ELSEVIER

Journal of Chromatography B, 713 (1998) 227–243

JOURNAL OF CHROMATOGRAPHY B

Review

Quality control in toxicological analysis

Davide S. Ferrara*, Luciano Tedeschi, Giampietro Frison, Gianpaolo Brusini

Centre of Behavioural and Forensic Toxicology, Istituto di Medicina Legale, University of Padova, Via Falloppio 50, I-35121 Padova, Italy

Abstract

The validity and effectiveness of quality control procedures are reviewed here in light of the principles of analytical toxicology, and of the professional responsibility involved in awareness of the profound influence which analytical results have in the fields of health and social security. Regardless of the methods used, laboratory work should aim essentially at achieving a very high degree of reliability. Factors contributing to the 'quality' of analytical results and methods used to check their reliability are discussed here. The technical background and organization of internal and external quality control procedures are presented, with particular reference to educational aspects, and to the ways in which computer and internet technologies may be exploited for further improvement of the effectiveness of these procedures.

Keywords: Reviews; Quality control; Proficiency testing program; Drugs of abuse

Contents

*Corresponding author

0378-4347/98/$19.00
PII: S0378-4347(98)00010-3

1. Introduction

Results provided by an analytical laboratory can be used to make administrative and/or legal decisions or to determine diagnosis and therapy, and must be as error-free as possible. For this reason quality assurance (QA) must be the prime objective of any analytical laboratory.

QA has been defined as a ''total integrated management program for assuring the reliability of data'' [1], denoting all the procedures, cognitive and mechanical, designed to minimize or identify all sources of preanalytic, analytic and postanalytic variation which may have an impact upon the provision of high-quality analytical results [2].

On the basis of this definition, an analytical laboratory should construct its own QA on three different levels, encompassing all aspects of the laboratory activities, including administrative functions.

(a) *Logistic–structural:* including laboratory environment, reagents, equipment, instrumentation, personnel and chain-of-custody procedures.

(b) *Analytical–interpretative:* including sampling (collection, selection and storage), specimen preparation and handling, availability of reference materials, selection of analytical methods, review of results by experienced personnel, and reporting formats.

(c) *Control:* including internal quality control and participation in external quality assessment schemes, to verify the effectiveness of QA planned activities for the achievement of error-free data.

There is no single generally accepted plan for establishing a laboratory's QA program. Every discipline has its own specific problems which require special consideration and treatment [3]. In any case, once quality goals have been defined, quality control (QC) procedures must always be designed with the aim of monitoring performance [4–6].

This model is even more true for toxicology laboratories, which are called upon to satisfy epidemiological, diagnostic and therapeutic requirements in various contexts ranging from clinical, workplace, legal and the world of sport, where possible analytical errors may have a great impact on an individual's livelihood, freedom and civil rights. It is, therefore, of paramount importance for an analytical toxicology laboratory to implement the constant use of an 'in-house' QC program to check that procedures used are under statistical control and that the data produced are of absolute reliability.

It should be stressed, however, that control does not necessarily imply quality: control by itself can only be used to monitor the current quality of the process, but it cannot improve the analytical quality properly [7]. In order to improve the quality of analytical procedures, introduction of better methods, equipment and/or improvement of personnel training are needed.

2. Establishing analytical quality

For a toxicology laboratory, the goal of testing may be stated as the reliable demonstration of the presence, or absence, of specified substances in an

analysed sample – that is, production of valid positive or negative results [8]. A quantitative analytical approach for specific substances is also sometimes performed [9]. Analytical quality may, therefore, be stated as the power of an analytical procedure as a whole to determine, with acceptable confidence (on the basis of the intended application), the identity of a particular substance in a given biological matrix. QC procedures can only be effective when all factors related to the creation of analytical quality have been clearly defined and carefully set up.

Toxicological testing usually includes a two-step process: preliminary screening to identify negative samples not containing substances or their metabolites, i.e., samples which do not require further analysis, and subsequent confirmation to identify specific substances in presumptive positive samples. To ensure accuracy and reliability, confirmation must be carried out by techniques based on physicochemical principles different from those used for screening. Immunochemical techniques offer the advantage of being applicable to a wide range of analytes at lower cost and are usually used for screening, whereas chromatographic techniques, providing higher analytical sensitivity and specificity, are mainly used for confirmation.

It should also be mentioned that, in analytical toxicology, every analytical procedure has a cutoff or threshold concentration [10] associated with it, that is, the defined concentration of analyte in a specimen at or above which the test is called positive and below which it is called negative, and which is usually significantly greater than the limit of detection of the assay. A proper cutoff value provides a safeguard to the analysis because it defines a positive or negative result well within the ultimate capability of the analytical method, and therefore helps minimize false results.

On the basis of these considerations, the creation of analytical quality must take into account *internal* factors, directly related to the laboratory's own management, as well as *external* ones. These factors may also be further subdivided into *permanent* (not subjected to batch-to-batch variations) and *variable* (which may be subjected to batch-to-batch changes) [11,12]. Table 1 summarizes this concept. Of course, since analytical quality is only relevant to *analytical results*, the table does not include other important factors related to *reported result*, such as chain-of-custody, review of results, etc., which must be seen in the more general context of the QA system [5,6].

On one hand, in the absence of external objective institutions which publish lists of the quality of commercial products, the quality of external factors is beyond control and each laboratory must try to choose the most appropriate products (reagents, instruments, calibrators) from those available on the market. Data available from external quality assessment schemes which summarize results from par-

Table 1
Factors in establishing analytical quality

Factors	Permanent	Variable
External	Instrumental equipment	Reagents (stability)
	Analytical principle	Reference materials (stability)
	Reagents (choice of source)	Calibrators
	Reference materials (choice of source)	Consumables
Internal	'In house' equipment, reagents, etc.	'In house' reagents (stability)
	Choice of techniques and methods for screening on the basis of:	Calibration functions
	sensitivity	Personnel training
	specificity	Equipment maintenance
	prevalence	Documentation
	predictive values	Environmental/human contamination
	Choice of techniques and methods for confirmation on the basis of:	
	limit of detection	
	limit of quantitation	
	Planned cutoff	
	Standardization of the method	

ticipants according to method and instrumentation may be of some help [12]. On the other hand, the laboratory is directly responsible for its internal factors, first by implementing the most appropriate methods, and later through monitoring analytical quality and detecting errors.

3. Overview of quality control

The practice of QC falls naturally into two categories: activities carried out within a laboratory, independently of external agencies (internal quality control (IQC)) and those undertaken in conjunction with other laboratories and agencies (external quality control (EQC)). Although intra- and inter-laboratory activities have the same overall aim, on the basis of the errors for which they are able to check, the emphasis often differs.

