226	4.8	3.9	6.3	1349	4.9	3.9	6.5	2576	100%
	0.0	0.1	0.0	982	0.0	0.0	0.0	1103	0.0	0.1	0.0	2085	81%
GT	26	13	108	1301	18	10	71	1426	22	11	86	2728	100%
	0	0	−5	1137	0	0	1	1247	0	0	−3	2385	87%
HDL-chol.	1.3	0.9	2.1	1386	1.7	1.0	2.6	1540	1.5	0.9	2.5	2928	100%
	0.0	0.0	0.0	1085	0.0	0.0	0.0	1239	0.0	0.0	0.0	2324	79%
Iron	19.7	10.1	34.0	1195	18.1	7.8	33.3	1357	18.9	8.7	33.6	2554	100%
	0.0	0.0	0.3	1043	−0.1	0.0	0.2	1189	0.0	0.0	0.2	2233	87%
LD	169	117	241	1170	166	120	232	1308	167	118	236	2479	100%
	−2	1	−11	789	−2	−1	−4	860	−2	0	−7	1649	67%
Magnesium	0.84	0.72	0.97	1083	0.83	0.71	0.94	1254	0.84	0.71	0.95	2339	100%
	0.01	0.01	−0.01	666	0.00	0.00	0.00	790	0.00	0.00	0.00	1457	62%
Phosphate	1.08	0.74	1.55	1352	1.15	0.85	1.48	1496	1.12	0.78	1.50	2850	100%
	0.01	0.02	0.02	981	0.00	0.00	0.01	1076	0.01	0.00	0.01	2059	72%
Potassium	4.1	3.6	4.7	1413	4.1	3.6	4.6	1568	4.1	3.6	4.7	2983	100%
	0.0	0.1	0.0	928	0.0	0.0	0.0	1036	0.0	0.0	0.0	1965	66%
Protein	71	63	79	1127	69	62	77	1204	70	62	78	2332	100%
	0	−1	−1	859	0	0	0	929	0	0	0	1789	77%
Sodium	141	136	146	1417	141	135	145	1576	141.1	135.8	145.7	2995	100%
	0.1	1.4	−0.2	816	0.1	1.4	0.0	909	0.0	1.4	−0.4	1725	58%
TIBC	63	49	82	425	67	50	95	496	65	49	90	921	100%
	0	0	2	260	1	0	0	317	0	0	0	577	63%
Triglyceride	1.09	0.49	3.23	1404	0.91	0.44	2.3	1566	0.99	0.45	2.87	2972	100%
	0.00	0.00	−0.01	1224	0.03	0.01	0.09	1371	0.01	0.02	0.00	2596	87%
Urate	335	231	475	1367	247	155	373	1531	288	166	451	2901	100%
	0	0	6	1169	−1	1	7	1306	0	0	2	423	85%

ALP = alcalic phosphatase; ALT = alanine transaminase; AST = aspartate transaminase; LD = lactate dehydrogenase; GT = γ = glutamyltransferase; CK = creatine kinase; TIBC = total iron-binding capacity.

N is the number of reference values used for calculation and "Rest" is % of all reference values used for calculating reference intervals. For each component there are two rows; the first is data for all reference values, the second shows for 50, 2.5 and 97.5 the change from the row above when laboratories not fulfilling the goal are deleted.

trend is not surprising; for most of the components, the reference ranges become narrower after deleting inferior analytical series. The effect is relatively small, with sodium and calcium as possible exceptions (Table III).

DISCUSSION

A surprisingly large number of series were removed for many components, partly reflecting the optimal bias goal as the most demanding alternative and perhaps also the heterogeneous measurement systems used to produce reference values.

All reference values are first accurately corrected with a highly commutable material measured with 10 replicates; i.e., in principle, all laboratory bias should have been eliminated.

The higher the ratio between the biological variation and the analytical variation, the fewer series are deleted.

It is likely that the series not fulfilling the bias goal has biases that are both negative and positive and in that respect level each other out. The rest-effect would be an increase of variance widening the reference intervals. The difference

