*Klin. Biochem. Metab., 17 (38), 2009, No. 2, p. 100–105.*

# Time to engage in measurement uncertainty*

**Stöckl D.**
*STT Consulting, Horebeke, Belgium*

**SUMMARY**
Measurement uncertainty is an important evolving concept in laboratory medicine. This concept is described in the ISO document "Guide to the Expression of Uncertainty in Measurement", often referred to as "GUM". GUM has two important aspects; a statistical one, which deals with the propagation of standard deviations; an analytical one, which counts on the expertise of the analytical chemist for addressing measurement uncertainty that extends beyond the obvious (sample related effects, stability of analytes, etc.). This paper gives a short introduction to the GUM concept and addresses the main challenges for its application in laboratory medicine (for example, the treatment of bias). Besides some basic GUM calculations, it describes a generic laboratory approach for calculating measurement uncertainty from available data (typically from manufacturers' data sheets). The latter example shows the "spirit of GUM" which is equally important to the "statistics of GUM". The dilemma connected to the treatment of bias (correct or not correct) is demonstrated by internal quality control data.
*Keywords:* measurement uncertainty; GUM; analytical expertise; bias; laboratory approach to GUM; internal quality control.

**SOUHRN**
**Stöckl D.: Nastal čas zabývat se neurčitostí měření**
Nejistota měření je důležitým konceptem laboratorní medicíny. Tento koncept je popsán v dokumentu ISO „Guide to the Expression of Uncertainty in Measurement", všeobecně známém pod zkratkou GUM. Výpočet nejistoty má dva důležité aspekty – jeden statistický, založený na zákonu o šíření standardních odchylek, a druhý analytický, zabývající se použitelností výsledku v důsledku různých vlivů působících v procesech měření. Tento článek je stručným úvodem do problematiky stanovení nejistot podle GUM a poskytuje laboratořím orientaci k aplikaci nejistoty měření v praxi. Kromě některých základních postupů vycházejících z GUM jsou popsány i přístupy vycházející z dostupných dat uváděných v pracovních návodech výrobců. Sdělení se také zabývá důležitým dilematem, jak postupovat při měření nejistoty s hodnotami bias. Rovněž demonstruje použití hodnot vnitřní kontroly kvality k výpočtu nejistot.
*Klíčová slova:* nejistota měření, GUM, bias, vnitřní kontrola kvality, aplikace v klinické laboratoři.

## Introduction

Measurement results are inherently variable due to the influences of random and systematic effects. This variability must be quantified so that the user of the results has knowledge of their reliability. The Guide to the Expression of Uncertainty in Measurement (GUM) [1] provides general rules for quantifying measurement variability. Measurement variability quantified by the rules of GUM is called measurement uncertainty (see also Table 1 for definitions [2, 3]). Because of confusion on what GUM is really about and the many personal interpretations of GUM, the concept is introduced by use of original GUM citations (indicated by "quotation marks" and paragraph number).

**Table 1.** Definitions

| |
|---|
| **Measurement uncertainty [2, 3]**<br>non-negative parameter characterizing the dispersion of the *quantity* values being attributed to a *measurand*, based on the information used |
| **Quantity [2, 3]**<br>property of a phenomenon, body, or substance, to which a number can be assigned with respect to a reference |
| Quantities are designated in laboratory medicine by the format "system–component (analyte); kind of quantity" (for example, serum–cortisol; amount-of-substance concentration equal to *x* µmol/l). |
| **Measurand [2, 3]**<br>quantity intended to be measured |