The overall aim of QC for an analytical laboratory is to ensure that its analytical results are of adequate accuracy for their intended application [13]. For this aim, it is necessary to select appropriate QC procedures that will assure that the desired analytical quality is achieved, checking for errors and/or changes in performance related to internal as well as external factors involved in the creation of analytical quality itself.

3.1. Terminology

Because several terms related to QC have been interpreted in different senses, the following is a list of the conventions which will be adopted in this paper [10,14–18]:

True positive (TP): a test result which states that one or more substances are present in the analysed sample when, in fact, they are present in the sample at a concentration greater than the cutoff value.

True negative (TN): a test result which states that no substances are present in the analysed sample when, in fact, they are present in the sample at a concentration less than the cutoff value, or not present at all.

False positive (FP): a test result which states that one or more substances are present in the analysed sample when, in fact, they are present in the sample at a concentration less than the cutoff value, or not present at all.

False negative (FN): a test result which states that no substances are present in the analysed sample when, in fact, they are present in the sample at a concentration greater than the cutoff value.

Sensitivity ($TP\times100/(TP+FN)$): the incidence of TP results obtained when a test is applied to positive samples, i.e., the probability that a positive sample is identified as such by a certain analytical technique.

Specificity ($TN\times100/(TN+FP)$): the incidence of TN results obtained when a test is applied to negative samples, i.e., the probability that a negative sample is identified as such by a certain analytical technique.

False positive rate ($FP\times100/(FP+TN)$): the percentage of samples identified as positive when they are, in fact, negative.

False negative rate ($FN\times100/(FN+TP)$): the percentage of samples identified as negative when they are, in fact, positive.

Prevalence: percentage of positive samples in examined population. This is an 'a priori' probability, pre-existing and independent of results obtained with analytical techniques used for later measurements.

Positive predictive value (PPV): the incidence of correct positive results supplied by a given analytical technique when it is applied to a population including both positive and negative samples:

$$\mathrm{PPV}=\frac{\text{Sensitivity}\times\text{Prevalence}}{((100-\text{Specificity})\times(100-\text{Prevalence}))+(\text{Sensitivity}\times\text{Prevalence})}$$

Negative predictive value (NPV): the incidence of correct negative results supplied by a given analytical technique when it is applied to a population including both positive and negative samples:

$$\mathrm{NPV}=\frac{\text{Specificity}\times(100-\text{Prevalence})}{((100-\text{Sensitivity})\times\text{Prevalence})+(\text{Specificity}\times(100-\text{Prevalence}))}$$

Accuracy: the closeness with which results agree with a known true value of the quantity being measured.

Limit of detection (LOD): the minimum amount of an analyte which can be detected with confidence by a testing procedure;

Limit of quantification (LOQ): the minimum amount of an analyte which can be quantified by a

testing procedure, while conforming to the required coefficient of variation of the procedure.

Confidence: the probability that the diagnostic hypothesis is true.

3.2. QC planning

Successful QC planning does not just happen, it requires a systematic planning process [19] to consider all critical performance information in both an orderly and a quick access manner. This means supporting data calculations, preparing graphical tools and control charts such that trends in results may be visualized, and making it easy to document QC recommendations. Control results must also be fully documented and stored for later inspection. The most quantitative, comprehensive, and flexible approach to this aim is to provide computer support [20]. Another advantage of computer-aided QC is that laboratories are promptly informed of the outcome of a round of the trial so that analysts can immediately take countermeasures if necessary. Computer-aided QC should also be implemented in EQC programs where a common weak point is often identified in their excessive slowness for ongoing monitoring of performances. Some recent Internet-based approaches to EQC have been introduced and will be described later.

3.3. Control samples

The quality of control materials or control samples, common to both IQC and EQC, is of major importance for QC itself. An improper control sample will almost always result in false or unusable data, together with incorrect conclusions. Control samples must be representative of the real test material, and homogeneous so that virtually identical and stable portions can be used over a period of time. Spiked materials (coming from healthy volunteers or non-biological ones) are usually better than artificial materials, on condition that they are stable and reproducible. In the absence of suitable control samples, it may be necessary to replicate analyses of real test materials taken from each analytical run [28].

Particular care should be used when control samples are prepared *in-house* from non-conventional biological matrices, as in the field of hair analysis for drugs of abuse testing, because of the inherent difficulty in incorporating substances being tested in the biological matrix [29].

3.4. Types of errors

Errors related to permanent factors (Table 1) cannot be revealed by any internal control system, so that IQC can only monitor performance related to variable factors, both external and internal. According to the Vocabulary of International Metrology [30], this kind of error mainly refers to random error, which is related to variations (inherent or increasing) in the analytical procedure as implemented in the laboratory. Changes in consumables and/or reagents, contamination, and personnel substitution may all be main sources of this kind of error, which is easily detected by IQC procedures. The same procedures are also useful in detecting systematic errors [30] when they are related to instrumental calibration, which is the same for all analysed samples in a run.

Instead, EQC has a larger element of quality assessment, providing a demonstration of the analytical quality achieved by the laboratory when its analytical results are compared with those obtained by others. Furthermore, the evaluation of systematic errors due to *non-specific reactions* and/or *interfering substances* in specimens, which is of paramount importance in analytical toxicology [31,32], especially when immunochemical tests are performed, may only be monitored by participating in EQC programs.

Both kinds of control have their own targets and inherent limitations and, for proper assessment of analytical quality, the implementation of only one kind of control is not enough.

3.5. Statistical control

Since an analytical process cannot be *directly* measured (unlike industrial processes, where measurement of parameters such as volume, length or speed may be directly obtained), the practice of QC means *indirect* control of performance and the key concept in this respect is that of *statistical control*. Results collected from test controls are plotted versus time or sequential run number on a control chart and

evaluated to see whether the measurement procedure is in-control or out-of-control.

This procedure is applicable only when an analytical system can be said to be in a state of *statistical control*, that is, if it is subjected to sources of variations that are stable. Only in such circumstances is it possible to describe and predict the size of variations likely to affect results [13]. However, it should be emphasized that the achievement of statistical control does not imply that the system is performing within the analytical quality required, but only that it is stabilized. If the analytical procedure is based on erroneous assumptions, the analytical results will be wrong even if no statistical variations are reported by the QC procedures.

3.6. Control charts

This graphical method of displaying control results was first developed in studies of economics [21] and later introduced as a statistical control method in clinical laboratories [22].

Control charts take the form of charts on which the variable of interest is usually plotted on the *y*-axis versus time or sequential run number [23,24]. Every time a new control value is collected as part of routine work, the observed value is added to the control chart. For a stable testing process, the new control measurements should show the same distribution as past ones [25]. The observation of unusual results means that something strange has happened in the analytical procedure.