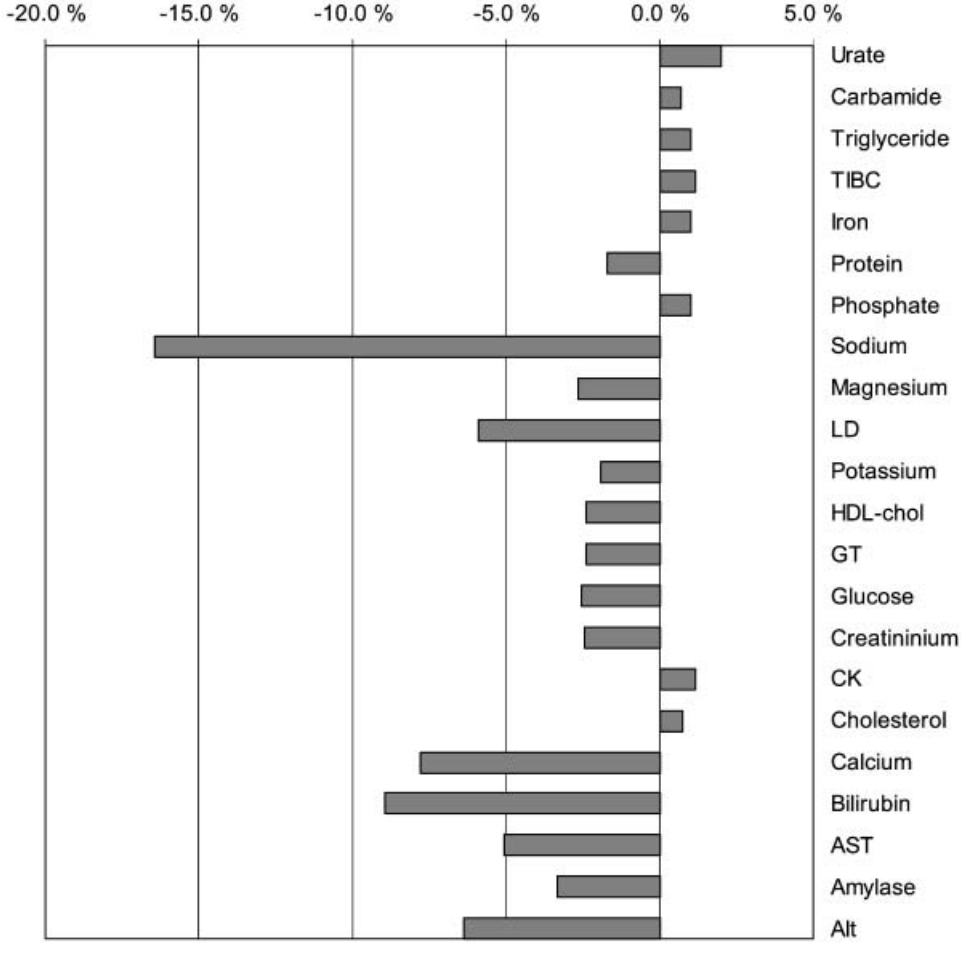

FIG. 1. Relative change in the difference between upper and lower reference limits after removing series not fulfilling bias quality goals. As expected, the obvious trend is that differences decrease the most extreme change for sodium. The changes are, however, rather small for most other properties.TIBC = total iron-binding capacity; LD = lactate dehydrogenase; GT = γ-glutamyltransferase; CK = creatine kinase; AST = aspartate transaminase; ALT = alanine transaminase.

TABLE III. Reference intervals for three electrolytes before and after exclusion of reference values. Exclusions have been done either according to bias goal for acceptance of series (34–42% excluded for each property) or because of a large deviation between measurements on different materials (serum/plasma, fresh/thawed) for the same individual.

	Calcium		Sodium		Potassium	
Before exclusion	2.15	2.54	135.8	145.7	3.6	4.7
After series exclusion	2.17	2.53	137.2	145.3	3.6	4.7
After material exclusions	2.17	2.51	136.7	144.8	3.6	4.6
Final agreement	2.15	2.51	137	145	3.6	4.6

between reference limits (H – L) after removing inferior series is decreased by 3% as a mean for the 23 components. As can be seen in Figure 1, only 8 of the components get larger differences and then by a maximum of 2%, while 16 components get narrower ranges, the most extreme for sodium with −16% (+1.4 mmol/L for L and −0.3 mmol/L for H) but also apparently important for calcium with −8% (+0.02 mmol/L for L and −0.01 mmol/L for H).

At a later stage in the project, a rule was introduced to delete analytical results that disagreed largely with results from serum and plasma for the same individual: if one result was outside and one inside the reference interval, the result outside was deleted if the absolute difference was greater than 1.5·CVb. For calcium, 81 reference values (3.0%) were deleted using this rule, for sodium 59 (2.1%) and for potassium 94 (3.4%). For the three electrolytes, the reference intervals before and after exclusion as described in this study and the final results for the project are presented in Table III. For sodium, the effect of deleting inferior series or deleting extreme values by material differences yields exactly the same result, narrowing the reference interval by 1 mmol/L at each end. For calcium, deletion of inferior series results in a 0.02 mmol/L higher upper limit than deletion of extreme values based on material comparisons. For potassium, there is no effect of series deletion, whereas material deletion results in a 0.1 mmol/L lower upper limit. This is interesting as the discrepancy for serum upper reference limit for NORIP (4.6 mmol/L) on the one side and Tietz (5.1 mmol/L) and Laurell (5.0 mmol/L) on the other cannot be explained by material deletion as the main cause.