One of the most important aspects of the GUM concept is that "it is assumed that the result of a measurement has been corrected for all recognized significant systematic effects and that every effort has been made to identify such effects" (3.2.4). However, "the result of a measurement after correction for recognized systematic effects is still only an estimate of the value of a measurand because of the uncertainty arising from random effects and from imperfect correction of the result for systematic effects" (3.3.1). Thus, "in general, the result of a measurement is only an approximation or estimate of the value of the measurand and thus is complete only when accompanied by a statement of the uncertainty of that estimate" (3.1.2). Measurement uncertainty is evaluated by two different methods; "Type A: method of evaluation of uncertainty by the statistical analysis of series of observations" (2.3.2); "Type B: method of evaluation of uncertainty by means other than the statistical analysis of series of observations" (2.3.3). "Type B is evaluated by scientific judgement based on all the available information on the possible variation of the measurand. The pool of information may include i) previous measurement data; ii) experience with or general knowledge of the behaviour and

* First published at http://www.westgard.com/guest41.htm.

properties of relevant materials and instruments; iii) manufacturer's specifications; iv) data provided in calibration and other certificates; v) uncertainties assigned to reference data taken from handbooks" (4.3). "The purpose of the Type A and Type B classification is to indicate the two different ways of evaluating uncertainty components and is for convenience of discussion only; the classification is not meant to indicate that there is any difference in the nature of the components resulting from the two types of evaluation. Both types of evaluation are based on probability distributions, and the uncertainty components resulting from either type are quantified by means of variances or standard deviations" (3.3.4). Note, "these categories are not substitutes for the words random and systematic" (3.3.3). "Type A uncertainty is obtained from a probability density function derived from an observed frequency distribution, while Type B uncertainty is obtained from an assumed probability density function based on the degree of belief that an event will occur" (3.3.5). It should be stressed that "the definition of uncertainty of measurement is not inconsistent with other concepts of uncertainty of measurement, such as a measure of the possible error in the estimated value of the measurand as provided by the result of a measurement. Nevertheless, "whichever concept of uncertainty is adopted, an uncertainty component is always evaluated using the same data and related information" (2.2.4). See Table 2 for a summary of the GUM concept.

**Table 2.** The GUM concept

| |
|---|
| GUM is CONSISTENT with other concepts, e. g., total error, BUT NOT THE SAME |
| Significant systematic effects must be corrected and NOT included in the calculations |
| Insignificant systematic effects may be neglected |
| Uncertainty is evaluated from the statistics of repeated measurements (Type A) AND from scientific judgement (Type B) |
| Types A & B uncertainties are not different in nature |
| Types A & B uncertainties are not substitutes for the words random and systematic |
| Types A & B uncertainties are quantified by variances or standard deviations |

Calculation of measurement uncertainty (adapted from the NIST website)

**Caution:** Some heavy stuff ahead. If you wish, continue below.

Before starting an uncertainty calculation, one has to define the measurement equation. Usually, the quantity $Y$ (called the measurand), is not measured directly, but is determined from $N$ other quantities $X_1, X_2, \ldots, X_N$ through a functional relation $f$, often called the measurement equation: $Y = f(X_1, X_2, \ldots, X_N)$. Included among the quantities $X_i$ are corrections (or correction factors), as well as other sources of variability, such as different observers, instruments, samples and sampling, laboratories, and times at which observations are made. Thus, the function $f$ of the above equation should express not simply a physical law but a measurement process, and in particular, *it should contain all quantities that can contribute a significant uncertainty to the measurement result*. An estimate of the measurand or *output quantity* $Y$, denoted by $y$, is obtained from the above equation using *input estimates* $x_1, x_2, \ldots, x_N$ for the values of the $N$ *input quantities* $X_1, X_2, \ldots, X_N$. Thus, the *output estimate* $y$ (= result), is given by $y = f(x_1, x_2, \ldots, x_N)$. The uncertainty of the measurement result $y$ arises from the uncertainties $u(x_i)$ (or $u_i$) of the input estimates $x_i$ that enter the equation. The combined standard uncertainty of the measurement result $y$, designated by $u_c(y)$ is the positive square root of the estimated variance $u_c^2(y)$ obtained from

$$u_c^2(y) = \sum_{i=1}^{N} (\partial f/\partial x_i)^2 \, u^2(x_i) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} (\partial f/\partial x_i)(\partial f/\partial x_j) \, u(x_i, x_j)$$

