UC 577,1; 61 ISSN 0354-3447

Jugoslov Med Biohem 24: 207–214, 2005 *Pregledni članak*[1]
Review article

ESTABLISHING REFERENCE LABORATORIES IN LABORATORY MEDICINE

Lothar Siekmann[2]

Institut für Klinische Biochemie, Universität Bonn, Sigmund-Freud-Str. 25, 53127 Bonn, Germany

Summary: The concept of measurement traceability provides probably the most important strategy to achieve standardisation in laboratory medicine aimed at comparable measurement results regardless of the method, the measurement procedure (test kit) and of the laboratory where analyses are carried out. Establishing networks of reference laboratories is – in addition to reference measurement procedures and reference materials – one of the biggest challenges in implementing the concept of measurement traceability. With respect to these requirements, the Joint Committee on Traceability in Laboratory Medicine (JCTLM), established by the BIPM, the IFCC and the ILAC, has launched two projects in its working groups. WG-1 has to date published tables of reference materials and reference procedures on the BIPM web-sites, whereas WG-2 is identifying reference measurement laboratories. There is general agreement now that reference laboratories should be identified
– according to the metrological level of the procedures applied where the principle of measurement is the most important criterion,
– on the basis of accreditation or at least compliance with ISO 15195 or ISO 17025 as calibration laboratory, and
– on the basis of their ability to demonstrate performance in regular inter-laboratory comparisons (ring trials).
To date, a data base on candidate reference laboratories has been collected containing information on the laboratory identity, the metrological level of the procedures and on the status of accreditation and the participation in networks or ring trials. The data base currently contains the addresses of about 60 laboratories. On average, each of the laboratories reported measurement capabilities for six different measurands resulting in about 360 entries. The IFCC has recently launched a ring trial program for reference laboratories for some thirty different measurands. Ring trial results not only demonstrate the competence of individual laboratories, but also reveal the equivalence or bias of different reference procedures.

Key words: traceability, standardisation, reference (calibration) laboratories

Introduction

In recent decades, considerable efforts have been devoted to the achievement of standardisation of measurement procedures and results in laboratory medicine. The concept of measurement traceability, which was developed in general analytical chemistry, provides probably the most important strategy to achieve standardisation in laboratory medicine aimed at comparable measurement results regardless of the method, the measurement procedure (test kit) and the laboratory where analyses are carried out. Essential tools for the implementation of the concept of measurement traceability in laboratory medicine are reference materials, reference measurement procedures and – one of the biggest challenges – networks of reference laboratories.

The In Vitro Diagnostic Directive (1) of the European Union stipulates that values assigned to calibrators and control materials must be traceable to reference materials and/or reference methods of higher order.

According to the 'Vocabulary in Metrology (VIM)' (2) and the 'Guide to the Expression of Uncertainty in Metrology (GUM)' (3) measurement traceability is defined as *property of the result of a measurement*

Address for correspondence:

Prof. Dr. L. Siekmann
Institut für Klinische Biochemie, Universität Bonn
Sigmund-Freud-Str. 25, 53127 Bonn, Germany

[1] Presented on Symposium for Balkan region on Education, Management and Standards in Laboratory Medicine, Belgrade, June 10–12th, 2005.

[2] Chair Traceability in Laboratory Medicine of the Scientific Division of the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC).

