Trends in Analytical Chemistry 60 (2014) 16–24

Contents lists available at ScienceDirect

Trends in Analytical Chemistry

journal homepage: www.elsevier.com/locate/trac

On the development of quality assurance

Jens E.T. Andersen *

Department of Chemistry, Technical University of Denmark, Kemitorvet Building 207, DK-2800 Kgs. Lyngby, Denmark

ARTICLE INFO

Keywords:
Certified reference material
Chemical analysis
Competence
Consensus science
Interlaboratory comparison
Method validation
Outlier
Precision
Quality assurance
Uncertainty

ABSTRACT

Contemporary research in quality assurance indicates that large uncertainties observed in interlaboratory comparisons to a large extent originate from a lack of competence of laboratory staff. This explanation is challenged by the present article for which six technologies and multiple series of experiments were investigated with respect to uncertainty of measurement and treatment of outliers. It was found that long-term precision was poor in comparison to short-term precision. The ratio of predicted uncertainty to observed uncertainty was determined as significantly above 1, and it was suggested that a correction factor is needed for the predicted uncertainty. This indicates that statistical control could be obtained only by treating many independent series of experiments and using pooled calibrations in the method validation. Retention of outliers in calculation of contents and calculation of uncertainties of certified reference materials (CRMs) gave results that differed significantly from those in the certificates of the CRMs.

Contents

1. Contemporary recommendations

Many scientifically intriguing problems arose after discovery of both the magnitude and the high frequency of discrepancies between results of interlaboratory comparisons (ILCs) [1–7]. In an attempt to account for differences between mean values and uncertainties of measurements, it was realized that fundamentals of metrology [8–11] and quality assurance [12–19] need further development. It was also realized that mathematical methods of statistics should be adapted to analytical chemistry or re-interpreted in terms of basic parameters that are associated with operational calibrations [20–25]. Several international organizations have been involved in the process of developing metrology and quality assurance and social responsibility for analytical chemistry [26]. Methods published by the International Union of Pure and Applied Chemistry (IUPAC) [27] have been implemented in structures of the International Standards Organization (ISO) that has provided a range of standards for industry [28]. *Bureau International de Poids et Mesures* (BIPM), Eurachem, Cooperation on International Traceability in Analytical Chemistry (CITAC) and Nordtest published several guidelines [12,23,29] to support accreditation and method validation according to, e.g., the International Conference on Harmonization (ICH) [16]. Despite the success of the methodology and good intentions, dissemination to a wider audience is a time-consuming process.

Introduction of traceability [18] and uncertainty budgets [12,23] give the researcher an opportunity to establish a complete overview of all uncertainties involved in the process of chemical analysis and to explain those uncertainties based on fundamental units of the SI system, such as meter and kilogram [30]. BIPM

* Tel.: +45 4525 2348; Fax: +45 4588 3136
E-mail address: jeta@kemi.dtu.dk.

http://dx.doi.org/10.1016/j.trac.2014.04.016
0165-9936/

introduces a list of concepts and terms of metrology in the guide International Vocabulary of Metrology, also denoted as Vocabulary in Metrology (VIM3), published by the Joint Committee for Guides in Metrology (JCGM); there, the concepts are explained in greater detail and new concepts are promoted [10]. The concept of analyte is abandoned and is substituted with measurand, that is the species that is intended to be measured [12,29].

In the area of metrology, there was also introduced a new standard that is specially designed for the use in analytical chemistry: the certified reference material (CRM), also frequently denoted standard reference material (SRM). The content can only be estimated by making measurements by different types of apparatus that are all supposed to provide the same result if an infinite number of measurements were performed [10]. However, for practical purposes, everything less than "infinite" is possible and the problem of assessing the number of repetitions required to obtain a reliable value is not fully elucidated; it has been suggested that it is possible to establish reliable levels of uncertainty with a low number of repetitions [31].

One of the main problems in QA is the assessment of a useful value for uncertainty of measurement. The Eurachem/CITAC Guide to Quantifying Uncertainty in Analytical Measurement (QUAM) [12] suggests that all uncertainties of every single unit operation must be included in the total standard uncertainty. An overview of all contributions should be created in the shape of an uncertainty budget that constitutes an integral part of method validation [32] that is fundamental to quality assurance. Further, a standard operating procedure (SOP) [33] is required in order to describe all unit operations whose uncertainties are supposed to be included in the uncertainty budget. These procedures are all very healthy for evaluation of traceability and to laboratory production, but the corresponding method validation (ICH) [16] becomes very demanding on resources.