To facilitate this observation, it is useful to insert on the chart some lines corresponding to the expected value (x), the 'control limits' (usually referred to as $x \pm 2\sigma$, where σ is the standard deviation) and 'warning limits' (usually called $x \pm 3\sigma$). When the control value falls within the control limits, the run can be classified as '*in-control*' and the analytical result can be accepted and reported. When the control value falls outside the warning limits, the run must be classified as '*out-of-control*' and the analytical result should be rejected and not reported.

One result may fall outside these limits and, in this case, requires no action, provided that the next result falls inside them. The observation of two consecutive results outside the warning limits indicates the need for remedial action. Both results falling on the same side of the limit, indicates increased bias; alternatively, if they are on opposite sides, there is stronger evidence of increased random error.

The main function of the chart is to act as a means of detecting *departure* from statistical control, not as a means of assessing the acceptability of accuracy of individual results [26,27].

4. Internal quality control

The main objective of a toxicology test is to determine with acceptable confidence the presence or absence of a particular substance or group of substances in a given sample with respect to a predefined cutoff value. To achieve this, the analyst must define what level of confidence is acceptable – 95% probability, 99% probability, or greater than 99%. The 95% confidence limit is accepted as a general standard in clinical laboratories. Civil law applications may require a confidence level equal to or greater than 95%. Applications in criminal law cases may require confidence levels greater than 99%. The power of the analytical method to meet these requirements may be stated as the '*analytical quality*' of the test and must be carried out on the basis of parameters such as sensitivity, specificity, predictive and cumulative predictive values, speed and cost [17].

The function of IQC is to detect changes in performance during routine operations compared with performance reported when the analytical quality of the test has been carefully set up. For this aim, control samples with usually known compositions are repeatedly tested using the routine analytical procedure to generate an objective numerical description of its operating state.

Control samples must be treated in the same way as routine samples, so that if the analytical procedure requires pretreatment techniques, controls must be pretreated too.

Furthermore, because analysts tend to be more cautious in their approach to known control samples with respect to routine samples, often repeating the analysis if results do not fit their expectations, *blind* QC samples should be included in the run and should appear as normal samples to laboratory analysts [33]. However, it is a matter of choice as to whether or not

the control material should be identifiable as such by the analyst [34,35], because the introduction of 'secrecy' into IQC may be counterproductive in that it can lead to the feeling that the analyst is the subject of the control test rather than the analytical procedure. Therefore, when good control is achieved over a period of time, the frequency of such samples may be reduced.

When an IQC process is implemented, it should monitor every part of the analytical procedure and, as noted previously, since in analytical toxicology a two-step analytical process is usually adopted (screening-confirmation) a double quality control approach is also necessary.

4.1. Screening tests

As already stated, most screening tests are qualitative but, since they must be compared with a predefined cutoff value, they do have their quantitative aspect [36]. Indeed, when a sample is tested for the presence of a given analyte and the result is negative, it is good practice to report that ''the analyte was not present at concentrations at or above the cutoff level'', instead of the generally used ''the sample does not contain the analyte''. It is this *quantitative* component of screening which must be tested for QC purposes.

4.1.1. Control of screening test

Each analytical run of specimens to be screened must include: (a) samples certified to contain no analyte(s) of interest (*negative controls*); (b) samples fortified with known concentrations of analyte(s) of interest (*positive controls*); (c) positive controls containing the analyte(s) of interest with concentrations equal to or near the cutoff value.

Furthermore, procedures to ensure that carryover does not contaminate the testing of an individual specimen should be implemented, for example, adding negative controls after known positive controls.

The number of positive and negative controls to include in each analytical run partly depends on the accuracy of the method used and the quality goals of the testing procedure. Because of the high quality required for toxicological analysis, a relatively high number of controls should be used for each run [37]. For example, the standard operating procedure for US military drug testing laboratories stipulates that 20% of samples in an immunochemical screening test must consist of standards and controls: thus, in a run of 200 samples, 10 must be negative controls, 5 positive controls, and 25 standards in various concentrations [38,39]. Federal workplace drug-testing programs in the USA requires that a minimum of 10% of all test samples should be QC samples [33].

4.1.2. Evaluation of screening test control

For every known standard concentration added to the analytical run, a single control chart should be prepared (or just one control chart reporting all concentrations). When the control sample is analysed the analytical result is plotted on the control chart and the difference between obtained and expected values must be evaluated as described in Section 3.6. If values falling out of 'warning limits' are evidenced, remedial action must be taken.

Negative control samples must be examined with particular attention, especially in forensic toxicology, because the need to minimize FPs as much as practicable is well recognized. When a negative control sample (considered free from interfering substances) is identified as positive, it is important to check for other negative samples added to the same run (or to check other blank samples if the positive was the last one). If no other positive results are reported, the aberrant value was probably due to a negative random event (e.g., increased analytical background noise). Instead, consecutive positive responses, obtained from negative controls placed after known positive controls, may be regarded as an indication of analytical carryover. In both cases, the analytical results of the run should be reported only after careful confirmation by other techniques, and a check of the analytical procedure may be needed.

Due to the existence of the cutoff value, in analytical toxicology the evidence of positive results for negative control samples must always be considered with particular care. Indeed, the amount of positive response coming from random or inherent variability of the analytical method, even if of negligible concentration, may increase the *real* concentration of the analyte and therefore be particularly insidious for the proper classification of an analytical

result as being a TP or FP compared with a cutoff value [40,41].

4.1.3. Evaluation of prevalence

The analytical quality of a given technique used in a screening test is assessed by considering parameters such as the sensitivity, specificity and predictive values of the technique itself. However, an important point to focus on is that predictive values have been shown to be related to the prevalence of the analytes in the samples being tested [14]. Test procedures with excellent predictive values are reliable predictors of the presence of a particular substance in samples related to a population of subjects in which the prevalent use of this specific substance is high but, if this prevalence varies, the related parameters will also vary [15].

It is therefore important to implement periodic monitoring of the prevalence of the substances analysed by a laboratory, because its variations may reflect on the analytical quality of the procedure. This kind of check is especially important for laboratories which perform regular and frequent analyses of well-defined classes of substances (e.g., in the field of workplace drug testing), while it is negligible for analytical procedures used in situations of occasional demand.

It should be noted, however, that such an approach does not lend itself to statistical interpretation and must be regarded as providing valuable information indicative of the accuracy of the analysis rather than the sound confirmation provided by a system of QC relying on control charts.

4.2. Confirmation tests

Because of the need to avoid as far as possible the incidence of FP results, the concept of confirmation is a cornerstone in analytical toxicology. Confirmation is especially important in forensic toxicology, because of the gravity of the health and judicial consequences which FP reports may involve.

