Bioquímia

Reporte breve
Metrología

The uncertainty concept and its implications for laboratory medicine

Anders Kallner, PhD MD*

MEASUREMENT PERFORMANCE

Laboratory measurement performance needs to be estimated in a uniform and standardized manner to allow comparison of results and become scientifically valid. A much-used measure is the variation of repeated measurements and the agreement of the result with a reference or true value. There are thus principally two types of variation of results of measurements, systematic and random. Information on both types needs to be attached to the result of the measurement; that information shall be informative and understood by the end-users. Provided some rules are observed that can be achieved by relying on the concept of *uncertainty*.

Systematic variations result in changes in the agreement between the obtained value and the true value, i.e. the *trueness* of the result. If the *bias*, the statistic used to measure trueness, can be assessed, the results can be compensated for the deviation. Bias is however difficult to assess and particularly so in biological systems since the true value is rarely known. Bias can then be expressed as the deviation from a reference value obtained by reference methods.

The random variation, *precision*, is the closeness of the average of the results of a large number of replicate measurements. The statistic that is used to describe precision numerically is *imprecision*. The overall concept precision can be subdivided into *repeatability, reproducibility* and *intermediate precision*. The repeatability describes the performance if the measurement is repeated without any changes in the conditions, reagents etc. whereas reproducibility is the performance if all conditions have been changed. Intermediate precision is when but a few of the conditions have been changed; those changed shall be stated.*

The *total error*, which is a long-standing concept in clinical chemistry, is the sum of a systematic and a random error contribution. It is usually reported in relative form, i.e. percentage. It can be criticized, e.g. bias is often both absolute and relative to the result and it is not always correct to add the two components linearly. Likewise the concept of *accuracy* comprises precision and trueness because it describes the closeness of the result of one measurement to a reference or true value. Accuracy cannot be given a numerical value. Accuracy must not be incorrectly used for trueness.

An alternative to using the concept of errors that may be embedded in the results is to estimate and state the *uncertainty* of the result. The most important differences between the 'Error model' and the 'Uncertainty model' are listed in *table I*. A well-known parameter that describes the dispersion of results is the standard deviation, which thus is equal to the standard uncertainty, a measure of the imprecision.

ESTIMATION OF IMPRECISION

Imprecision can be estimated by several different methods. It is often obtained under repeatability con-

* Department of Clinical Chemistry
Karolinska University Hospital, Stockholm SWEDEN
E-mail: anders.kallner@karolinska.ki.se

Correspondencia:
Anders Kallner, PhD MD
Department of Clinical Chemistry
Karolinska University Hospital, Stockholm SWEDEN
E-mail: anders.kallner@karolinska.ki.se

Recibido: 16-03-2007
Aceptado: 28-03-2007

* Definitions of a selection of metrological terms used in the document are given in the Appendix.

Table I. Comparison of key elements of the error model and uncertainty model.

Error model	Uncertainty model
Single unknowable value.	Defines an interval within which the true value is assumed.
Could be corrected if known.	Cannot be used for correction or corrected.
Two types of component: • systematic error • random error	One type of component.
Applies to a single quantity value.	Applies to all values obtained according to a given procedure.

ditions, e.g. the quantity, e.g. the concentration is measured several times in a reference or patient sample in one series. Samples of several different concentrations are often used to create an 'imprecision profile' since there is no universal rule that either the absolute or relative variation is constant within a measuring interval. Another technique is to measure samples of different concentrations in duplicates and estimate the uncertainty from the difference between the paired results. In clinical laboratories, the results of measuring the quantity of control materials during a certain period of time e.g. approaching intermediate conditions are frequently used in the calculation. Only rarely, the imprecision is estimated under reproducibility conditions, e.g. all conditions changed.

If given the choice and possibility, it is an advantage to use patient materials to estimate the imprecision. The end-user also needs to know how the imprecision has been estimated to benefit optimally from that knowledge. It is therefore important that an agreement is reached between laboratories and end-users (clinicians) on how the imprecision shall be estimated.

