

Investigating out-of-specification test results of chemical composition based on metrological concepts

Ilya Kuselman · Francesca Pennechi ·
Cathy Burns · Aleš Fajgelj · Paolo de Zorzi

Received: 21 July 2009 / Accepted: 6 November 2009 / Published online: 22 November 2009
© Springer-Verlag 2009

Abstract A metrological background for investigating out-of-specification (OOS) test results of chemical composition is discussed. When an OOS test result is identified, it is important to determine its root causes and to avoid reoccurrence of such results. An investigation of the root causes based on metrological concepts would be beneficial. It includes (1) assessment of validation data of the measurement process, (2) evaluation of the measurement uncertainty contributions, and (3) assessment of metrological traceability chains critical for measurement parameters and environmental conditions influencing the test results. The questions, how can the validation data be applied for this investigation, and how can measurement uncertainty contributions and/or metrological traceability chains change a probability of OOS test results, are analyzed.

Keywords Out-of-specification test results · Validation · Measurement uncertainty · Metrological traceability · Probability

Introduction

In pharmaceutical industry, out-of-specification (OOS) test results are results that (after rounding off) fall outside the specifications of established acceptance criteria [1]. By analogy, measurement or test results obtained in other industries and such fields as environmental and/or food analysis, which do not comply with regulatory, specification or legislation limits, can be named also OOS test results.

The problem of OOS test results was known for analysts working in quality control laboratories since the 1920s, but it was not understood until the 1990s that a lack of statistical and metrological thinking is the main aspect of the problem [2]. In 1993, Barr Laboratories (a generic-drug manufacturer) was sued by US government regarding a set of issues influencing the product quality, including the way the company dealt with OOS test results. Among the issues were averaging OOS with in-specification test result values to get a passing result, conducting multiple retests with no defined end point, performing inadequate failure investigations, maintaining an ineffective program for process validation and lacking analytical method validation, etc. Judge Wolin's ruling (the Barr Decision) was that following an OOS test result, an investigation must be initiated before any retesting can be done [3, 4].

Identifying OOS test results is described in the FDA Guidance [1] as the laboratory (Phase 1) investigation. It includes responsibility of the analyst and his or her supervisor, conditions of the testing in the laboratory, etc.

I. Kuselman (✉)
The National Physical Laboratory of Israel (INPL),
Givat Ram, 91904 Jerusalem, Israel
e-mail: ilya.kuselman@moital.gov.il

F. Pennechi
Istituto Nazionale di Ricerca Metrologica (INRIM),
73 Strada delle Cacce, 10135 Turin, Italy

C. Burns
Food and Drug Administration (FDA), 6th Ave and Kipling St,
DFC-Bldg 20, Denver, CO 80225, USA

A. Fajgelj
International Atomic Energy Agency (IAEA),
Wagramer Strasse 5, P.O. Box 100, 1400 Vienna, Austria

P. de Zorzi
Istituto Superiore per la Protezione e la Ricerca Ambientale
(ISPRA), Via Castel Romano, 100, 00128 Rome, Italy

In this guidance, a failure of the quality control in the analytical laboratory is not classified as an OOS test result.

OOS test results from a statistical point of view are discussed in the paper [5], false-negative and false-positive decision errors when assessing compliance of a test result with specification and statutory limits—in ref. [6], identifying OOS test results using Shewhart control charts—in ref. [7].

Currently, the majority of analysts realize that the measurement uncertainty concept is very important because of necessity to balance the cost of measurements versus the product quality risk [8, 9]. For example, to assess compliance of a test result within legislation limits in food and feed in Europe, the analyst should report not only an analyte concentration, but also the associated measurement uncertainty [10]. The value obtained by subtracting the uncertainty from the reported concentration is used to assess compliance with the upper legislation limit. When the compliance assessment is made on the basis of a measurement result accompanied by information on the uncertainty associated with the result, the rules developed in the EURACHEM/CITAC Guide [11] are applicable for identifying OOS test results. Similar rules are included in the ILAC Guidelines [12]. JCGM Guide on the role of measurement uncertainty in conformity assessment is under development [13].

After identification of an OOS test result, it is important to determine its root causes with the purpose to avoid any repetition of the occurrence when the appearance of a next OOS test result is possible or even inevitable. The FDA Guidance [1] formulates recommendations for such incidences including production process review, additional laboratory testing using a pre-defined procedure, reporting testing results, and concluding the investigation with identification of the root causes. Thus, this document establishes an empirical organizational approach to the full-scale (Phase 2) investigation and decisions which can be accepted at the different stages of this investigation.