For this aim, the requirement that the analytical technique used for confirmation analysis must be based on different physico-chemical principles from those used for screening tests is well-founded in good laboratory practice and provides an important additional level of assurance towards the correctness of results. With respect to the analytical techniques used for screening tests, a confirmation technique must offer a higher degree of specificity for the analyte in question and its LOD and LOQ should be well below the cutoff values used for screening [42]. Gas chromatography–mass spectrometry (GC–MS), providing specific identification of single substances at very low concentrations as well as accurate quantitative results, is usually adopted as the preferred technique [10], especially in the field of forensic toxicology, although alternative techniques may be used on condition that they meet the requirements of analytical quality.

4.2.1. Control of confirmation test

Each analytical run of samples to be confirmed must include: (a) samples certified to contain no analyte(s) of interest (*negative controls*); (b) samples fortified with known concentrations of analyte(s) of interest (*positive controls*); (c) positive controls containing the analyte(s) of interest with concentrations equal to or near the LOD and LOQ.

Because of the very meaning of *confirmation*, it is self-evident that a positive result when testing a negative control sample cannot be accepted as a confirmatory technique. If FPs are reported by a confirmation technique, *all* analytical results coming from the run must be rejected and the analytical system must be checked.

As for screening tests, control charts should be prepared to illustrate statistical variations of the analytical results when testing standards at known concentrations. Furthermore, when confirmatory tests are required for quantitative determinations, control charts are also needed to document the linearity and precision of the method. Data should be evaluated as described in Section 3.6 and, if values falling out of ‘warning limits’ are evidenced, remedial action must be taken.

5. External quality control

As stated above, in the case of errors related to permanent factors – that is if the method was not accurately set up (e.g., an impure standard was used, a measuring device was miscalibrated, etc.), or in order to verify whether other methods perform more

acceptably – IQC does not suffice [43]. Furthermore, interfering substances in real samples may not be reflected by the control samples used for IQC purposes.

External quality control programs (alternative terms are *Proficiency Testing Programs* (PTP) or *External Quality Assessment Schemes* (EQAS)) permitting a laboratory to monitor its performance by comparison with other laboratories, are better for evaluation of variables due to permanent factors.

These programs are distinct from IQC but, providing external evidence of the quality of the laboratory's analytical performance, one of the main purposes of PTPs is also to strongly encourage the proper use of IQC and to incorporate an external reference to guard against bias [44]. PTPs are also different from other interlaboratory tests, such as collaborative trials used for validating standard methods [45,46] or certification trials used to establish the true value of an analyte concentration in a reference material [47].

5.1. General context of proficiency testing

Proficiency testing must be seen in the general context of accreditation because, in order to gain accreditation, a toxicology laboratory must demonstrate an effective QA system, which includes participation in relevant PTPs [5,6]. Within QA, proficiency testing defines an inter-laboratory program aimed at assessing the accuracy of analytical data and the reliability of the methods used for qualitative and/or quantitative assay of samples [48,49].

In PTPs, samples are distributed for analysis to laboratories, a few times each year, by organizations that are external to the laboratories. These organizing bodies are responsible for the design of the scheme, the preparation, validation and distribution of test materials together with instructions to the participating laboratories, the collection and statistical analysis of the data obtained from the tests, and the feedback of results to participants.

This general organization is common to existing schemes but it must be emphasized that the single parts of the whole are often different, particularly as performance indices are frequently uneven. It is therefore difficult to appreciate rapidly the meaning of an index in an unfamiliar program. This is especially important when a laboratory starting a new analytical procedure uses results coming from EQAS to evaluate the best methods available. Therefore, to improve the usefulness of PTPs, the universal use of a standardized method for their assessment is greatly to be desired. [50].

5.2. Types of proficiency testing

Two distinct types of proficiency testing may be identified on the basis of their aims:

Educational, the aims of which are to improve the performance of participating laboratories, to supply a system to check the reliability of analytical data, to provide technical and scientific feedback, to encourage remedial action when shortcomings in performance are detected, and to set up a system for exchange of scientific information;

Accreditation, the aim of which is to provide a rational basis for the selection or licensing of laboratories for a specific task and, likewise, to disqualify laboratories from a specific task should their performance on it fall below a certain standard.

While participation in an educational program is usually on a voluntary basis, participation in accreditation programs is compulsory in those countries where they are used, by government agencies, as part of programs for laboratory accreditation [51]. In both cases, it is the major responsibility of the organizing body to guarantee the anonymity of the participating laboratory. In this sense, it is also good practice to identify participating laboratories by means of individual codes.

In analytical toxicology, the PTPs which have been operational over the last 20 years [39,52–58] have been organized by scientific or professional agencies and scientific institutions. Only in the United States, however, does legislation compel laboratories to document IQC data, the precision and accuracy of laboratory tests, and participation in inter-laboratory programs. In Europe, a questionnaire sent in 1991 to about 300 laboratories working in the field of analytical toxicology [9] indicated that only 50% were taking part in some sort of EQC program. This fact should not be interpreted only as the low participation of laboratories in such programs but it

may also be due to their low availability/sporadic activity.

5.2.1. Accreditation programs

Proficiency testing, as an instrument of evaluation or regulation of analytical laboratories, is included in the regulation of some countries such as the USA. Comparison or split samples are analysed both by the laboratory seeking accreditation and by the government laboratory, the results from the latter representing the standard. The laboratory being accredited may be required to maintain QC charts, and each analyst may be required to maintain individual charts. These findings are reviewed by the government's QC officer to ascertain whether or not the laboratory meets established performance standards. This kind of PTP is only part of the requirements for accreditation. An extensive description of accreditation programs is beyond the scope of the present paper and readers are encouraged to review the existing literature [4,33,59].

5.2.2. Open and blind proficiency testings

A further subdivision may be made between *open* and *blind* proficiency testing [60]. Proficiency test samples are identical for both these subtypes but, in the case of blind testing, the laboratory is unaware of the test. Samples arrive at the laboratory exactly as routine specimens do, are processed and tested, and the results reported in a routine fashion. The surrogate user then reports results back to the organizing body, which compares and scores the laboratory results. It is well known as laboratories that do well in open proficiency testing sometimes perform poorly in blind testing [61].

5.3. Proficiency testing program design

A general format for voluntary (educational) PTPs should include the following features [62]:

(a) the organizing body prepares a homogenous sample(s) and distributes portions to participating laboratories for analysis;

(b) the laboratory is required to examine the sample(s) within a specified time and to submit the results to the organizing body (if a standard method is not required by the organizing body, the laboratory is requested to describe the method used);

(c) the organizing body collates the data, performs statistical analysis, and sends a report to the participating laboratories (these reports (over time) should allow a participating laboratory to assess its performance at a particular time and to compare it with its performance in the past, to compare its performance with that of other laboratories, and to evaluate the performance of methods used by all participating laboratories).