Laboratories that are accredited according to the EN/ISO 15189 or ISO/IEC 17025 are obliged to report the uncertainty in their measurements. The EN/ISO 15189 states (5.6.3): "The laboratory shall determine the uncertainty of results, when possible and relevant. Uncertainty components that are of importance shall be taken into account." This statement is generally understood as implying that an uncertainty budget shall be established – when possible – and a combined uncertainty estimated.

UNCERTAINTY BUDGET, COMBINED AND EXPANDED UNCERTAINTY

The common understanding of "budget" is probably an act or a document that deals with the future use of available resources. When talking about an uncertainty budget in measurements, however, this means a description of the size, nature and functional relationship of the sources of uncertainty. Comprehensive discussions of the concept of uncertainty and how it shall be estimated and interpreted can be found in references 1 to 4 of which "Guide to the expression of uncertainty in measurement", GUM, is the parent document.[1]

Briefly, the 'GUM' method is to identify and quantify the standard uncertainty of all processes that together constitute the measurement procedure (input variables) and their functional relationship. The standard uncertainties are then combined according to certain rules to form the combined uncertainty. The standard uncertainty of a process is the same as its standard deviation and can be estimated either by any of the methods mentioned above, e.g. repeated measurements (Type A) or by assuming other distributions e.g. a rectangular distribution that allows the estimation of the standard uncertainty (Type B). Both process can be applied to the entire procedure and thus give the combined uncertainty directly. A most important assumption in the GUM is that all known biases have been eliminated or compensated, leaving an uncertainty of the success that can be added as another input variable to the combined uncertainty. The combined uncertainty thus describes the total uncertainty and is favorably compared with the total error concept *(Table I)*.

In their universal form, the 'uncertainty propagation rules'[1-4] are based on partial derivatives of the function that describes the relation between the input variables. Few laboratorians are familiar with this mathematical procedure. There is however a convenient method of numerical approximation of the general rules. [2] This procedure has been realized in a Microsoft Excel sheet.[5] For simple operations like addition-subtraction and multiplication-division manual procedures, which are applicable the rules are:

$C = A + B;$
estimate u_C *if*
u_A *and* u_B
$u_C = \sqrt{(u_A)^2 + (u_B)^2}$

And

$C = A \times B$
estimate u_C *if*
u_A *and* u_B
$\frac{u_C}{C} = \sqrt{\left(\frac{u_A}{A}\right)^2 + \left(\frac{u_B}{B}\right)^2}$

Where u_a, u_b and u_c represent the standard uncertainty of input variable A and B and the product C, respectively.

Estimates of a preanalytical variation and the uncertainty of the bias elimination may be entered as separate input variables into the uncertainty budget. Their functional relations may be additive or multiplicative.

The combined uncertainty is estimated as 1 SD, e.g. the result will be found within the interval delineated by the reported value ± the combined uncertainty with a probability of about 67%. If a higher probability of finding the true value within the limits is desired, then the combined uncertainty shall be multiplied by a 'coverage factor' (k). This gives the expanded uncertainty (U). A coverage factor of 2 is usually taken as resulting in a confidence interval of 95%. If an expanded uncertainty is reported, then the coverage factor must also be reported. However, like in scientific literature, it is most convenient, although not always conventional, to report the combined uncertainty only.

REPEATABILITY, REPRODUCIBILITY AND INTRA- OR INTERLABORATORY UNCERTAINTY

The combined uncertainty does not always satisfy the needs of the laboratories or the clinical end users of laboratory data. The laboratories need to know the repeatability performance and the intermediary imprecision (e.g. between series or after calibrations, or change of reagent or calibration lots) to properly manage monitoring of the quality. Laboratories and clinicians alike also need to know the intralaboratory variation e.g. the variation taking all these factors and the possibility of using several different instruments into account. The clinician is particularly interested in the latter aspect.