Another approach, based on metrological concepts and amplifying recommendations of the FDA Guidance for Phase 2 or the full-scale investigation, was initiated in a new joint project of the IUPAC Interdivisional Working Party on Harmonization of Quality Assurance [14] and CITAC [15]. Principles of this approach are discussed in the present position paper of the project task group.

Hypotheses on a product quality and OOS test results

Not depending on knowledge in mathematical statistics, a decision maker on product quality should oppose null hypothesis H_0 that the quality is satisfactory, and an

alternative hypothesis H_1 about unsatisfactory product quality [16]. For example, when an upper specification limit is discussed, there are $H_0: c_{\text{true}} \leq c_{\text{upp.s.l}}$ against $H_1: c_{\text{true}} > c_{\text{upp.s.l}}$, where c_{true} is the true value of the concentration c of an analyte in the lot or batch of the product, and $c_{\text{upp.s.l}}$ is the upper specification limit for the analyte concentration. The true value is unknown and decisions are made using test results c_{test} . Distribution of c_{true} values in different product batches and corresponding distributions of c_{test} values for two of these batches are illustrated in Fig. 1 as probability density functions (pdf) $f(c)$, truncated normal for simplicity. Centers of the distributions are shown by dotted lines and marked with $c_{\text{true-av}}$ and $c_{\text{test-av}}$, respectively. The index “av” means average or mean, when the number of batches and/or the number

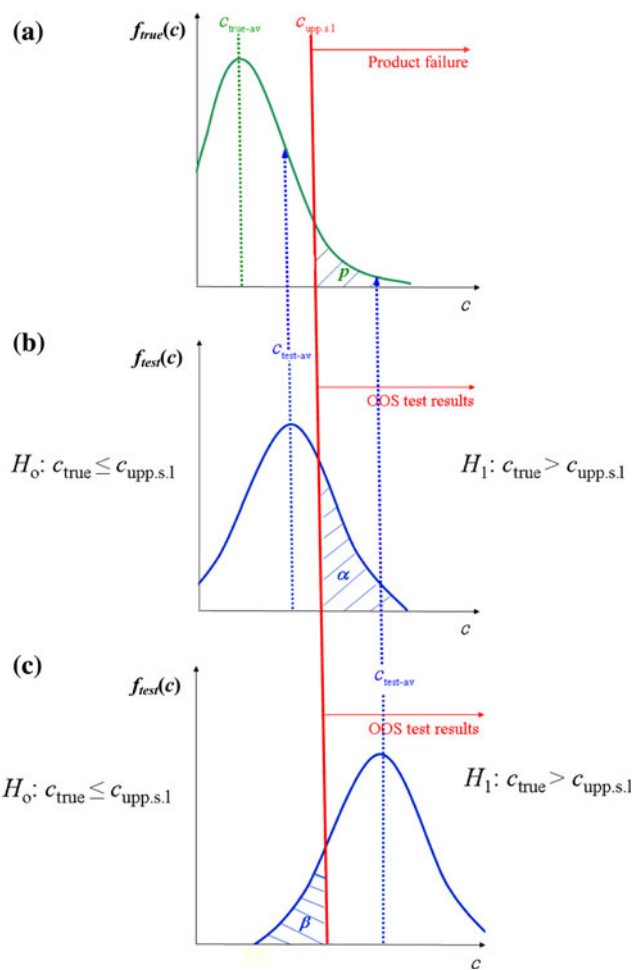


Fig. 1 OOS test results, producer’s risk α and consumer’s risk β . H_0 and H_1 are the null and the alternative hypotheses on the product quality, respectively; c is the analyte concentration in the product; c_{true} , $c_{\text{upp.s.l}}$ and c_{test} are the c true value, the upper specification limit and the test result, respectively; $f_{\text{true}}(c)$ is pdf of c_{true} values in **a**, and $f_{\text{test}}(c)$ is pdf of c_{test} values for two product batches under testing in **b** and **c**, respectively; $c_{\text{true-av}}$ and $c_{\text{test-av}}$ are the average values of c_{true} and c_{test} , respectively; p is the probability of the product failure