However, this general design is not in itself sufficient to ensure either that every participating laboratory can test its performance with respect to its own analytical situation, or that PTP data can be used to monitor the 'ongoing' situation. Furthermore, unlike the accreditation program, an educational PTP should always consider the use of control samples with compositions mimicking as far as possible real samples. This gives a laboratory an effective tool to evaluate the possible impact of metabolites or interfering substances on its analytical methods, since they are not usually implemented in control samples for IQC purposes. A more detailed design for a PTP in analytical toxicology is described in the following sections.

5.3.1. Program protocol

The organizing body must create and distribute to intending participants a protocol clearly stating the purpose of the program, the type of analytes planned by it, the frequency of batch shipments, the time needed and the methods used for processing data, the outcome of statistical analysis, and the elapsed time for its shipment to laboratories.

5.3.2. Choice, preparation, validation and distribution of control materials

Natural materials are the best choice when selecting control samples but, due to the difficulty in obtaining both large amounts of materials and in suitable analyte concentrations, a valid alternative is to use spiked drug-free biological matrices with the analytes and possible interferents of interest. In no case should control materials be prepared by the producers of kits, calibrators or analytical instruments.

The main consideration in the preparation of materials is that they should, as far as possible, be representative of the type of materials normally

analysed, with respect to the composition of the matrix and the concentration range of the analytes. It should also reflect the composition in metabolites of real samples and consider the possible presence of interfering substances [63,64], as may happen during routine analysis.

For example, if a PTP requires the analysis of methadone or morphine in urine, at least 2-ethylidene-1,5-dimethyl-3,3-diphenylpyrrolidine (EDDP) and morphine-3-*O*-glucuronide, as their respective main metabolites, should also be added. If the analysis of amphetamines is required, analogues such as 3,4-methylendioxyamphetamine, 3,4-methylendioxyethylamphetamine or 3,4-methylendioxymethamphetamine (which are currently found on most illicit markets) should also be added, and it is also a good practice to add substances such as phenylpropanolamine or phentermine [65,66] which may cause cross-reactivity problems with some immunochemical techniques used for amphetamine screening tests.

As an example, Table 2 lists the panel of substances with relative cutoffs, adopted by the Italian PTP for drugs of abuse testing in urine [67].

Moreover, when such programs are based on educational assumptions, they should also include some informational aspects. For example, when the illicit drug *N*-methyl-3,4-methylendioxyphenyl-2-butanamine (MBDB) appeared in Italy for the first time, it was immediately added to some urine samples within the national PTP on drugs of abuse [67]. Since it was unknown to most of the approximately 200 participating laboratories, the first analytical results were disastrous, with very high numbers of FPs due to erroneous identification of MBDB with the other amphetamines included in the program at that time. Fast feedback by the organizing body and subsequent proper countermeasures by participants meant that, during later batches, there was a rapid improvement in performance (Fig. 1).

Sufficient volumes of control samples should be prepared, so that participating laboratories receive adequate aliquots for conducting multiple determinations or confirmation tests.

A basic step in the preparation of control samples is composition validation, which may be conducted in various ways. The closest approach to obtaining 'true' value of each single analyte is to use a consensus value produced by a group of qualified reference laboratories, external to the organizing body, using the best possible methods. In some cases, due to the high cost of this approach, consensus value may be produced from each round of the proficiency test on the basis of results obtained by participants. In this case, the consensus is usually estimated as the mean of the test results after any outliers have been rejected. This method is less expensive but not safe, because may provide consensus values biased by the general use of faulty methods [68].

Once sample composition has been validated, all materials must be checked to verify their homogeneity. After bulk materials have been subdivided for distribution, a random selection of 5–10% of the containers should be taken and the contents of each subjected to replicate analysis during the entire period planned for the analytical trial. This enables between-samples and between-time variations to be estimated by analysis of variance. Before shipment to participating laboratories, each sample must be analysed to check its composition in analytes and, when required, its sterility.

Materials thus prepared should be stored and distributed under conditions which minimize the effects of any sample instability.

To be effective in maintaining analytical quality standards of a laboratory, the minimum frequency for distribution of samples should be four rounds per year. Of course, for the same reason, once the trial has been completed, prompt feedback to the laboratory is necessary.

5.3.3. Choice of cutoff values

As stated previously, screening tests in toxicological analysis require cutoff values to guard against possible FPs and this also means that cutoff values must be planned by a PTP. Cutoffs may be autonomously set by the organizing body or taken from the prevailing values stated on an international scale by scientific organizations [9,33], and participating laboratories must perform their analyses on the basis of those cutoffs. Allowing a laboratory to communicate analytical results by comparison with its own internal cutoffs (as happens in countries where no cutoff values are defined by a central authority) is a practice to be avoided because this approach implies

Table 2
Substances and cutoff concentrations

Classes and Single substances	Cutoff (ng/ml)
Amphetamines and analogues	
Amphetamine	1000
Methamphetamine	1000
3,4-Methylendioxyamphetamine	1000
3,4-Methylendioxyethylamphetamine	1000
3,4-Methylendioxymethamphetamine	1000
Barbiturates	
Amobarbital	500
Butalbital	500
Butabarbital	500
Phenobarbital	500
Secobarbital	500
Benzodiazepines	
Diazepam	500
Nordiazepam	500
Oxazepam	500
Nitrazepam	500
7-Aminonitrazepam	500
Flunitrazepam	500
7-Aminoflunitrazepam	500
Flurazepam	500
Desalkylflurazepam	500
N-hydroxyethylflurazepam	500
Lorazepam	500
Triazolam	500
α-Hydroxytriazolam	500
Cannabinoids	
11-nor-9-COOH-Δ^9-THC	50
11-nor-9-COOH-Δ^9-THC-glucuronide[a]	
Cocaine	
Benzoylecgonine	150
Cocaine[a]	
Ecgonine methylester[a]	
Methadone	
Methadone	300
2-Ethylidene-1,5-dimethyl-3,3-diphenylpyrrolidine[a]	
Opiates	
Codeine	300
Morphine	300
6-Monoacetylmorphine[a]	
Morphine-3-*O*-glucuronide[a]	
Morphine-3-*O*-sulphate[a]	
Interfering substances	
Phentermine	
Phenylpropanolamine	
Ephedrine	
Selegiline	
Chlorpromazine	
N-Methyl-3,4-methylendioxyphenyl-2-butanamine	
Dantoine	
Ethylmorphine, etc.	

[a]Metabolites added with the function of mimicking real samples.