CLSI EP 15 advised an efficient and simple protocol that allows estimating the components of the combined uncertainty.[6] The protocol comprises repeated measurements of the quantities on several occasions, usually on different days. The original recommendation is three replicate measurements in five series, but by increasing the number of replicates, series confidence of the results can be increased. We have developed a Microsoft Excel spreadsheet that will carry out these calculations.

A similar approach[5] can be attempted to estimate the interlaboratory uncertainty, which is of interest if patients use different laboratories, and the physicians thus have to evaluate results that may not be comparable. It might then be appropriate to find means to harmonize the measurement procedures to increase the transferability of the results. Since Proficiency Testing (PT) schemes rather assess the accuracy of measuring the test sample than the bias of the laboratory, they are not helpful for the individual laboratory to take rational corrective actions.

CLINICAL USE OF THE UNCERTAINTY

The clinician uses the laboratory data for monitoring the status of a patient or for classifying the patient in relation to a reference value, e.g. in diagnosis. In both cases their concern is if the value obtained from the laboratory differs from a previous result or a biological reference value. It becomes important to advise a method to objectively estimate the least significant difference between two results. The principles for this have been outlined above. Thus the criteria for a significant difference between two values are that it shall be larger than the uncertainty of the difference Δ:

$$\Delta > k \times u_\Delta$$
$$u_\Delta = \sqrt{(u_A)^2 + (u_B)^2}$$

Where k is the coverage factor.

If the sample has been analyzed by the same laboratory it is feasible to assume that the uncertainty of both results is the same and the criteria changes to:

$u_\Delta > k \times u_A \times \sqrt{2}$. As a rule of thumb, considering k = 2 and the $\sqrt{2}$ equal to 1.4, the difference between two consecutive results should be 3 times the combined uncertainty to rule out – at a confidence level of 95% - that the difference is due to laboratory factors.

If compared with a reference value, the reasoning is the same. However, it is usually agreed that a reference value is without uncertainty; its value is the result of a consensus or other decision. Therefore,

the least significant difference in this case is reduced to $u_{\Delta} > k \times u_{A}$.

However, it is important to remember that all the uncertainty estimates are based on statistical considerations that may themselves be liable to considerable uncertainties! The physicians' judgment will therefore be the basis for a final decision on diagnosis and treatment of a patient. The laboratories – in collaboration with the clinicians - should find the necessary level of uncertainty to make a rational use of the results possible. A useful basis for such discussions may be found in the report by Ricos et al.[7] The table in the Ricos' publication is based on biological variations and in many cases the laboratories can perform much better than the table indicates.

APPENDIX

Definitions

Metrology – the science of measuring – requires an exact vocabulary based on globally agreed definitions. Several international groups that are actively involved in measurements have published a standard document. This is the *International Vocabulary of Basic and General Terms in Metrology (VIM)*, first published by ISO in 1993. A revised 3rd edition will most likely appear during 2007. Other standards and recommendations follow the VIM, e.g. ISO 5725 and the CLSI documents published after 2003.

Ideally the terms should be understood intuitively but also be useful in all disciplines that deal with measurements. This is not always possible and moreover there may be difficulties to intuitively accept the difference in meaning of a word in scientific and everyday languages.

The definitions may be difficult to read at first sight, partly due to the ambition to create generally applicable definitions. The vocabulary has been translated into major languages but the official languages are English and French.

The ISO format of the entries has been retained and the terms are sorted in alphabetical order.

Accuracy: closeness of agreement between the result of a measurement and a true value.

Combined standard uncertainty: standard measurement uncertainty that is obtained from the measurement results of the input quantities in a measurement function.

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.

Coverage factor: number larger than or equal to one by which a combined standard measurement uncertainty is multiplied to obtain an expanded measurement uncertainty.

Expanded measurement uncertainty: product of a combined standard measurement uncertainty and a factor larger than the number one.

Intermediate precision condition: condition of measurement in a set of conditions that includes the same measurement procedure, same location, and replicated measurements on the same or similar objects over an extended period of time, but may include other conditions involving changes.