of tests are large enough. The dotted lines $c_{\text{test-av}}$ reach the true values c_{true} of the analyte concentration in the particular product batches under testing, i.e. considered coinciding with them. The upper specification limit is represented by solid vertical line. The shaded area under the $f_{\text{true}}(c)$ curve to the right side of $c_{\text{upp.s.l}}$ in Fig. 1a, shown by pointer, equals to the probability p of the product failure. OOS test results $c_{\text{test}} > c_{\text{upp.s.l}}$ are shown by pointers in Fig. 1b, c. The shaded area under the $f_{\text{test}}(c)$ curve in Fig. 1b is the probability α of type 1 error in the decision on the product quality. This error, named also “false-positive”, appears when $c_{\text{true}} = c_{\text{test-av}} \leq c_{\text{upp.s.l}}$, while $c_{\text{test}} > c_{\text{upp.s.l}}$ and hypothesis H_1 is not rejected. Probability α of type 1 error is the producer’s risk. To decrease probability α the batch testing can be repeated. However, when abused, it becomes “reflexive retesting” [2].

Type 2 error in the decision on the product quality, named also “false-negative”, is possible when product failure is analyzed, $c_{\text{true}} = c_{\text{test-av}} > c_{\text{upp.s.l}}$, while $c_{\text{test}} \leq c_{\text{upp.s.l}}$ and hypothesis H_0 is not rejected. This situation is illustrated in Fig. 1c. The shaded area under the $f_{\text{test}}(c)$ curve to the left side of $c_{\text{upp.s.l}}$ is probability β of type 2 error. It is the consumer’s risk. A decision maker lacking a conscientious philosophy is able to produce a passing test result (to increase probability β) by retesting and looking for type 2 error, i.e. by “testing into compliance” [2].

Any OOS test result can indicate a product failure, or be caused by measurement/analytical (metrological) problems. Thus, probability α of type 1 error in decision based on OOS test result is the metrologically related producer’s risk. Since the product should be rejected when investigation of OOS test result does not support the product quality, the consumer’s risk (probability β of type 2 error) is practically absent in such a case.

When a number of produced batches is statistically significant, the distribution shown in Fig. 1a can be approximated by the test result distribution (from batch to batch). Since the distribution of test results for one batch (like in Fig. 1b, c) is known, as a rule, the probability p of the product failure can be evaluated as proposed in the JCGM Guide draft [13].

Using this scheme of the distributions and the decisions, the full-scale OOS test result investigation can be directed at the analysis of the main metrological factors influencing the distribution parameters and probabilities of type 1 and type 2 errors.

Metrological approach

When a result of testing [17] is quantitative and equal to the measurement result, the metrological approach requires

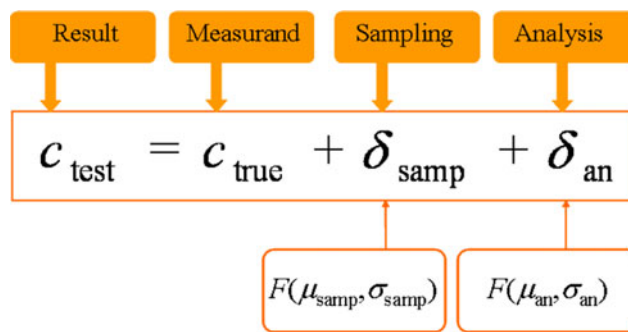


Fig. 2 Model of the measurement result. Values δ_{samp} and δ_{an} are the contributions of the test result c_{test} caused by sampling and analysis, respectively; F is the distribution function of δ_{samp} and δ_{an} with centers μ_{samp} and μ_{an} , and variances σ_{samp}^2 and σ_{an}^2 , respectively

defining first of all the measurand, i.e. the quantity intended to be measured [18]. In an analytical quality control laboratory, it is the amount-of-substance concentration c_{true} of the analyte in a product batch. A model of the analytical measurement/test result c_{test} shown in Fig. 2 includes contributions caused by two stages of testing: sampling δ_{samp} and analysis δ_{an} . Distributions functions F associated with these contributions can be very different [19]. However, for well studied and widely used measurement (sampling and analysis) methods they are mostly normal or can be transformed into normal. Values μ_{sample} and μ_{an} in Fig. 2 are the means of the distributions associated with the measurement errors δ_{samp} and δ_{an} , respectively, while σ_{sample}^2 and σ_{an}^2 are the corresponding variances. Taking into account this model, the full-scale investigation of OOS test results based on the metrological concepts should include:

1. assessment of validation data for sampling and chemical analysis,
2. evaluation of the measurement uncertainty contributions from different stages of the test,
3. assessment of the metrological traceability chains important for the measurement parameters and environmental conditions influencing the test results.