The reference intervals based on the concept of correction with a common calibrator (commutable serum pool) are almost insensitive to the deletion of the worst analytical series. This was not obvious to the project group before this study was carried out. Later in the project, it was therefore decided to keep analytical series that did not fulfil bias goals as described before calculating reference intervals. The benefit of doing so is that a large number of reference values are important when evaluating the effects of partitioning because the uncertainty of calculated reference limits increases as the number of reference values decreases.

REFERENCES

1 Solberg HE, editor. International Federation of Clinical Chemistry, Expert Panel on Theory of Reference Values, and International Committee for Standardization in Haematology, Standing Committee on Reference Values. Approved recommendation (1986) on the theory of reference values. Part 1. The concept of reference values. J Clin Chem Clin Biochem 1987; 25: 337–42. Clin Chim Acta 1987; 165: 111–8. Labmedica 1987; 4: 27–31. Ann Biol Clin 1987; 45: 237–41.

2 Solberg HE, Stamm D, editors. International Federation of Clinical Chemistry, Expert Panel on Theory of Reference Values. IFCC recommendation. The theory of reference values. Part 4. Control of analytical variation in the production, transfer and application of reference values. Eur J Clin Chem Clin Biochem 1991; 29: 531–5. J Autom Chem 1991; 13: 231–4. Ann Biol Clin 1991; 49: 487–90. Clin Chim Acta 1991; 202: S5–12.

3 Solberg HE, editor. International Federation of Clinical Chemistry, Expert Panel on Theory of Reference Values, and International Committee for Standardization in Haematology, Standing Committee on Reference Values. Approved recommendation (1987) on the theory of reference values. Part 5. Statistical treatment of collected reference values. Determination of reference limits. J Clin Chem Clin Biochem 1987; 25: 645–56. Clin Chim Acta 1987; 170: S13–32.

4 Hyltoft Petersen P, Blaabjerg O, Irjala K, editors. Assessing quality in measurements of plasma proteins. The Nordic Protein Project and Related Projects. Uppsala J Med Sci 1994; 99: 195–389.

5 Bäck S-E, Nilsson J-E, Fex G, Jeppson J-O, Rosén U, Tryding N, von Schneck H, Norlund L. Towards common reference intervals in clinical chemistry. An attempt at harmonization between three hospital laboratories in Skåne, Sweden. Clin Chem Lab Med 1999; 378: 573 – 92.
6 Fuentes-Arderiu X, Ferré-Masferrer M, González-Alba JM, Escolà-Aliberas J, Balsells-Roselló D, Blanco-Cristobal C, González-Cruz E, Ibarz-Escuer M, Latorre-Marcellán P, Lugo-Arocena J, Mar-Medina C, Muros-de-Fuentes M, Vicens-Manero M. Related articles, links multicentric reference values for some quantities measured with Tina-Quant reagents systems and RD/Hitachi analysers. Scand J Clin Lab Invest 2001; 61: 273 – 6.
7 Fuentes-Arderiu X, Ferré-Masferrer M, González-Alba JM, Villarino-González MI, Arrimadas-Esteban E, Cabrero-Olivé D, Cándenas-Arroyo M, García-García D, García-Lario JV, Idoate-Cervantes I, León-López C, López-Lazareno N, Mar-Medina C, Martí-Marcet I, Mauri-Dot M, Pérez-Valero V, Reta-Manterola A, Sánchez-Eixeres MR. Multicentric reference values for some quantities measured with the Elecsys 2010 analyser. Clin Chim Acta 2001; 304: 143 – 6.
8 Gowans EMS, Hyltoft Petersen P, Blaabjerg O, Hørder M. Analytical goals for the acceptance of common reference intervals for laboratories throughout a geographical area. Scand J Clin Lab Invest 1988; 48: 757 – 64.
9 Hyltoft Petersen P, Gowans EMS, Blaabjerg O, Hørder M. Analytical goals for the estimation of non-Gaussian reference intervals. Scand J Clin Lab Invest 1989; 49: 727 – 37.
10 Fraser CG, Hyltoft Petersen P, Libeer J-C, Ricós C. Proposals for setting generally applicable quality goals solely based on biology. Ann Clin Biochem 1997; 34: 8 – 12.

Received: 27 January 2004
Accepted: 10 March 2004