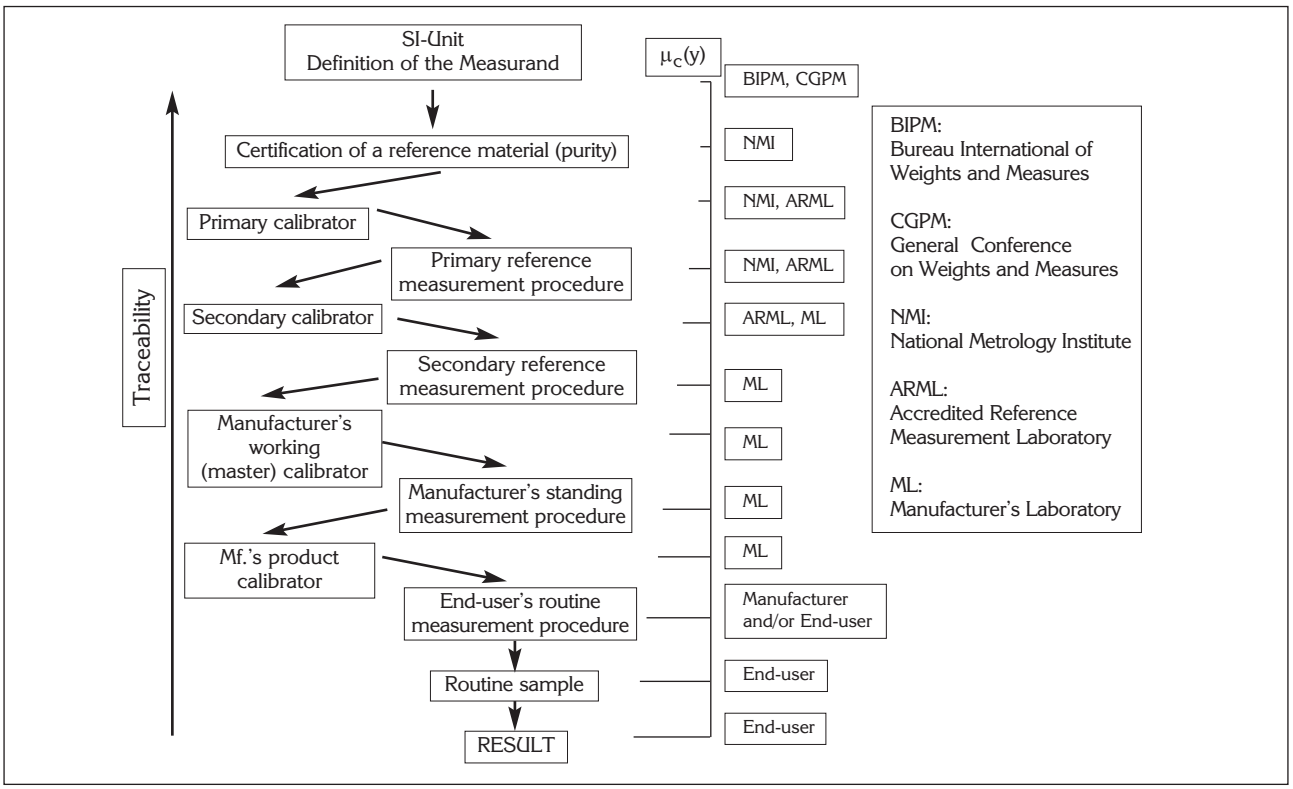


Figure 1 Calibration hierarchy and traceability in laboratory medicine according to ISO IEC 17511 (4)

or the value of a standard whereby it can be related to stated references, usually national or international standards, through an unbroken chain of comparisons all having stated uncertainties

Traceability of a value attributed to a routine sample, a calibrator or a control material is established by a series of comparative measurements using measurement procedures and reference materials in a chain of decreasing hierarchical order as shown in *Figure 1*. This is described in detail in ISO standard 17511 (4). Since each link in the traceability chain contributes to the uncertainty of the result it is advisable to omit as many steps as possible. In terms of metrology it would be ideal to omit all in-between steps of the traceability chain and to measure the routine sample directly by use of a primary reference procedure; this of course is not feasible.

The complete traceability chain as presented here is valid only for those measurable quantities which can have a value expressed in SI units. When primary or secondary calibrators are not available, the traceability chain for many measurands in laboratory medicine ends at a lower level, e.g. at the manufacturer's standing measurement procedure. In a case where a manufacturer detects a new diagnostic marker and defines the measurable quantity by establishing a measurement procedure for this marker, the manufacturer's measurement procedure will form the top of the traceability chain. Nevertheless, even in this simple situation, the principles of the traceability concept are applicable.

An inevitable precondition for establishing traceable results to calibrators and control materials is the specificity of the measurement procedures applied. Results of measurement cannot be traceable when the procedure applied partially detects components which are not consistent with the definition of the measurand.

The implementation of this concept obviously requires reference measurement procedures, reference materials and reference laboratories. With respect to these requirements, the Joint Committee on Traceability in Laboratory Medicine (JCTLM), established by the International Bureau of Weights and Measures (BIPM), the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) and the International Laboratory Accreditation Cooperation (ILAC), has launched two projects in its working groups. WG-1 has to date published tables of reference materials and reference procedures on the BIPM web-sites, whereas WG-2 is identifying reference measurement laboratories (5).

In the hierarchical scheme of laboratories, the National Metrology Institutes form the top level, followed by accredited calibration (reference) laboratories and, finally, the testing (routine) laboratories (*Figure 2*).

There is general agreement now that reference laboratories should be identified

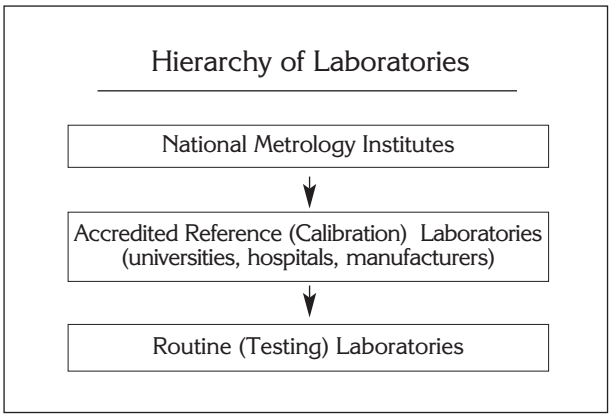


Figure 2 Hierarchical structure of calibration and testing laboratories in laboratory medicine

– according to the metrological level or principle of measurement,
– on the basis of accreditation or at least compliance with ISO 15195 (6) or ISO 17025 (7) as calibration laboratory, and
– on the basis of their ability to demonstrate performance in regular inter-laboratory comparisons (ring trials).

To date, a data base on candidate reference laboratories has been collected which currently contains the addresses of about 60 laboratories. On average, each of the laboratories reported measurement capabilities for six different measurands resulting in about 350 to 360 entries.

Laboratories were asked to report details on the measurement principle, the calibrator and the control materials as well as the literature used.

The most common principle of measurement used as reference measurement procedure for the determination of metabolites and substrates, e.g. glucose, creatinine, urea, uric acid, cholesterol and total glycerol was isotope dilution mass spectrometry (IDMS); a small number of laboratories applied an HPLC technique.

Enzymes are usually analysed by the IFCC primary reference procedures; some Japanese laboratories applied national standard methods.

For electrolytes, e.g. sodium or potassium, we find a variety of different metrological principles comprising flame emission and atomic absorption spectrometry, ion chromatography and gravimetry.

Steroid and thyroid hormones as well as therapeutic drugs were exclusively determined by IDMS.