Flow chart of this approach is presented in Fig. 3. The chart leads from investigation of the measurement method and the sampling procedure with their validation data to the metrological requirements to the measurement results,

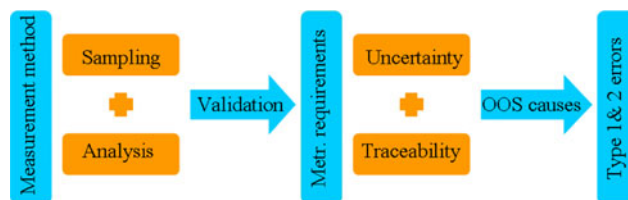


Fig. 3 Flow chart of the metrological approach

associated uncertainty and traceability chains; then to causes of OOS test results and their influence on probabilities of errors in decisions on the product quality.

The metrological approach developed in this project is not dealing with cases of qualitative (e.g. organoleptic) testing and human errors. If a product failure is caused by technological problems, this approach cannot be also directly useful. However, when a significant measurement uncertainty arising from sampling is identified, an optimization of the technological parameters may be required to increase the product homogeneity.

Assessment of validation data

Validation is widely used procedure in pharmaceutical industry. There are the FDA draft guidance for industry process validation including validation of sampling procedures [20], the ICH guideline for validation of analytical procedures [21], recommendations for analytical method and measuring equipment validation [22], etc. In other industries and analytical fields, validation is regulated by EURACHEM guide [23], AOACI validation programs and other national and international documents [24].

The most common validation parameters are repeatability, reproducibility, trueness and bias, limit of detection, selectivity and sensitivity, defined in the international vocabulary of metrology [18], as well as linearity and limit of quantification [21], robustness and ruggedness [25].

Investigating OOS test results, one should verify, where the specified requirements and the validation data are adequate for an intended use. The absence of adequacy can be a root cause of the OOS test results. Another question to be checked, are the validation data complete enough to evaluate contributions of the associated measurement uncertainty?

Evaluation of measurement uncertainty contributions

Measurement uncertainty is non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand, based on the information used [18]. Evaluation of measurement uncertainty can be done using repeatability, reproducibility and trueness estimates from the validation data [26]. A number of examples of uncertainty calculation in the field of environmental analysis are available in the handbook [27]. Other methods for quantifying uncertainty in analytical measurement are described in the guide [28]. Methods and approaches for evaluating measurement uncertainty arising from sampling are discussed in the guide [29].

The current situation in the pharmaceutical industry is that the traceability of analytical measurement results to

certified values of the reference standards is required, still without evaluating their associated measurement uncertainties [30]. However, an evolution of understanding of metrological concepts in this field is already started [31]. The first application of measurement uncertainty concept to USP's compendial reference standards, as certified reference materials, is published in the paper [32] awarded recently by CITAC [33].

The two most important measurement uncertainty aspects and questions in the full-scale investigation of OOS test results are: (1) is the test result value significantly larger than the associated measurement uncertainty? and (2) are the measurement uncertainty contributions of the same order? Any negative answer on one or both these questions can indicate a cause of the OOS problem. If a dominant contribution is detected while answering the second question, this contribution should be thoroughly studied. Figure 4 illustrates how the ratio between the contributions influences the “picture” of OOS test results and corresponding producer's risk α . For simplicity, the sampling and analysis contributions in Fig. 4 have normal distributions with centers μ_{sample} and μ_{an} equal to zero and variances σ_{samp}^2 and σ_{an}^2 . At these assumptions $c_{\text{test-av}} = c_{\text{true}}$ for the product batch under testing in both cases described below. Axis c is expressed in units of σ_{an} , while $\sigma_{\text{an}} = 1$, i.e. $c/\sigma_{\text{an}} = c$; $c_{\text{test-av}} = 10\sigma_{\text{an}} = 10$, $c_{\text{upp.s.1}} = 13\sigma_{\text{an}} = 13$. Curve 1 is the pdf for case 1, when $\sigma_{\text{samp}} = 1/3\sigma_{\text{an}} = 1/3$ and combined standard uncertainty $u = (\sigma_{\text{samp}}^2 + \sigma_{\text{an}}^2)^{1/2} = 1.05$. Curve 2 is the pdf for case 2, when $\sigma_{\text{samp}} = 3\sigma_{\text{an}} = 3$ and $u = 3.16$. The values $\alpha_1 = 0.002$ and $\alpha_2 = 0.171$ are the producer's risks for cases 1 and 2, respectively. It is clear, that the measurement