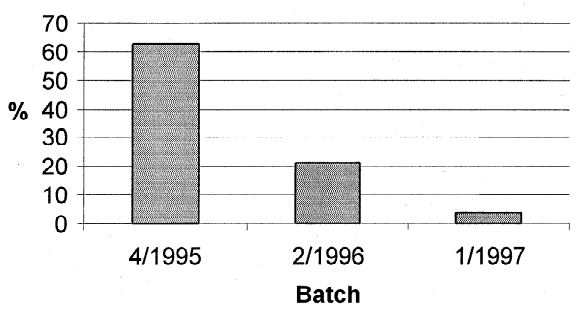

Fig. 1. Percentage of FPs for MBDB in three subsequent analytical trials of Italian PTP for drugs of abuse in urine.

excessive fragmentation of the statistical evaluations, which is not very helpful for all the participants. If a laboratory participates for quantitative determination of analytes (this implies confirmation testing), it should be allowed to communicate its LOQ in performing this analysis and to use this limit to evaluate the presence or absence of a substance.

5.3.4. Choice of participation

The greater the number of laboratories participating in a PTP, the more effective statistical evaluation of the general performance of participants and methods will be. However, when there are many participating laboratories, there are also great differences among their analytical approaches and the substances being tested. This means that an effective PTP must be sufficiently flexible to offer participating laboratories the possibility of choosing a kind of tailor-made participation, on the basis of their effective needs.

In this respect, a participating laboratory should be able to choose among *identification* of *Classes* of substances, qualitative *identification* of *Single* substances, and *quantification* of *Single* substances. Furthermore, for each batch of samples, laboratories may decide among the ones planned by the program, for which Class of substances to participate and, within those Classes, for which single analytes.

5.3.5. Analytical trials

Adequate, pre-formatted analysis forms must be supplied to participating laboratories. These forms must be simple enough to be correctly compiled by all participants but sufficiently detailed for a good description of the entire analytical procedure; the best method here is to provide for compilation by predefined codes. The code list must be as comprehensive as possible, supplied to participating laboratories together with the protocol, and also provided together with every single batch reporting possible updatings. The analysis forms should at least allow participating laboratories to provide:

(a) an indication of the presence or absence of Classes of substances planned by the program;

(b) an indication of the presence or absence of Single substances (only those planned by the program, not possible others added to act as interfering substances);

(c) quantification of Single substances (only those planned by the program, not others added to act as interfering substances); in this case the analysis forms should also report the unit of measurement required by the program;

(d) communication of sample pretreatment procedures used for screening, and confirmation tests if performed;

(e) communication of analytical technique(s) used for screening, and confirmation tests if performed.

As far as possible a laboratory should perform PTP analyses in the same way in which it performs routine analyses. However, because analysts are over-careful with PTP samples, this is a difficult task to accomplish. As described previously, blind proficiency testing is necessary in this case.

Analyses must be performed within a specified time (usually planned considering the necessary requirements of fast feedback, but also avoiding placing too great a burden on a laboratory) and the forms sent back to the organizing body before the indicated deadline.

5.3.6. Data collection and analysis

The need to store data coming from every single batch for later consultation and comparison, the often complex statistical procedures required to be carried out in an error-free way, and the absolute need to avoid accidental exchanges between data from different laboratories, all indicate that this phase must be carried out using a computer-aided system.

All analysis forms from participating laboratories should be inserted in a proper data-base by two different operators. After data entry, a control software program compares the two databases to check

for possible mistakes. In the same way, subsequent analysis of data must be conducted using both databases and a tailor-made program of statistical evaluation. It is good practice (although expensive in terms of time) that reports are compared with the respective analysis forms before shipment.

5.3.7. Feedback to participants

Usually, the first stage in producing a score from a result X (a single measurement of analyte concentration in a test material) is obtaining the estimate of the bias, thus: bias $= x - X$, where x is the true concentration or amount of analyte. Most PTPs proceed by comparing the estimate of the bias with a standard error. An obvious approach is to form a z-score given by: $z = (x - X)/\sigma$, representing the maximum allowable variation consistent with valid data [69].

However, the above sections describe a kind of program in which the general parameters used for PTPs, such as z-score, q-score [43], Standard Deviation Index (SDI) [70], etc., are not suitable for correct scoring of participating laboratories.

Indeed, unlike IQC, in which the quantitative component of a screening test can be used for an evaluation of the analytical result, in EQC for analytical toxicology, participating laboratories can choose to report results only in terms of the presence or absence of the Class or Single substance to be analysed. Of course, a different approach must be used for evaluation of quantitative results. Furthermore, it must always be kept in mind that, when designing an educational program, the first aim is to improve laboratory performance and not rank laboratories on the basis of their actual performance. In this respect, the best approach is to make use of a reporting fashion as comprehensive as possible containing all necessary information a laboratory can use for better evaluation of its performance.

Reports should therefore contain:

(a) the composition (quali-quantitative) of the analysed batch, both *real* and *communicated* by the laboratory; real composition must also include substances added for special purposes (cross-reactions, interferences) and different from those planned by the program;

(b) every single error made, clearly highlighted;

(c) the FP rate obtained by the laboratory (this parameter is important in evaluating possible interferences or cross-reactions which cannot be identified by IQC procedures);

(d) FN rates for all the Classes and/or Single substances for which the laboratory has participated with qualitative identification only;

(e) global FP rates and FN rates (all Classes and Single substances) obtained by all laboratories participating in the trial;

(f) global FP rates and FN rates (all Classes and Single substances) obtained by all analytical techniques used during the trial (these data should also be accompanied by a description of pretreatment procedures, if performed);

(g) besides the above items, laboratories participating for quantitation must also receive: z-score indexes of their performance and the SDI, which expresses the difference between a laboratory test result and the overall average in terms of the number of standard deviations from the overall mean; in creating the SDI, great attention must be paid to the analytical technique used by the participating laboratory; in general, quantitation from immunochemical techniques should be refused for this kind of PTP or, at least, two different SDIs should be created.

In addition, reports for participating laboratories should also contain a warning system for those laboratories whose performance falls below a minimum standard of analytical quality.

5.3.8. Complementary services

Global reports or workshops on, for example, an annual basis should be planned, to present results obtained by participating laboratories during the past year of activity and to discuss method developments and topics of analytical toxicology [71–73].

A *direct line* between the organizing body and participating laboratories, providing information on PTP-related topics (for example, changes in planned cutoffs, list of substances to be analysed, etc.) should also be implemented.