Additional information is requested concerning the status of accreditation or compliance with relevant standards. Currently, only very few laboratories in laboratory medicine are accredited as reference laboratories – or calibration laboratories according to either ISO 15195 or ISO 17025. It should be stressed that ISO 17025 concerns all types of laboratories, routine (testing) and calibration laboratories, and that it contains some separate paragraphs dedicated specifically to calibration laboratories. In the accreditation process it should be emphasised that compliance with the calibration laboratory requirements is essential. ISO 15195 is dedicated particularly to calibration laboratories in laboratory medicine, here called reference laboratories. ISO 15195 uses the terminology of laboratory medicine but the requirements are the same as formulated in ISO 17025 for calibration laboratories.

The management system requirements formulated in ISO 15195 are similar to those requested for any type of laboratory; however, the *technical requirements* are dedicated to the metrological aspects that have to be observed by laboratories responsible for 'calibration'.

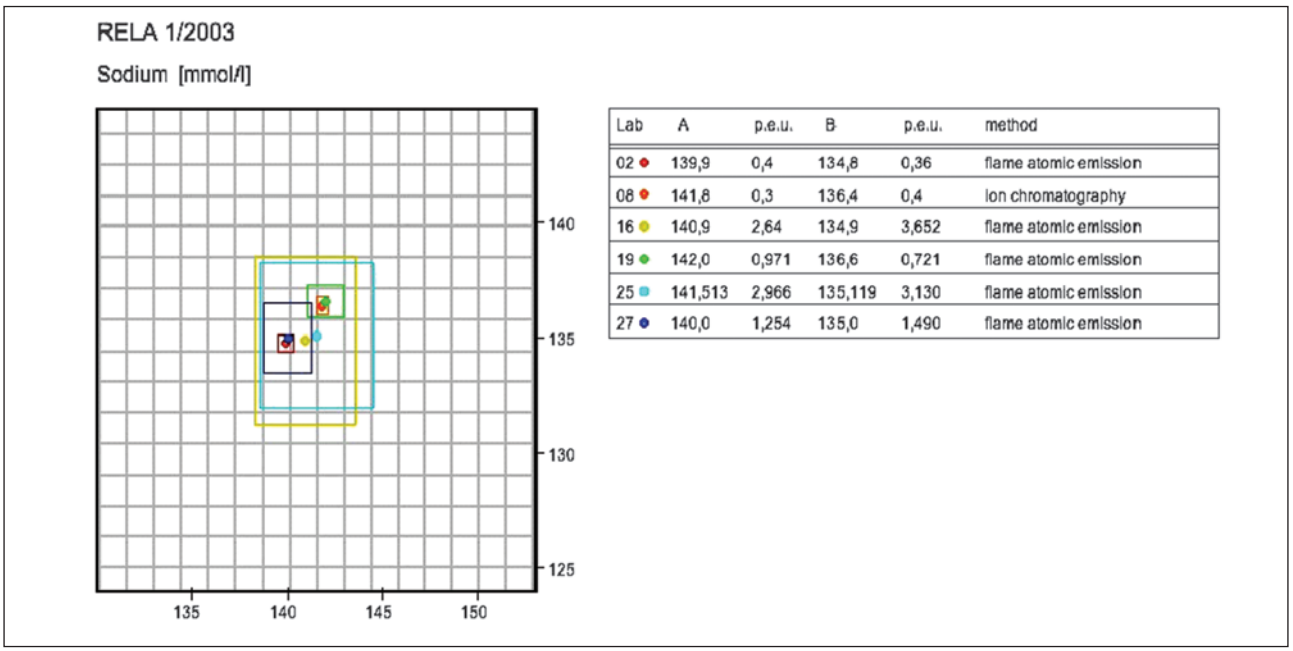
Another section of the questionnaire concerned the *frequency of measurements*. Surprisingly, one laboratory declared that it performs more than 8000 measurements using IFCC enzyme reference procedures per annum. Considering the efforts necessary in performing reliable reference measurements, there was some doubt in that statement and it turned out that this laboratory actually performed routine measurements using commercial IFCC traceable test procedures.

The last items of the questionnaire concerned the participation in a *network* or in *ring trials*. Evidently, the candidate reference laboratories only occasionally performed comparative measurements, e.g. in the IMEP programme or in national networks conducted by a National Metrology Institute. The CCQM offered a small number of ring trials (cholesterol, glucose and creatinine) where only National Metrology Institutes were accepted as participants.

In view of this, in 2003, the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) launched a ring trial program for reference laboratories. Currently, laboratories are asked to register for the second ring trial through a questionnaire in the form of an Excel data sheet.