This equation is based on a first-order Taylor series approximation of the measurement equation and is referred to as the *law of propagation of uncertainty*. The partial derivatives of $f$ are often referred to as *sensitivity coefficients* and $u(x_i, x_j)$ is the covariance associated with $x_i$ and $x_j$. The partial derivatives of $f$ with respect to the $X_i$ are equal to the partial derivatives of $f$ with respect to the $X_i$ evaluated at $X_i = x_i$; $u(x_i)$ is the standard uncertainty associated with the input estimate $x_i$. Note, standard uncertainties of Type A equal standard deviations typically estimated in the laboratory ($u(x_i) = SD(x_i)$). If the probability distribution of $y$ and its combined standard uncertainty $u_c(y)$ is approximately normal (Gaussian) then the interval $y \pm u_c(y)$ should encompass ~68% of the values that could reasonably be attributed to $Y$ ($Y = y \pm u_c(y)$). If other confidence levels are desired, the standard uncertainty may be expanded (expanded uncertainty: $U = k\, u_c$) by use of a coverage factor, $k$ (e. g., 2 or 3: approximate confidence of 95 or > 99%).

**Continue**

GUM uncertainty calculations can be made as sophisticated as desired. I will present here only the "lightweight" version. At the end, a laboratory approach is given for estimating measurement uncertainty using available data.

However, if you are required to do in-depth GUM calculations (manufacturer; in-house method developer), it is recommended to consult the internet resources given below and to purchase the GUM.

**GUM calculations "lightweight"**

Boiled down, GUM is about propagation of standard deviations/variances. This is also part of standard clinical chemistry textbooks [4]. The 2 simpliest cases of error propagation will be addressed, namely, sums and multiplications/divisions (both follow the same rule) (see also Table 3).

1. The measurand is described by the function
$y = a + b$; with SDa = 5, SDb =10
$u(y)$ = SQRT(5 x 5 + 10 x 10) = SQRT(125) = 11.2;
SQRT = square root

CAVE: Do **not** use the CV of the methods for measuring a and b.
2. The measurand is described by the function $y = a/b$;
with a = 70, b = 40 ($y$ = 70/40 = 1.75),
SDa = 5 (CVa = 100*[5/70] = 7.1%), SDb =10 (CVb = 25%)
$u(y) = y$ x SQRT($[5/70]^2$ + $[10/40]^2$) = 1.75 x 0.26 = 0.455 (= 26% of $y$ =1.75)
Note, the relative variances are propagated; therefore, the CV can be used:
$u(y)$ (%) = SQRT($[7.1]^2$ + $[25]^2$) = 26% (26% of 1.75 = 0.455)
These equations shall be applied to the calculation of the uncertainty of the anion gap and the creatinine clearance.

**Table 3.** Basic rules for propagation of standard deviations (random errors)

| |
|---|
| 1. The measurand is described by the function $y = a + b$; with SDa = 5, SDb =10<br>$u(y)$ = SQRT(5x5 + 10x10) = SQRT(125) = 11.2; SQRT = square root<br>CAVE: Do not use the CV of the methods for measuring a and b. |
| 2. The measurand is described by the function $y = a/b$;<br>with a = 70, b = 40 ($y$ = 70/40 = 1.75), SDa = 5 (CVa = 100*[5/70] = 7.1%), SDb =10 (CVb = 25%)<br>$u(y) = y$ x SQRT($[5/70]^2$ + $[10/40]^2$) = 1.75 x 0.26 = 0.455 (= 26% of $y$ =1.75)<br>Note, the relative variances are propagated; therefore, the CV can be used:<br>$u(y)$ (%) = SQRT($[7.1]^2$ + $[25]^2$) = 26% (26% of 1.75 = 0.455) |