NOTE
1. A specification should contain the conditions changed and unchanged, to the extent practical.
2. In chemistry, the term' inter-serial intermediate precision condition of measurement' (between series imprecision) is sometimes used to designate this concept.

Quantity: property of a phenomenon, body, or substance to which a number can be assigned with respect to a reference.

Repeatability: property of a measuring system to provide closely similar indications for replicated measurements of the same quantity under repeatability conditions.

Measurand: quantity intended to be measured.

Measurement accuracy: closeness of agreement between a measured quantity value and a true value of a measurand.

Measurement bias: systematic measurement error or its estimate, with respect to a reference quantity value.

Measurement error: measured quantity value minus a reference quantity value.

Measurement precision: closeness of agreement between indicators obtained by replicate measurements on the same or similar objects under stated specified conditions.

NOTE
Measurement precision is usually expressed numerically by measures of imprecision, such as standard deviation, variance, or coefficient of variation under the specified conditions of measurement.

Measurement repeatability: repeatability, measurement precision under the set of repeatability conditions of measurement

Measurement reproducibility: reproducibility, measurement precision under reproducibility conditions of measurement.

Measurement traceability: property of a measurement result whereby the result can be related to a stated reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty.

Measurement trueness: closeness of agreement between the average of an infinite number of replicate measured quantity values and a true value of the measurand.

Measurement uncertainty: parameter characterizing the dispersion of the quantity values being attributed to a measurand, based on the information used.

Random error: component of measurement error that, in replicate measurements, varies in an unpredictable manner.

Repeatability condition: condition of measurement in the set of conditions that includes the same measurement procedure, same operators, same measuring system, same operating conditions and same location, and replicated measures on the same or similar objects over a short period of time.

Reproducibility condition: condition of measurement in a set of conditions that includes different locations, operators, measuring systems, and replicated measurements on the same or similar objects.

Standard uncertainty: measurement uncertainty expressed as a standard deviation.

Systematic error: component of measurement error that in replicates measurements remains constant or varies in a predictable manner.

Type A evaluation of measurement uncertainty: evaluation of a component of measurement uncertainty by a statistical analysis of quantity values obtained under defined conditions of measurement precision.

Type B evaluation of measurement uncertainty: evaluation of a component of measurement uncertainty determined by means other than a Type A evaluation of measurement uncertainty.

Uncertainty budget: statement of a measurement uncertainty, of the components of that measurement uncertainty, and of their calculation and combination.

REFERENCES

1. *Guide to the expression of uncertainty in measurement*. Geneva: ISO; 1993.
2. Eurachem/Citac. Guide quantifying uncertainty in analytical measurement 2000. Available from: URL: *http//: www.eurachem.ul.pt* (visited 11-03-07).
3. EA. Guidelines on the expression of uncertainty in quantitative testing (EA 4-16). 2004 Available from: URL: *http//: www.european-accreditation.org* (visited 11-03-07).
4. Australasian Association of Clinical Biochemists. Uncertainty of measurement in quantitative medical testing 2004. Available from: URL: *http//: www.AACB.au* (visited 11-03-07).
5. Kallner A, Khorovskaya L, Pettersson T. A method to estimate the uncertainty in measurements in a conglomerate of laboratories/instruments. *Scand J Clin Lab Med.* 2005; 65: 551-567.
6. Clinical and Laboratory Standards Institute. User Verification of Performance for Precision and Trueness; Aproved Guideline – Second Edition. CLSI document EP15-A2(e) [ISBN 1-56238-574-7].
7. Ricos C, Álvarez F, Cava F, García-Lacio JV, Hernández A, Jiménez CV, et al. Current databases on biological variation; pros, cons and progress. *Scan J Clin Lab Invest.* 1999; 59: 491-500. Table updated on the homepage of Dr J Westgard, available from: URL: HTTP//:www.westgard.com (visited 11-03-07).