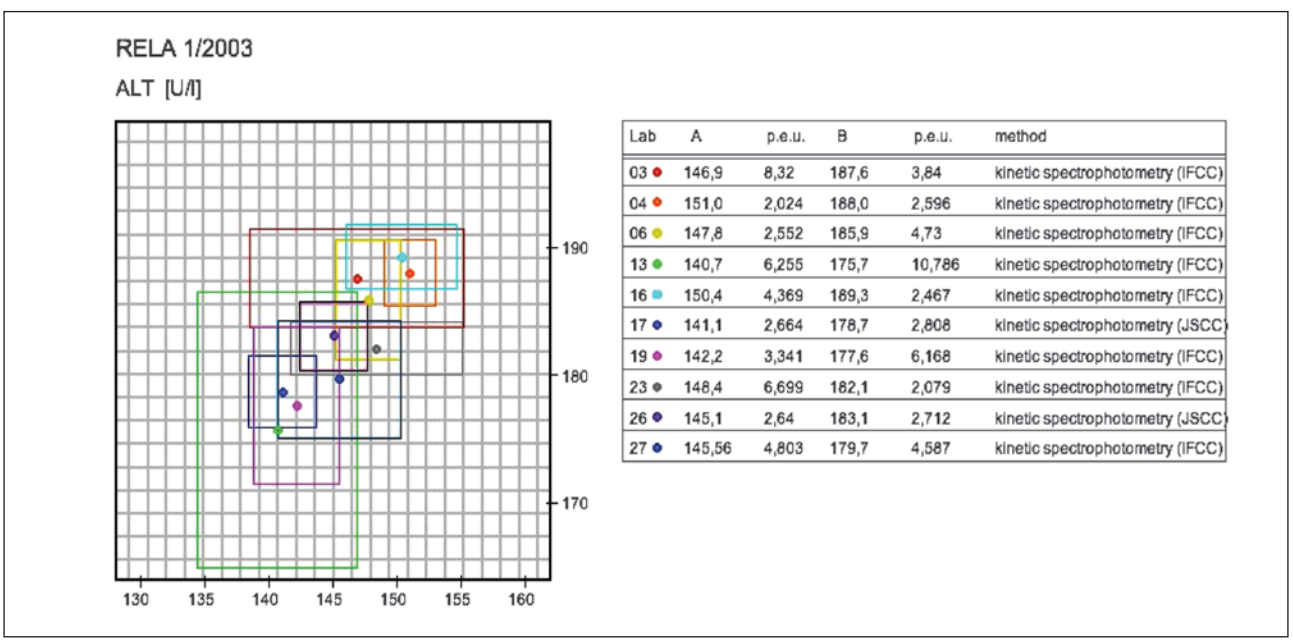
Ring Trials for Reference Laboratories are offered for some thirty different measurands according to the following rules:

a) Ring trials will be offered for the following *groups of measurands:*
   – metabolites and substrates,
   – electrolytes, basic
   – enzymes,
   – hormones,
   – therapeutic drugs.



Lab	A	p.e.u.	B	p.e.u.	method
02	139,9	0,4	134,8	0,36	flame atomic emission
08	141,8	0,3	136,4	0,4	ion chromatography
16	140,9	2,64	134,9	3,652	flame atomic emission
19	142,0	0,971	136,6	0,721	flame atomic emission
25	141,513	2,966	135,119	3,130	flame atomic emission
27	140,0	1,254	135,0	1,490	flame atomic emission

Figure 3 Ring trial results from candidate reference laboratories for *sodium* in serum. Each dot in the diagram represents the two results of each laboratory for the two different sera (abscissa: sample A, ordinate: sample B). The rectangles around the dots show the uncertainty for each laboratory



Lab	A	p.e.u.	B	p.e.u.	method
03	146,9	8,32	187,6	3,84	kinetic spectrophotometry (IFCC)
04	151,0	2,024	188,0	2,596	kinetic spectrophotometry (IFCC)
06	147,8	2,552	185,9	4,73	kinetic spectrophotometry (IFCC)
13	140,7	6,255	175,7	10,786	kinetic spectrophotometry (IFCC)
16	150,4	4,369	189,3	2,467	kinetic spectrophotometry (IFCC)
17	141,1	2,664	178,7	2,808	kinetic spectrophotometry (JSCC)
19	142,2	3,341	177,6	6,168	kinetic spectrophotometry (IFCC)
23	148,4	6,699	182,1	2,079	kinetic spectrophotometry (IFCC)
26	145,1	2,64	183,1	2,712	kinetic spectrophotometry (JSCC)
27	145,56	4,803	179,7	4,587	kinetic spectrophotometry (IFCC)

Figure 4 Ring trial results from candidate reference laboratories for *ALT* in serum. Each dot in the diagram represents the two results of each laboratory for the two different sera (abscissa: sample A, ordinate: sample B).

b) Ring trials will be offered regularly (every 6–12 months),

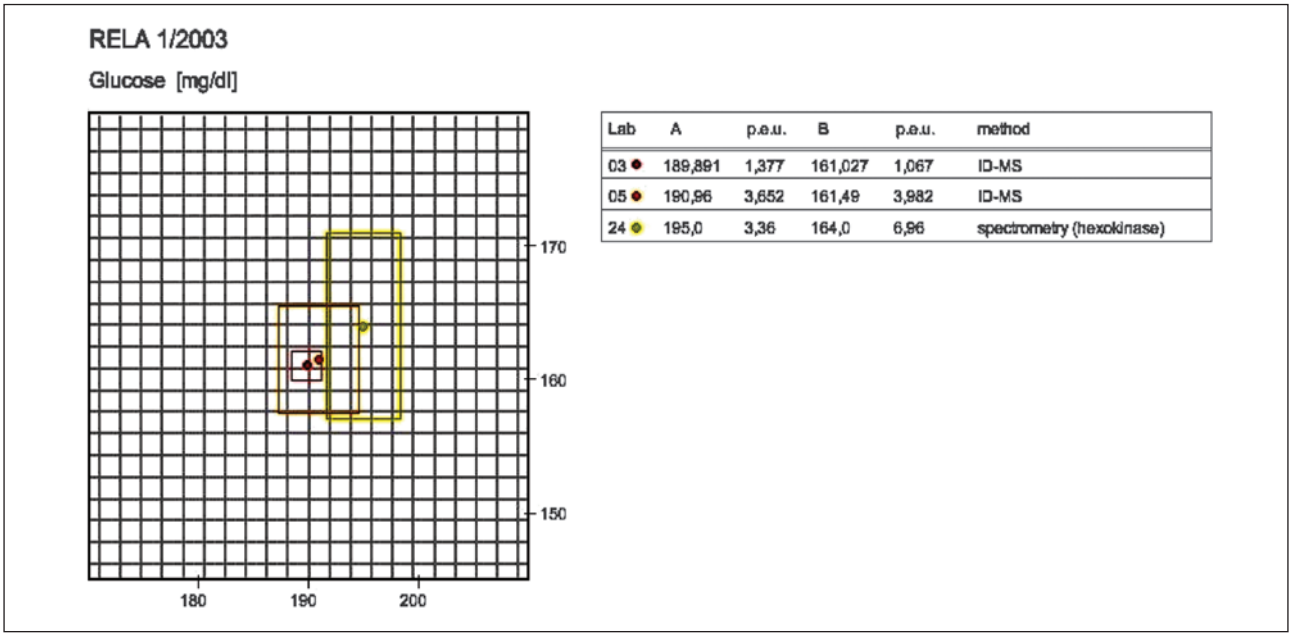
c) The deadline for reporting results after distribution of samples will be five months.