**Applications** [see also 4]
*Anion gap (AG)*
AG = ($[Na^+]$+$[K^+]$) – ($[Cl^-]$+$[HCO_3^-]$)
For daily practice, potassium is frequently ignored, leaving the equation:
AG = ($[Na^+]$) – ($[Cl^-]$+$[HCO_3^-]$) mmol/l
For example: AG = 140 – (106 + 22) = 12 mmol/l
SD$[Na^+]$ = 1.3 mmol/l; SD$[Cl^-]$ = 1.2 mmol/l;
SD$[HCO_3^-]$ = 0.7 mmol/l
SD[AG] = SQRT($1.3^2$ + $1.2^2$ + $0.7^2$) = 1.9 mmol/l
This value equals the standard uncertainty for the anion gap:
$u$(AG) = 1.9 mmol/l; using a coverage factor of 2 (approximately 95% probability), the expanded uncertainty is $U$ = 3.8 mmol/l. This is equivalent to a total error (TE), in the absence of a systematic error (SE), calculated as 2 times the random error (RE): TE = 2 x RE.

**Creatinine clearance**
Ccr = (U x V)/S mL/min
Creatinine clearance (Ccr; mL/min); urine creatinine (U; µmol/l); volume urine/minute (V; mL/min); serum creatinine (S; µmol/l)
With CV(U) = 3%; CV(V) = 10%; CV(S) = 3%
CV(Ccr) = SQRT($3^2$ + $10^2$ + $3^2$) = 11%
At a Ccr of 80 mL/min, $u$(Ccr) = 8.8 mL/min (= 11% of 80 mL/min)

**So why the fuss about GUM, there's nothing new?**
While GUM computations may be simple, the GUM philosophy encourages the analyst to look "what is behind the input data", in particular, to address the Type B uncertainty. The problem with the anion gap, for example, is that it may widely vary with instruments. Therefore, care should be taken that the involved measurement procedures are correctly standardized and the correct reference interval is used for its interpretation [5, 6]. This line of thought is continued in the example below.

Note, some of the most important advantages and disadvantages of the GUM concept are listed in Tables 4 and 5.

**Table 4.** Advantages of the GUM concept

| |
|---|
| When properly done, it gives realistic information about test performance because it considers all sources of measurement variability |
| It teaches error propagation rules |
| It may identify the most important source of test variation |

**Table 5.** Disadvantages of the GUM concept

| |
|---|
| Establishment of the full measurement function may be very difficult |
| Error propagation rules may become complex, in particular, when input elements are correlated |
| Bias is not included in the calculations (> Move to total error, for example) |

## Laboratory approach using available data

**The assay**
*Measurand:* Serum/plasma–testosterone;amount-of--substance concentration (nmol/l)
*Intended use:* Immunoassay for the in vitro quantitative determination of testosterone in human serum and plasma.
*Clinical applications (selection):*
**Women:** Diagnosis of androgenic syndrome, polycystic ovaries, tumors.
**Men:** Suspected reduced testosterone production (hypogonadism, estrogen therapy).
*Test principle:* Competitive assay: 2 incubations, separation, wash, detection, clean.
*Master calibration:* Commercial testosterone (no primary reference material available) using a 5 point calibration curve (spline function).

Calibration based on method comparison with an isotope dilution mass spectrometry reference procedure using 40 native sera available, but not implemented. Reason: risk of clinical misinterpretation; international standardization awaited (see Table 6).
*Customer calibration:* 2 point, with every new reagent lot.
*Recalibration:* 14 days if actual kit lasts longer; 2 months when using the same lot.
*Quality control:* 2 levels, once per day (rules and acceptance criteria defined by user).