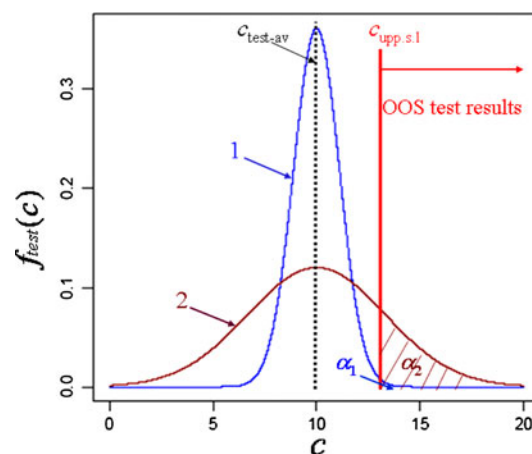


Fig. 4 OOS test results and producer's risks at different ratios of the measurement uncertainty contributions. Axis c is expressed in units of σ_{an} , while $\sigma_{\text{an}} = 1$; $c_{\text{test-av}} = 10\sigma_{\text{an}}$, $c_{\text{upp.s.1}} = 13\sigma_{\text{an}}$; curve 1 is the pdf for case 1, when $\sigma_{\text{samp}} = 1/3\sigma_{\text{an}}$; curve 2 is the pdf for case 2, when $\sigma_{\text{samp}} = 3\sigma_{\text{an}}$; $\alpha_1 = 0.002$ and $\alpha_2 = 0.171$ are the producer's risks for cases 1 and 2, respectively

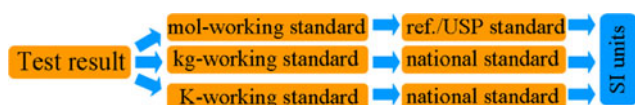


Fig. 5 Metrological traceability requirements

uncertainty arising from sampling in case 2 is the cause of the OOS test results and of the significant producer's risk.

Assessment of metrological traceability chains

Metrological traceability is a property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty. Traceability chain is a sequence of measurement standards and calibrations that is used to relate a measurement result to a reference [18]. There are the EURACHEM/CITAC guide [34] on this topic, the IUPAC project on metrological traceability of measurement results in chemistry [35] under development, and other documents and publications [36].

Assessment of metrological traceability chains is important for measurement parameters and environmental conditions influencing the test results. For example, traceability chains to SI units of measurement results of quantity of matter (mol), of mass (kg) and of temperature (K) as shown in Fig. 5, should be realized for practically every chemical test. The reason is that a test portion is quantified by mass, measuring instruments are calibrated by certified reference materials, and temperature is to be under control. In particular, the pharmaceutical industry's practice of using a one-point calibration raises questions regarding the traceability chain of the measurement result to mole. This calibration consists of comparison of responses of the measuring system obtained for the test portion and a working standard. The working standard is certified by comparison with a USP's or other reference standard. Commutability [18] between the reference and the working standards, and adequacy [37] of the working standard to the substance under analysis should be a point for investigation, first of all for impurities and degradation products.

Any broken metrological traceability chain can lead to OOS test results.

When investigation of an OOS test result based on the metrological concepts does not identify or explain the reason of the event, an investigation of the production process (technological parameters) is necessary.

Conclusions

1. Investigation of OOS test results based on the metrological concepts should answer the following

questions: (a) are the specified requirements and the validation data of the testing process (including sampling and chemical analysis) adequate for an intended use and for evaluation of contributions of the associated measurement uncertainty? (b) on which stage of the testing process the measurement uncertainty contributions are dominant and could cause the OOS test results? (c) which metrological traceability chains are broken and could also cause the OOS test results?

2. The metrological approach to investigation of OOS test results is not dealing with qualitative testing, human errors and technological problems of the product failure.