Since reference procedures are laborious and time-consuming it is necessary to allow at least five months for analyses.

Evaluation and reporting by the EQAS organisation will take about one month ensuring that results will be available before the next ring trial is initiated.

d) Two different samples will be distributed in each ring trial; results are evaluated as YOUDEN diagrams.

e) Each participating laboratory will receive an overall evaluation as well as a report on its individual performance. According to the rules set out by the



Lab	A	p.e.u.	B	p.e.u.	method
03	189,891	1,377	161,027	1,067	ID-MS
05	190,96	3,652	161,49	3,982	ID-MS
24	195,0	3,36	164,0	6,96	spectrometry (hexokinase)

Figure 5 Ring trial results from candidate reference laboratories for *glucose* in serum. Each dot in the diagram represents the two results of each laboratory for the two different sera (abscissa: sample A, ordinate: sample B).

JCTLM Executive, results and addresses will no longer be confidential. However, we will offer the participants the possibility to withdraw individual laboratory results for particular measurands one month after a preliminary report of the results and before reporting the data on the web-site.

f) Key measurands will be selected for each group of measurands in each ring trial.

This is necessary to collect statistically sound information from the ring trials.

Example:

The EQAS organisation will offer ring trials for all five enzyme activity measurements according to the IFCC 37 °C reference procedures. Due to the workload, not all laboratories will participate for all measurands. At the worst, Lab.A will analyse ALT, Lab.B AST, Lab.C GGT, Lab.D CK and Lab.E LDH. Because of the limited number of participating laboratories it will be difficult to collect a statistically relevant number of results for each of the measurands necessary to demonstrate comparability of results from different laboratories.

Therefore, the EQAS organiser will select one key measurand from each group of measurands for every ring trial occasion. It is highly recommended to provide results for this key measurand. Participation for all other measurands is voluntary. The selected key measurand for each group of measurands will change from one ring trial to the next.

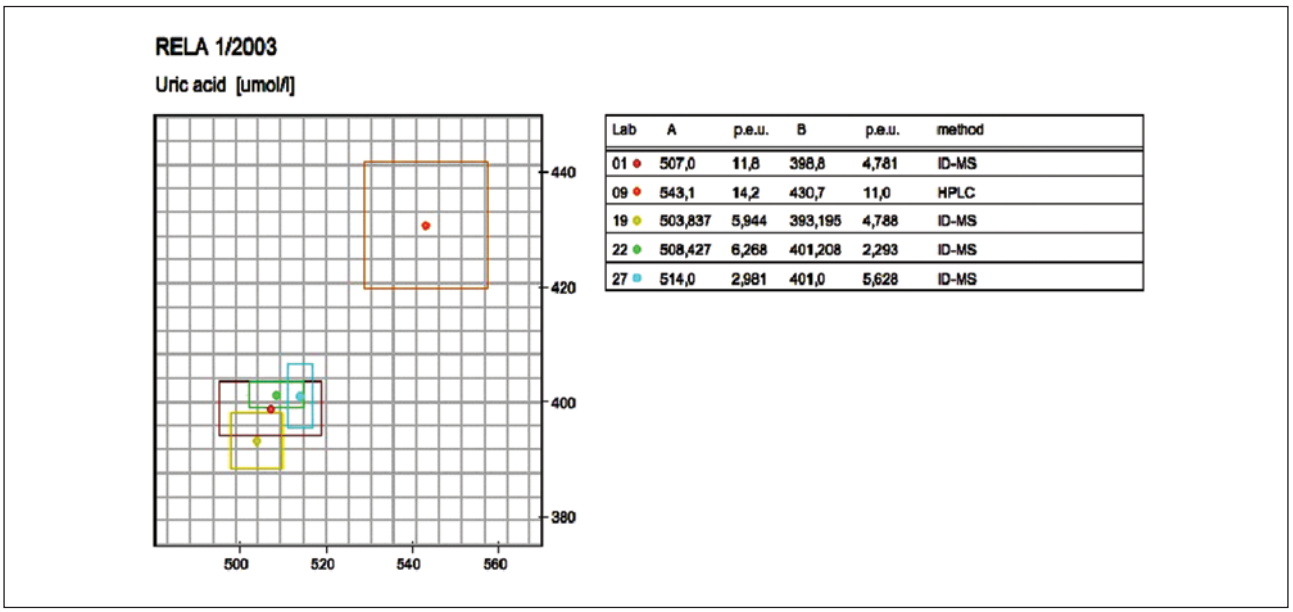
In the first ring trial conducted in October 2003, 28 laboratories registered for various measurands. Of these, 26 sent in their results. As was to be expected, the majority of results were reported for the key measurands, e.g. cholesterol, sodium, ALT and progesterone. In total, 92 pairs of results were received and evaluated in YOUDEN diagrams.