**Table 6.** Uncertainty given by manufacturer ($u_c$ = SQRT[$u_{Cal}^2 + u_{Impr}^2$])

| | Uncertainty (%), Levels | | |
|---|---|---|---|
| **Uncertainty** | **0.8 nmol/l** | **6 nmol/l** | **20 nmol/l** |
| $u_{Cal}$ | 4 | 1.5 | 2 |
| $u_{Impr}$ | 12 | 7.1 | 5 |
| $u_c$ | 12.6 | 7.2 | 5.4 |
| *U* | 25.3 | 14.4 | 10.8 |

$u_{Cal}$ = calibration uncertainty
$u_{Impr}$ = uncertainty imprecision (including variation from 3 reagent batches)
$u_c$ = combined uncertainty (note: due to the squaring principle, the combined uncertainty is entirely dominated by the imprecision)
*U* = expanded uncertainty, coverage factor 2

## Review of the information

The laboratory realizes that the uncertainty estimate of the manufacturer does not include data on:
– trueness;
– sample related effects (matrix; interferences);
– effects of limits for linearity, recovery, method comparability;
– uncertainty at low female concentrations;
– stability of calibrators and reagents.

The laboratory decides to obtain additional information from the manufacturer's technical documentation (see Table 7) and the scientific literature.

**Table 7.** Information from manufacturer's technical documentation

| | | |
|---|---|---|
| *Analytical specificity* | Cross reactivity data, but no indication of their relevance for patients' specimens. | |
| *Interference* | Interference limit 10% (concentrations given for lipids, etc). No interference with common drugs. | |
| *Linearity* | Limit 10% | |
| *Recovery* | Limit 10% | |
| *Method comparison* | Acceptable slope: 0.9 – 1.1 | |
| *Limit of detection (LoD)* | 0.1 nmol/l | |
| *Quantitation limit (CV = 20%)* | 0.4 nmol/l | |
| *Reportable range range (from LoD to calibration maximum)* | 0.1–60 nmol/l | |
| *Expected values* | Male: 10–28 nmol/l<br>Female: 0.2–3 nmol/l | |
| *Stability data calibrator and reagent lots* | Some decline during maximum recommended time (no limits given). | |
| *Lot-to-lot criteria calibrators and reagents* | No information available. | |
| *Imprecision (EP 5)* | *Intra-assay CV (%)* | *Total CV (%)* |
| Human, 0.8 nmol/l: | 4.6 | 7.4 |
| Human, 24 nmol/l: | 1.1 | 1.7 |
| Low control, 6.5 nmol/l: | 1.7 | 2.6 |
| High control, 20 nmol/l: | 0.9 | 1.6 |
| Note: those data are not considered as long-term imprecision (done with one lot, over 20 days). | | |

## Scientific information

Consultation of the scientific literature revealed the risk of considerable sample related effects, in particular, for females [7]. The incidence of antibody interference seems to be low [8, 9], while interferences due to cross-reactivity seem to be more common [10–12]. Interferences to consider are dehydroepiandrostenedione sulphate and testosterone conjugates.

## Laboratory approach for estimating measurement uncertainty using all above information

The laboratory could verify the imprecision data, but decided to modify the uncertainty estimates of the manufacturer in the following way:
– the uncertainty "point" estimates were converted into intervals and one range was added;
–the estimate in the low range was expanded for sample related effects and considering the quantitation limit of the assay (total effect: factor of 2);
– the estimate in the low-medium range was expanded by sample-related effects (in the order of the total imprecision) and the imprecision was interpolated;
– an uncertainty of 5% was added in all ranges to account for recovery and linearity;
– the lower end of the working range was increased to 0.25 nmol/l (relative big difference between the LoD and the quantitation limit).

A risk analysis was done for interferences and a policy was written.

The trueness problem was discussed with the manufacturer and their rationale was accepted.

The laboratory keeps the following uncertainty estimates in its files (Table 8).