5.4. Limitations of a proficiency testing program

PTP is not in itself sufficient to ensure the production of high-quality data. First, the interpretation of data from PTP is subject to statistical uncertainty, and the criteria on which decisions will

be based are to some degree arbitrary. Second, too much time often elapses between analysis of test samples and communication of performance results to each participating laboratory. Third, there is no way of determining whether samples are examined by routine test procedures or if they receive special treatment. The fact that participating laboratories approach analysis with more care and attention may have a great impact on statistical certainty but, unfortunately, these practices are difficult to eliminate. In this respect, blind testing would be helpful to avoid these problems but, as a fourth problem, already high costs in carrying out open PTPs would be further increased.

5.5. Improvement by new technologies

The worldwide communication by electronic means offered by the Internet provides an obvious enhanced opportunity for a solution to some PTP-related problems. For example, planning an appropriate analysis form '*on-line*' would avoid any problem related to bad data entry by the organizing body (and, as a consequence, of incorrect statistics). In this way, since the Internet not only potentially reaches all the world but also does so very rapidly, problems related to dead time between analysis form shipment and feedback would be avoided, because the entire process would be carried out by an automated system from receipt of analysis forms to shipment of reports. Thus, the PTP would increase its efficacy in monitoring the ongoing performance of a laboratory. Furthermore, using appropriate password systems, Web pages could be set up for every single participating laboratory so that they can access them to monitor their overall performance in one step only (avoiding having to browse among many reports), from the beginning of participation until the present moment. Alternatively, a free access area could contain continuously updated information about the program as a whole, together with global statistics summarizing results obtained by all laboratories and analytical methods both in single batches and as a general trend in time. Lastly, a discussion forum for PTP users could also easily be created, to improve the exchange of scientific information and the usefulness of participation.

This kind of Internet-based PTP has been carried out in Italy since 1994 by the Centre of Behavioural and Forensic Toxicology of the University of Padova, with about 200 participating laboratories [67].

6. Concluding remarks

The inherently high degree of reliability related to analytical toxicology implies a higher degree of quality and consequent appropriate implementation of a system of control which, although it cannot improve analytical quality by itself, is useful in detecting errors.

IQCs can only detect errors related to variable factors, so that a scheme for integration of internal and external QC programs should be planned on the basis of quality requirements as well as of specific laboratory needs.

For this aim, PTPs must be designed to allow each single participating laboratory to determine its appropriate quality goals, and should be based on the universal use of a standardized method for assessment of results. The introduction into proficiency schemes of new technological tools, such as the global communication offered by the Internet, constitutes an important step towards this goal.

References

[1] A. Cross-Smiecinski, L.D. Stetzenbach, Quality Planning for the Life Science Researcher, CRC Press, Boca Raton, FL, 1994.
[2] P. Bachner, Clin. Lab. Med. 6 (1986) 613.
[3] R.J. Mesley, W.D. Pocklington, R.F. Walker, Analyst 116 (1991) 975.
[4] Revisions of the Laboratory Regulations for the Medicare, Medicaid, and Clinical Laboratory Improvement Act of 1967: Final Rule with Comment Period, USA Federal Register, March 14, 55 (1990) 9538.
[5] International Organization for Standardization, General Requirements for the Competence of Calibration and Testing Laboratories, ISO Guide 25, 3rd ed., Geneva, 1990.
[6] The Joint European Standards Institution, General Criteria for the Operation of Testing Laboratories, EN 45001, CEN/Cenelec, Brussels, 1989.
[7] P. Hyltoft Petersen, C. Ricos, D. Stockl, J.C. Libeer, H. Baadenhujisen, C. Fraser, L. Thienpont, Eur. J. Clin. Chem. Clin. Biochem. 34 (1996) 983.
[8] R. Blanke, Clin. Chem. 33 (1987) 41B.