References

1. US FDA (2006) Guidance for industry. Investigating out-of-specification (OOS) test results for pharmaceutical production
2. Kuwahara S (2007) *BioPharm Int* 20(11):42–52
3. Summary of Judge Wolin's Interpretation of GMP Issues Contained in the Court's Ruling in USA vs Barr Laboratories. GMP Institute, USA (1993) <http://www.gmp1st.com/barrsum.htm>
4. Hoinowski AM, Motola S, Davis RJ, McArdle JV (2002) *Pharm Technol* 26(11):40–50
5. Köppel H, Schneider B, Wätzig H (2007) *J Pharm Biomed Anal* 44:718–729
6. Desimoni E, Brunetti B, Clerici A (2008) *Accred Qual Assur* 13:653–662
7. Nunnaly BK, Nunnaly DF, McConnel JS (2008) *BioPharm Int* 21(6):66–74
8. ICH Q9 (2005) Quality Risk Management
9. ICH Q10 (2008) Pharmaceutical Quality System
10. Report on the Relationship between Analytical Results, Measurement Uncertainty, Recovery Factors and the Provisions of EU Food and Feed Legislation (2004) http://ec.europa.eu/food/food/chemicalsafety/contaminants/report-sampling_analysis_2004_en.pdf
11. EURACHEM/CITAC Guide (2007) Use of uncertainty information in compliance assessment
12. ILAC G8 (2009) Guidance on the reporting of compliance with specification
13. JCGM 106 (2009) Draft guide. Evaluation of measurement data—the role of measurement uncertainty in conformity assessment
14. IUPAC Project 2008-030-1-500 (2009) <http://www.iupac.org/web/ins/2008-030-1-500>
15. CITAC Publications/Documents Under Development (2009) IUPAC/CITAC Project. Investigating out-of-specification test results of chemical composition based on metrological concepts. <http://www.citac.cc>
16. Mittag HJ, Rinne H (1993) *Statistical methods of quality assurance*. Chapman & Hall, London, pp 119–150
17. ISO 17000 (2004) *Conformity assessment—vocabulary and general principles*
18. ISO Guide 99 (2007) *International vocabulary of metrology. Basic and general concepts and associated terms (VIM)*. JCGM 200 (2008) <http://www.bipm.org/en/publications/guides/vim>
19. Kuselman I (2006) *Accred Qual Assur* 10:659–663
20. FDA (2008) Draft guidance for industry process validation. General principles and practices

21. ICH Q2(R1)(2005) Validation of analytical procedures: text and methodology
22. Huber L (1999) Validation and quantification in analytical laboratories. Interpharm Press, Buffalo Grove
23. EURACHEM Guide (1998) The fitness for purpose of analytical methods: a laboratory guide to method validation and related topics
24. De Bievre P, Günzler H (eds) (2005) Validation in chemical measurement. Springer, Berlin
25. Hendriks MMWB, de Boer JH, Smilde AK (eds) (1996) Robustness of analytical chemical methods and pharmaceutical technological products. Elsevier, Amsterdam
26. ISO/TS 21748 (2004) Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation
27. Magnusson B, Näykki T, Hovind H, Kryssel M (2003) Handbook for calculation of measurement uncertainty in environmental analysis. NORDTEST Report TR 537, Nordtest Tekniikantie 12, FIN-02150 Espoo, Finland
28. EURACHEM/CITAC Guide (2000) Quantifying uncertainty in analytical measurement, 2nd edn
29. EURACHEM/CITAC Guide (2007) Measurement uncertainty arising from sampling. A guide to methods and approaches
30. Kuselman I, Weisman A, Wegscheider W (2003) Accred Qual Assur 8:530–531
31. Williams RL (2006) J Pharm Biomed Anal 40:3–15
32. Manning RG, Lane S, Dressman S, Hauck WW, Williams RL (2007) Pharm Forum 33(6):1–13
33. Dressman S, Koch WF, Manning RG, Lane S, Hauck WW, Williams RL (2009) CITAC News 22–24
34. EURACHEM/CITAC Guide (2003) Traceability in chemical measurement. A guide to achieving comparable results in chemical measurement
35. IUPAC Project 2001-010-3-500 (2001) <http://www.iupac.org/web/ins/2001-010-3-500>
36. De Bievre P, Günzler H (eds) (2005) Traceability in chemical measurement. Springer, Berlin
37. Kuselman I (2004) Accred Qual Assur 9:591–596