Measurement of uncertainty by the methods of IUPAC [27] and ISO [28] is to some extent based upon the same statistical methods as those in QUAM [12] but with less focus on traceability and uncertainty budgets. The total measurement uncertainty of the uncertainty budget, which is denoted as the expanded uncertainty [12,23], cannot explain the large deviations between uncertainties found in ILCs [1,2,4–7,34]. These same deviations cannot be explained by the IUPAC and ISO methods, so it is important to consider some of the potential drawbacks of both IUPAC/ISO and Eurachem, which may need further investigations. Specifically, a few potential drawbacks of ISO 5725 [28] and related methods may be summarized, as follows:

- uses complicated mathematics;
- uses the concept of error as a measure of statistical spread in data;
- promotes weighted regression;
- promotes rejection of outliers;
- predicts a centroid value that has never been verified experimentally;
- assumes that many repetitions automatically provide good accuracy;
- does not distinguish clearly between predicted and measured uncertainty; and,
- assumes homoscedasticity of variance that has never been verified experimentally.

The complicated mathematics may exclude some users from understanding the concepts and including them in their everyday work. Mathematics must be simplified in order to secure wide dissemination, and introduction of weighted regression should be accompanied by examples or case studies that explain the importance of such procedures using real data [22]. A result should not be delivered to a customer with a weighted uncertainty because it prevents comparison with results of other laboratories. Similarly, it is unreasonable to reject outliers from a data set if the rejection was based entirely on statistical methods [35], such as Grubb's test for outliers or Cochran's test for outliers [28]. Rejection of outliers requires that the average value remains unaffected and that another laboratory rejects the same outliers in order to arrive at the same average value, which is impossible in practice.

Despite the success and the progress of procedures presented in QUAM [12], there are a number of features that are unattractive, inconvenient or unnecessary, as QUAM:

- uses two types of uncertainties: type A and type B;
- introduces triangular and square distributions;
- has no formal policy on treatment of outliers;
- has no information on the central-limit theorem and the number of repetitions;
- does not distinguish clearly between predicted and measured uncertainty;
- omits uncertainty of long-term variations in the uncertainty budget; and,
- assumes that many repetitions automatically provide good accuracy.

Since the introduction of QUAM [12], scientists no longer operate with systematic uncertainty (systematic error in old terminology) but they treat only uncertainty that originates from random variations introduced by preparing samples, solutions and measurements. QUAM recommends measured uncertainty (Type A uncertainty) and estimated uncertainty (Type B uncertainty). Should bias of a minor magnitude, such as bias of standard addition, be ignored or become relegated to random uncertainty? However, results of unknown uncertainty, which are supposed to be constituents of the uncertainty budget, are evaluated by triangular or square distributions that lead to adjustments by a factor of 0.41 or 0.58, respectively [12,23]. In any case, triangular distribution and square distribution approximations that are supposed to alleviate the effort of estimating uncertainties have the opposite effect; users are faced with problems of deciding whether one or the other factor should be applied to their analysis of uncertainty, thus losing the main focus of their investigation, which is to ensure correspondence with reality. This is the principle of scientific methodology; theory must comply with experiment. Otherwise, the result is invalid. Exactly the same can be stated about uncertainty of measurement; predicted uncertainty must equal measured uncertainty [25,36].

Accuracy is the aim of analytical chemistry, not precision. Occasionally, the content of a sample may be well known (e.g., in inorganic chemistry, dissolution of high-purity silver in ultrapure nitric acid produces a solution of well-known concentration). However, this does not automatically mean that we measure this particular concentration with our apparatus, such as atomic absorption spectrometry (AAS), inductively-coupled plasma optical emission spectrometry (ICP-OES) or ICP-mass spectrometry (ICP-MS). With the current state of analytical chemistry and quality assurance, it is guaranteed that it is possible to obtain significantly different values obtained by the apparatus and the values may also differ significantly from the known quantity value [1,2,4–7,34]. Thus, every single apparatus may possess its own inherent level of accuracy and it is a task of science to investigate the origin of possible bias in order to improve the performance of the apparatus and to make correct decisions.

Proximity to the known quantity value is denoted as trueness, which may be represented by a numerical number while accuracy is not associated with a corresponding number [10]. Accuracy cannot be given a number because the true quantity value of a sample is unknown.