**Table 8.** Uncertainty estimates kept in the laboratory files

| | Uncertainty (%), Ranges (nmol/l) | | | |
|---|---|---|---|---|
| Uncertainty | 0.25–0.7 | 0.7–2 | 2–10 | > 10 |
| $u_{Cal}$ | 4 | 2 | 1.5 | 2 |
| $u_{Impr}$ | 15 | 10 | 7.1 | 5 |
| $u_{Sample}$ | 20 | 10 | 7.1 | – |
| $u_{Lin\text{-}Rec}$ | 5 | 5 | 5.0 | 5 |
| $u_c$ | 25.8 | 15.1 | 11.3 | 7.3 |
| *U* | 52 | 30 | 23 | 15 |

$u_{Cal}$ = calibration uncertainty
$u_{Impr}$ = uncertainty from long-term imprecision
$u_{Sample}$ = sample related effects
$u_{Lin\text{-}Rec}$ = linearity & recovery
$u_c$ = combined uncertainty
*U* = expanded uncertainty, coverage factor 2

## Beyond GUM

The long-term internal quality control data indicated a somewhat high lot-to-lot variation ($u$ = 10%). The laboratory made total error calculations and simulations by introduction of biases. It found that biases of 10% changed the results "to be acted upon" by 50%. While this was deemed too high, no solution could be found. The laboratory increased its quality assurance efforts and introduced a quality control rule with an increased power.

**Limitations of GUM**

As outlined above, bias is not covered by the GUM calculations but needs to be corrected. However, the treatment of bias (existing or input in total error models) is vital to the laboratory. For example, to investigate the effect of reagent batch-to-batch variations on patient data. Figure 1 below shows a test with a batch-to-batch CVbb of 10% (= 10 at a value of 100) and a within-batch CVwb of 5%. It was created by simulating 20 random numbers with a SD of 10 and a mean of 100. Then, for each of the 20 values (batch means are indicated by bars), 20 random numbers were simulated with a SD of 5. The figure would represent quality control data obtained with a batch lasting 20 days and doing 1 QC sample a day. Further, it is assumed that the mean of the stable process is known to be 100.

According to GUM, 2 possibilities exist. If the observation time is extended over all 20 batches, the biases of the individual batches become random and the total CVtot becomes 11.2%. The laboratory may decide to keep in its files that the process has an uncertainty of 11.2%, without considering the bias introduced when changing reagent batches. This, however, would give a false impression about the test performance, because the bias in each reagent batch may have a profound influence on diagnostic decisions [13]. If the observation time is 2 batches, considerable systematic effects would be seen from time to time (batch 2, for example). This is the reasoning why GUM deprecates the distinction between "random" and "systematic": it may depend on the observation time. According to GUM, one would correct the second batch giving a mean of 120 (bias = 20%). The laboratory, however, is usually unable to correct for batch-to-batch variations. Nevertheless, it needs to know the effect of a 20% bias on the patient results. If such a bias would increase the false positives by 50%, for example, it may require the manufacturer to tighten his batch-to-batch variations. Also, the laboratory needs a model that accounts for bias in order to select the appropriate quality control rules. Such a model, for example, is the total error approach used in the Westgard software products.

Contrary to the GUM philosophy, it is vital for the laboratory to distinguish between random and systematic effects. When systematic effects have to be taken into account, other concepts must be used for describing measurement variability, such as the total error concept. Thus, in my opinion, the different concepts are complementary and not contradictory. GUM alone, however, is unsufficient for managing real-world situations in the clinical laboratory.

**A note on Quality control**

The above example shows a dilemma of quality control: shall the laboratory use a CV of 11.2% or a CV of 5% as input value for the QC process? If a CV of 11.2% is chosen, typical QC rules seldom will give alarms. If a CV of 5% is chosen, typical QC rules will indicate problems regularly. But then, what to do? Currently, there is no easy answer to the problem. Obviously, for QC purposes, one could change the target value of the quality control sample, however, this changes nothing for the bias of the patient samples.

In the future, it would be desirable that manufacturers keep the between-batch variation in the same order as the within-batch variation. For comparison, Figure 2 shows a QC chart with CVbb = CVwb = 5%. The total CVtot is 7.1%.