Each dot in a YOUDEN diagram for the measurement of *sodium* in serum reflects the two results reported from one laboratory, whereby the value for sample A may be read from the abscissa and for sample B from the ordinate (*Figure 3*). The grey scales in the diagrams amount to 1% of the mean values. The rectangles around the dots reveal the individual expanded uncertainties.

For *sodium*, five laboratories used flame emission spectrometry and one laboratory ion chromatography. The small relative standard deviation of the reference laboratory results of about 0.6% among the six laboratories reflects an excellent agreement.

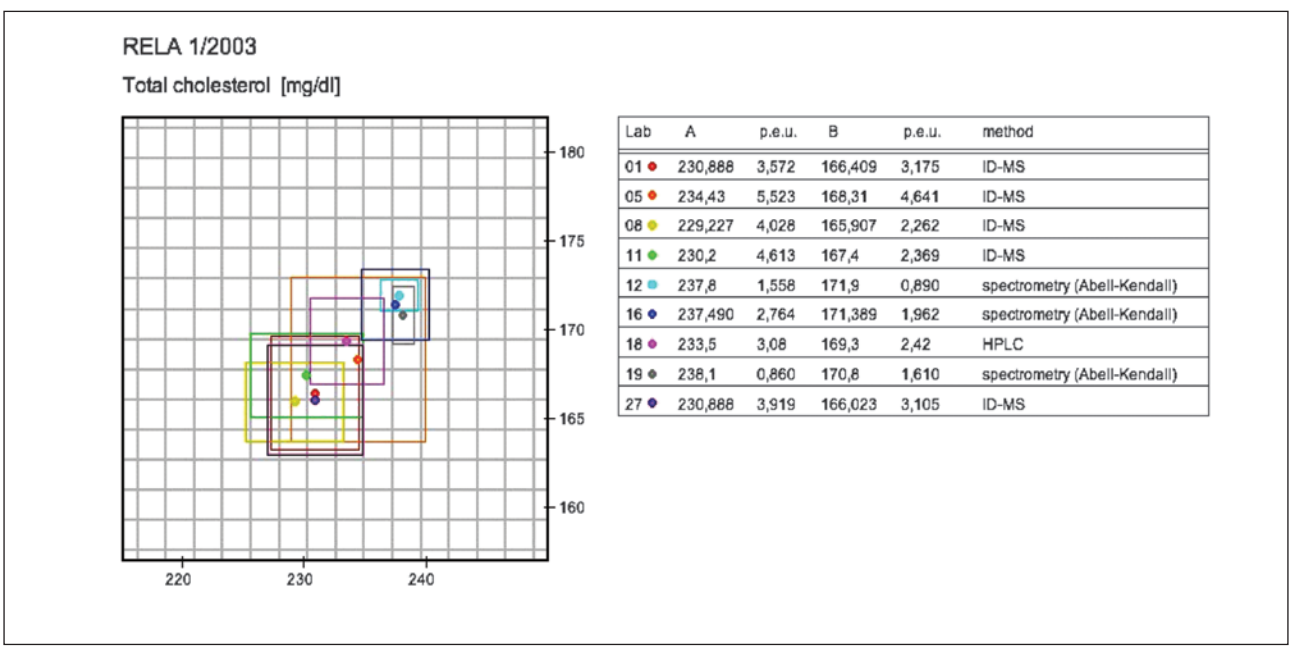
For *enzymes*, the situation is much more difficult. The results for ALT show a considerable dispersion between the laboratories (2.6–3.1% rel. standard deviation) as can be shown in *Figure 4*. This may be explained by the fact that the enzyme procedures do not use calibrators but are based on absolute measurements of the reaction rates where many inherent types of uncertainties, e.g. from the adjustment of temperature, pH, absorbance and wavelength play an important role. Therefore, the uncertainty of the measurement of enzymes is inevitably larger than for other measurands.

For *glucose*, only three sets of results were reported (*Figure 5*). It seems that there is a bias of the laboratory which used the spectrophotometric



Lab	A	p.e.u.	B	p.e.u.	method
01	507,0	11,8	398,8	4,781	ID-MS
09	543,1	14,2	430,7	11,0	HPLC
19	503,837	5,944	393,195	4,788	ID-MS
22	508,427	6,268	401,208	2,293	ID-MS
27	514,0	2,981	401,0	5,628	ID-MS

Figure 6 Ring trial results from candidate reference laboratories for *uric acid* in serum. Each dot in the diagram represents the two results of each laboratory for the two different sera (abscissa: sample A, ordinate: sample B).