[9] R. de la Torre, J. Segura, R.A. de Zeeuw, J. Williams, In collaboration with the Toxicology Experts Working Group, Ann. Clin. Biochem. 34 (1997) 339.
[10] NIDA, Technical, Scientific and Procedural Issues of Employee Drug Testing—Consensus Report, DHHS Publication, Rockville, 1990.
[11] J.C. Libeer, H. Baadenhujisen, C.G. Fraser, P. Hyltoft Petersen, D. Stockl, C. Ricos, L. Thienpont, Eur. J. Clin. Chem. Clin. Biochem. 34 (1996) 665.
[12] P. Hyltoft Petersen, O. Blaabjerg, K. Irjala, A. Icen, K. Bjoro, Uppsala J. Med. Sci. 98(C.H. deVerdieer, T. Groth, P. Hyltoft Petersen (Eds.), Medical Need for Quality Specifications in Clinical Laboratories) (1993) 241.
[13] Analytical Methods Committee, Analyst 114 (1989) 1497.
[14] R.S. Galen, S.R. Gambino, Beyond Normality: the Predictive Value and the Efficiency of Medical Diagnoses, Wiley & Sons, New York, 1975.
[15] P.F. Griner, R.J. Mayewski, A.I. Mushlin, P. Greenland, Ann. Intern. Med. 94 (1981) 553.
[16] S.K. Lwanga, C.Y. Tye, Teaching Health Statistics, World Health Organization, Geneva, 1986, p. 69.
[17] V.R. Spiehler, C.M. O'Donnell, D.V. Gokhale, Clin. Chem. 34 (1988) 1535.
[18] S.D. Ferrara, L. Tedeschi, G. Frison, G. Brusini, F. Castagna, B. Bernardelli, D. Soregaroli, J. Anal. Toxicol. 18 (1994) 278.
[19] J.O. Westgard, E.F. Quam, P.L. Barry, Clin. Lab. Sci. 3 (1990) 273.
[20] J.O. Westgard, Med. Lab. Observ. 26 (1994) 55.
[21] W.A. Shewhart, Economic Control of the Quality of Manufactured Products, McMillan, London, 1931.
[22] S. Levey, E.R. Jenning, Am. J. Clin. Pathol. 20 (1950) 1059.
[23] A.J. Duncan, Quality Control and Industrial Statistics, 5th ed., Irwin, Homewood, 1986.
[24] T.P. Ryan, Statistical Method for Quality Improvement, Part II, Wiley & Sons, New York, 1989.
[25] M. Thompson, R.J. Howarth, Analyst 105 (1980) 1188.
[26] G.B. Wetherhill, Sampling Inspections and Quality Control, 2nd ed., Chapman & Hall, London, 1977.
[27] O.L. Davies, P.L. Goldsmith, Statistical Method in Research and Production, 4th ed. (revised), Oliver & Boyd for ICI, Edinburgh, 1972.
[28] M. Thompson, R.J. Howart, Analyst 101 (1976) 690.
[29] M. Montagna, H. Muntau, A. Polettini, S. Facchetti, in: Proceedings of the International Conference and Workshop for Hair Analysis in Forensic Toxicology, Abu Dhabi, United Arab Emirates, November 1995, p. 370.
[30] BIPM, IEC, IFCC, ISO, IUPAC, OIML, International Vocabulary of Basic and General Terms in Metrology, 2nd ed., ISO, Geneva, 1993.
[31] A.H.B. Wu, Ann. Clin. Lab. Sci. 25 (1995) 319.
[32] M.A. ElSohly, A.B. Jones, J. Anal. Toxicol. 19 (1995) 450.
[33] Mandatory Guidelines for Federal Workplace Drug Testing Programs; Final Guidelines; Notice, USA Federal Register, April 11, 53 (1988) 11978.
[34] J.P. Dux, Handbook of Quality Assurance for the Analytical Chemistry Laboratory, Van Nostrand Rehinold Co., New York, 1986.
[35] J.K. Taylor, Quality Assurance of Chemical Measurements, Lewis Publisher Inc., Chelsea, 1987.
[36] T.A. Kubic, J.A. Buscaglia, in C.G.G. Aitken and D.A. Stoney (Eds.), The Use of Statistics in Forensic Science, Ellis Horwood, New York, 1990, p. 207.
[37] J.O. Westgard, G.G. Klee, in: N.W. Tietz (Ed.), Textbook of Clinical Chemistry, Saunders, Philadelphia, 1986, p. 424.
[38] Navy Drug Screening Laboratory Standard Operating Procedures (SOP) Manual, Naval Medical Command, US Navy, Washington, October 1, 1986.
[39] J. Irving, Clin. Chem. 34 (1988) 637.
[40] J.P. Egan, Detection Theory and ROC Analysis, Academic Press, New York, 1977.
[41] J.A. Swets, Evaluation of Diagnostic System: Methods from Signal Detection Theory, Academic Press, New York, 1982.
[42] G.M. Lawson, Clin. Chem. 40 (1994) 1219.
[43] S.D. Ferrara, in: Proceedings of the European Military Conference on Drugs of Abuse: International Meeting on Assessment and Control of Drugs Abuse in Young Population and Mutual Relationship with Military Organizations in European Countries, Senger and Associates, Palo Alto, 1989.
[44] Analytical Methods Committee, Analyst 117 (1992) 97.
[45] M.J. Welch, L.T. Sniegoski, C.C. Allgood, Forensic Sci. Int. 63 (1993) 295.
[46] P. Kintz, Forensic Sci. Int. 70 (1995) 105.
[47] P. Kintz, V. Cirimele, Forensic Sci. Int. 84 (1997) 151.
[48] S.D. Ferrara, I.C. Dijkhuis, in: S.D. Ferrara (Ed.), Il Laboratorio di Farmacologia e Tossicologia Clinica, CG Ed. Med. Sci., Torino, 1989, p. 857.
[49] S.D. Ferrara, Forensic. Sci. Int. 63 (1993) 305.
[50] I.C. Dijkhuis, B. Widdop, M. Moeller, in: Report of a Workshop on Proficiency Testing Programs for Psychoactive Substances, XXXV Meeting of The International Association of Forensic Toxicologists, Padova, Italy, August 1997.
[51] H.J. Barth, in: G.A. Berman (Ed.), Testing Laboratory Performance: Evaluation and Accreditation, NBS Publication 591, National Institute of Standards and Technologies, Gaithersburg, 1980, p. 169.
[52] I.C. Dijkhuis, Lancet ii (1982) 1341.
[53] C.S. Frings, D.J. Battaglia, R.M. White, Clin. Chem. 35 (1989) 891.
[54] US Department of Health, Education and Welfare, Proficiency Testing Summary Analysis: Toxicology of Drugs of Abuse, Technical reports, Atlanta, 1974–80.
[55] J. Segura, R. de la Torre, M. Congost, J. Cami, Clin. Chem. 35 (1989) 879.
[56] CBFT, Drugs of Abuse Proficiency Testing Summary Analysis, University of Padova, Technical Reports, Padova, 1982–96.
[57] J.F. Wilson, B.L. Smith, P.A. Toseland, J. Williams, D. Burnett, A.D. Hearst, I.D. Watson, A.N. Horn, Ann. Clin. Biochem. 31 (1994) 335.
[58] K.H. Schaller, J. Angerer, G. Lenhert, Int. Arch. Occup. Environ. Health 62 (1991) 537.
[59] W. Horwitz, in: S.L. Inhorn (Ed.), Quality Assurance Practices for Health Laboratories, American Public Health Association, Washington, 1977, p. 616.

[60] S.D. Ferrara, in: B. Kaempe (Ed.), Proceedings of the XXIX Meeting of The International Association of Forensic Toxicologists, MacKenzie, Copenhagen, Denmark, 1991, p. 69.
[61] H.J. Hansen, S.P. Caudill, J. Boone, J. Am. Med. Assoc. 253 (1985) 2382.
[62] F.M. Garfield, Quality Assurance Principles for Analytical Laboratories, 5th ed., AOAC International, Arlington, 1991.
[63] D.L. Colbert, Br. J. Biomed. Sci. 51 (1994) 136.
[64] M.P. Bosomworth, Br. J. Biomed. Sci. 50 (1993) 150.
[65] J. D'Nicuola, R. Jones, B. Levine, M.L. Smith, J. Anal. Toxicol. 16 (1992) 211.
[66] G.J. Turner, D.L. Colbert, B.Z. Chowdry, Ann. Clin. Biochem. 28 (1991) 588.
[67] S.D. Ferrara, in: Proceedings of the International Conference: The European Drug Testing Environment: Emerging Issues and Challenges, Seville, Spain, November 1994, p. 39.
[68] J.O. Westgard, J.J. Seehafer, P.L. Barry, Clin. Chem. 40 (1994) 1909.
[69] N.T. Crosby, J.A. Day, W.A. Hardcastle, D.G. Holcombe, R.D. Treble, F.E. Prichard, in: E.J. Newman (Ed.), Quality in the Analytical Chemistry Laboratory, Wiley & Sons, New York, 1995, p. 24.
[70] J.O. Westgard, http://www.westgard.com/lesson14.htm, 1996.
[71] S.D. Ferrara, L. Tedeschi, Ann. Ist. Super. Sanità 18 (1982) 727.
[72] S.D. Ferrara, L. Tedeschi, in: E. Pistocchi, C. Baccini (Eds.), Proceedings of the National Congress, Il Laboratorio di Chimica Clinica e le Tossicodipendenze, Cesena, Italy, March 1982.
[73] S.D. Ferrara, L. Tedeschi, in: S.D. Ferrara, I. Sunshine, A. Fuller (Eds.), Drugs of Abuse, Syva Publishing, Palo Alto, 1987, p. 41.