## References


1. ISO/IEC Guide 98:1995. Guide to the expression of uncertainty in measurement (GUM). International Organization for Standardization: Geneva, 1995.
2. ISO/IEC Guide 99:2007. International vocabulary of metrology – Basic and general concepts and associated terms (VIM). International Organization for Standardization: Geneva, 2007.
3. JCGM 200:2008. International vocabulary of metrology – Basic and general concepts and associated terms (VIM). International Bureau of Weights and Measures (BIPM); Joint Committee for Guides in Metrology (JCGM): Paris, 2008 (electronic document freely available at: http://www.bipm.org/en/publications/guides/vim.html).
4. **Kringle, R. O.** *Statistical Procedures*. In Burtis, C. A., Ashwood, E. R. [eds]. *Tietz Textbook of Clinical Chemistry,* 2nd edition, Chapter 12, p. 419–422. Philadelphia: Saunders, 1994.


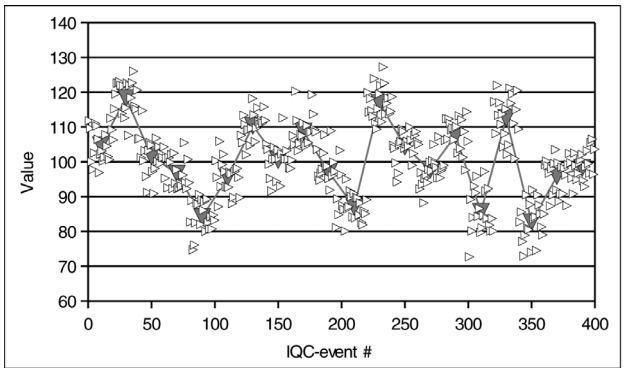


**Fig. 1.** Quality control data: 20 reagent batches with a CVbb of 10%, each measured 20 times with a CVwb of 5%; mean of the stable process = 100.

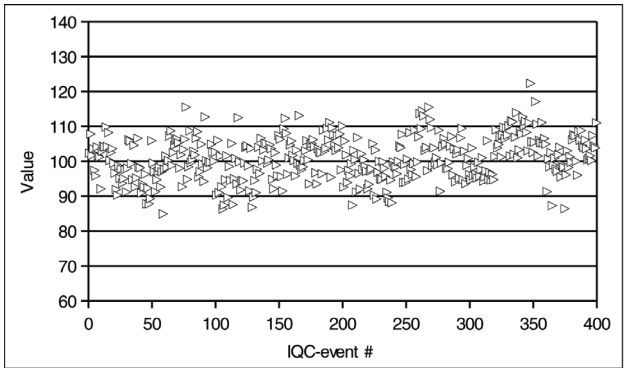


**Fig. 2.** Quality control data: 20 reagent batches with a CVbb of 5%, each measured 20 times with a CVwb of 5%; mean of the stable process = 100.