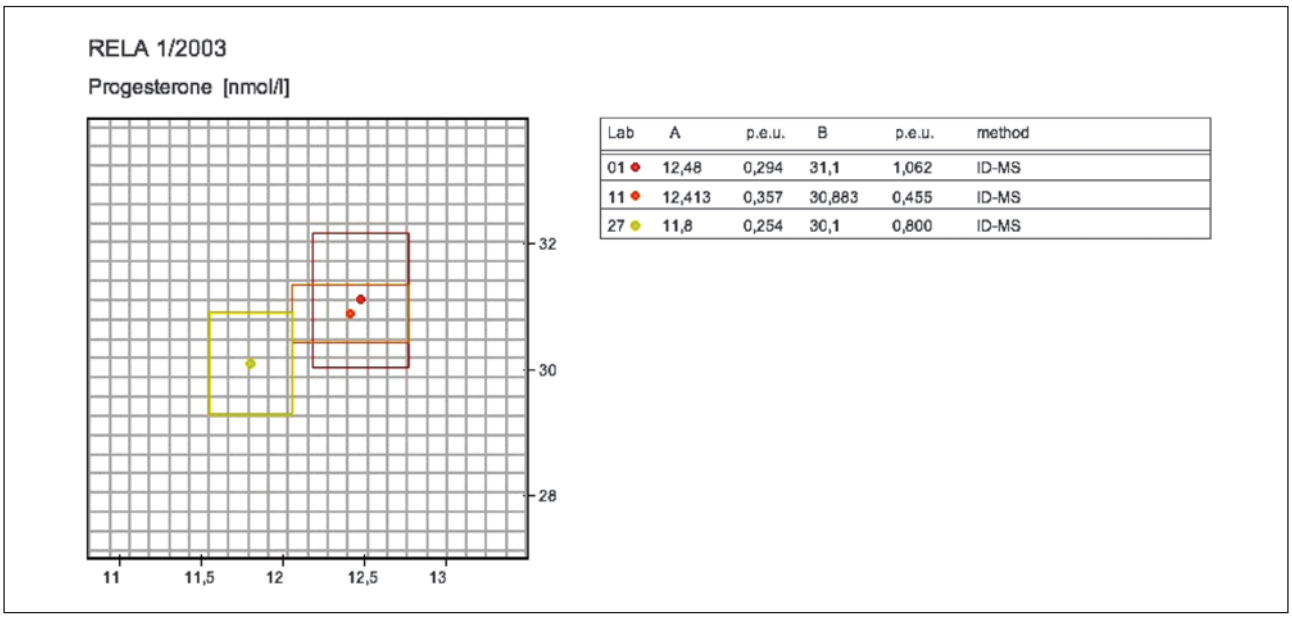
Lab	A	p.e.u.	B	p.e.u.	method
01	230,888	3,572	166,409	3,175	ID-MS
05	234,43	5,523	168,31	4,641	ID-MS
08	229,227	4,028	165,907	2,262	ID-MS
11	230,2	4,613	167,4	2,369	ID-MS
12	237,8	1,558	171,9	0,890	spectrometry (Abell-Kendall)
16	237,490	2,764	171,389	1,962	spectrometry (Abell-Kendall)
18	233,5	3,08	169,3	2,42	HPLC
19	238,1	0,860	170,8	1,610	spectrometry (Abell-Kendall)
27	230,888	3,919	166,023	3,105	ID-MS

Figure 7 Ring trial results from candidate reference laboratories for *total cholesterol* in serum. Each dot in the diagram represents the two results of each laboratory for the two different sera (abscissa: sample A, ordinate: sample B).

hexokinase-gluc-6-phosphate dehydrogenase procedure, whereas the two sets of results from IDMS laboratories are in excellent agreement.

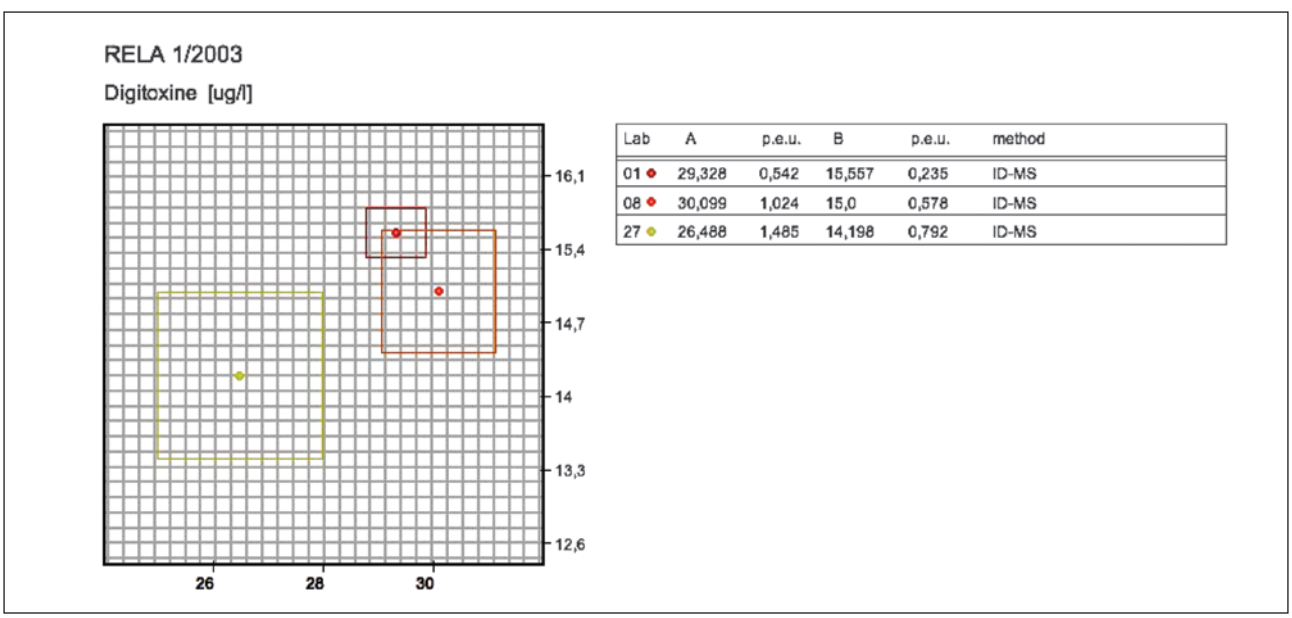
For *uric acid*, four laboratories using IDMS showed excellent agreement, whereas another laboratory performing HPLC showed considerable bias of its results (*Figure 6*).