5. **Kraut, J. A., Madias, N. E.** Serum anion gap: its uses and limitations in clinical medicine. *Clin. J. Am. Soc. Nephrol.,* 2007, 2, p. 162–174.
6. **Paulson, W. D., Roberts, W. L., Lurie, A. A., Koch, D. D., Butch, A. W., Aguanno, J. J.** Wide variation in serum anion gap measurements by chemistry analyzers. *Am. J. Clin. Pathol.,* 1998, 110, p. 735–742.
7. **Taieb, J., Mathian, B., Millot, F., Patricot, M. C., Mathieu, E., Queyrel, N., Lacroix, I., Somma-Delpero, C., Boudou, P.** Testosterone measured by 10 immunoassays and by isotope-dilution gas chromatography-mass spectrometry in sera from 116 men, women, and children. *Clin. Chem.,* 2003, 49, p. 1381–1395.
8. **Kuwahara, A., Kamada, M., Irahara, M., Naka, O., Yamashita, T., Aono, T.** Autoantibody against testosterone in a woman with hypergonadotropic hypogonadism. *J. Clin. Endocrinol. Metab.,* 1998, 83, p. 14–16.
9. **Torjesen, P. A., Bjøro, T.** Antibodies against [125I] testosterone in patient's serum: a problem for the laboratory and the patient. *Clin. Chem.,* 1996, 42, p. 2047–2048.
10. **Middle, J. G.** Dehydroepiandrostenedione sulphate interferes in many direct immunoassays for testosterone. *Ann. Clin. Biochem.,* 2007, 44, p. 173–177.
11. **Heald, A. H., Butterworth, A., Kane, J. W., Borzomato, J., Taylor, N. F., Layton, T., Kilpatrick, E. S., Rudenski, A.** Investigation into possible causes of interference in serum testosterone measurement in women. *Ann. Clin. Biochem.,* 2006, 43, p. 189–195.
12. **Stanczyk, F. Z., Cho, M. M., Endres, D. B., Morrison, J. L., Patel, S., Paulson, R. J.** Limitations of direct estradiol and testosterone immunoassay kits. *Steroids,* 2003, 68, p. 1173–1178.
13. **Thienpont, L. M.** Calculation of measurement mncertainty-Why bias should be treated separately. *Clin. Chem.,* 2008, 54, p. 1587.

*Internet resources*
http://physics.nist.gov/cuu/Uncertainty/index.html
http://physics.nist.gov/Pubs/guidelines/
http://www.measurementuncertainty.org/mu/guide/index.html

*Do redakce došlo 16. 1. 2009.*

*Adresa pro korespondenci:*
*Dietmar Stöckl, PhD*
*STT Consulting*
*Abraham Hansstraat 11*
*B-9667 Horebeke, Belgium*
*e-mail: dietmar@stt-consulting.com*

**Dietmar Stöckl, PhD.**

Dietmar Stöckl se narodil v roce 1954 v Castrop-Rauxel v Německu. Vystudoval univerzitu v Kolíně nad Rýnem a v roce 1982 se zde habilitoval prací z oblasti hmotnostní spektrometrie (MS). V letech 1985–1987 pracoval v oblasti MS biologických sloučenin jako vědecký pracovník a odborný asistent na univerzitě v Göttingenu. Pak řadu let vedl referenční laboratoř pro oblast laboratorní medicíny při INSTAND e.V. Düsseldorf, kde mimo jiné zavedl do rozsáhlého rutinního použití v oblasti reference právě analytickou techniku MS. Po roce 1995 nadále spolupracuje s INSTAND Düsseldorf, vede konzultační středisko STT Consulting a současně pracuje jako vědecký pracovník na katedře analytické chemie farmaceutické fakulty univerzity v Gentu v Belgii. Toto pracoviště je referenční laboratoří pro oblast klinické chemie a podílí se celosvětově na tvorbě referenčních metod měření a ustavení referenčních hodnot (RMP).

Odborně se zabývá externím hodnocením kvality, referenčními metodami, jejich standardizaci, statistickými postupy v oblasti reference a kontrolou kvality v oblasti laboratorní medicíny. Je autorem a spoluautorem několika set odborných publikací uveřejněných prakticky ve všech základních časopisech klinické chemie ve světě a řady monografií. Působil v mnoha mezinárodních pracovních skupinách a komisích IFCC (EQA – working groups; Analytical Quality – WG aj.). V současné době působí jako člen kolegia CLSI USA (Clinical and Laboratory Standards Institute) a je spoluautorem mnoha tam vzniklých Doporučení a Standardů pracovní skupiny IFCC pro tvorbu inovovaného Doporučení o vyjadřování nejistot v laboratorní medicíně (GUM – Guide to the Expression of Uncertainty in Measurement) a v řadě dalších odborných a edukačních orgánů IFCC.

*J. Kratochvíla*