From a metrological point of view, the ring trial result for *cholesterol* was quite interesting. At first glance, there seems to be a considerable dispersion of results (*Figure 7*). However, a more detailed inspection shows that there are two groups of results, one from the *IDMS laboratories* and the second with higher values of the *Abell-Kendall* users. A statistical evaluation demonstrated an acceptable small relative standard deviation of less than 1% between the IDMS labo-



Lab	A	p.e.u.	B	p.e.u.	method
01	12,48	0,294	31,1	1,062	ID-MS
11	12,413	0,357	30,883	0,455	ID-MS
27	11,8	0,254	30,1	0,800	ID-MS

Figure 8 Ring trial results from candidate reference laboratories for *progesterone* in serum. Each dot in the diagram represents the two results of each laboratory for the two different sera (abscissa: sample A, ordinate: sample B).



Lab	A	p.e.u.	B	p.e.u.	method
01	29,328	0,542	15,557	0,235	ID-MS
08	30,099	1,024	15,0	0,578	ID-MS
27	26,488	1,485	14,198	0,792	ID-MS

Figure 9 Ring trial results from candidate reference laboratories for *digitoxine* in serum. Each dot in the diagram represents the two results of each laboratory for the two different sera (abscissa: sample A, ordinate: sample B).

ratories. Among the Abell-Kendall laboratories, the dispersion of results is even smaller. However, a significant bias of almost 3% between the two principles of measurement could be observed.

Although *progesterone* was announced as a key measurand, only three laboratories were prepared to report results (*Figure 8*). Two of them are in good agreement, one came up with considerably lower values, in particular for the sample A.

Also, for *digitoxine*, only three laboratories reported results, whereby two of them were in good agreement (*Figure 9*).

It can be summarised that the JCTLM will identify reference laboratories on the basis of their *status of accreditation* or compliance with ISO 15195, on the results of *regularly performed ring trials*, on the *principle of measurement* applied.

USPOSTAVLJANJE REFERENTNIH LABORATORIJA U LABORATORIJSKOJ MEDICINI

Lothar Siekmann

Institut für Klinische Biochemie, Universitat Bonn, Sigmund-Freud-Str. 25, 53127 Bonn, Germany

Kratak sadržaj: Koncept merenja sledljivosti obezbeđuje verovatno najznačajniju strategiju za postizanje standardizacije u laboratorijskoj medicini s ciljem dobijanja uporedivih rezultata za metodu, merne procedure (test kit) i laboratorije u kojima se analize izvode. Uspostavljanje mreže referentnih laboratorija je – osim referentnih mernih procedura i referentnih materijala – jedan od najvećih izazova u implementiranju koncepta merne sledljivosti. Uvažavajući ove zahteve, Združeni komitet za sledljivost u laboratorijskoj medicini (Joint Committee on Traceability in Laboratory Medicine, JCTLM) koji su uspostavili BIPM, IFCC i ILAC, postavio je dva projekta svojim radnim grupama. WG-1 je objavila tabele referentnih materijala i referentnih procedura na BIPM web-stranici, dok je WG-2 identifikovala referentne merne laboratorije. Za sada postoji opšti stav da se referentne laboratorije identifikuju:

– na osnovu metrološkog nivoa primenjene procedure kod koje je princip merenja najznačajniji kriterijum,
– na osnovu akreditacije ili najmanje saglasnosti sa ISO 15195 ili ISO 17025 kao kalibracione laboratorije, i
– na osnovu njihove sposobnosti da dokažu sposobnost izvođenja u redovnim interlaboratorijskim izvođenjima (kružni trajali).

Da bi se uspostavila baza podata o kandidovanoj referentnoj laboratoriji sakupljaju se informacije o identitetu laboratorije, metrološkom nivou procedura i o statusu akreditacije i učestvovanja u mreži kružnih trajala. Do sada se u bazi nalaze podaci koji sadrže adrese oko 60 laboratorija. U proseku svaka od ovih laboratorija je sposobna za merenje šest različitih grupa sa oko 360 jedinica. IFCC je nedavno uspostavila program ispitivanja u referentnim laboratorijama oko trideset različitih sastojaka. Rezultati ovakvih ispitivanja ne samo da pokazuju kompetentnost individulalne laboratorije, već ukazuju na ekvivalentnost ili odstupanja od različitih referentnih procedura.

Ključne reči: sledljivost, standardizacija, referentne (kalibracione) laboratorije

References

1. European Parliament and the Council of European Union (1998) OFF J Eur Commun No L 331: 1–37.
2. BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML. International vocabulary of basis and general terms in metrology. English and French. 2nd ed. Geneva: ISO, 1993 (abbreviated VIM).
3. ISO, Guide to the expression of uncertainty in measurement, 1st ed. Geneva, 1995 (abbreviated GUM).
4. EN ISO 17511:2003, In vitro diagnostic medical devices – Measurement of quantities in biological samples – Metrological traceability of values assigned to calibrators and control materials http://www.bipm.org/en/committees/jc/jctlm-db/
5. EN ISO 15195:2003, Laboratory medicine – Requirements for reference measurement laboratories.
6. ISO IEC 17025:1999, General requirements for the competence of testing and calibration laboratories.