UNITED NATIONS INTERNATIONAL DRUG CONTROL PROGRAMME
Vienna

GLOSSARY OF TERMS FOR QUALITY ASSURANCE AND GOOD LABORATORY PRACTICES

UNITED NATIONS
New York, 1995

NOTE

Words printed in italics are glossary entries elsewhere or are variations of these entries. They may occur more than once in a definition but they will be printed in italics only once. "Specimen" in the context of this Glossary refers to biological specimens. "Sample" refers to a sample of seized material.

Symbols of United Nations documents are composed of capital letters combined with figures. Mention of such a symbol indicates reference to a United Nations document.

Figures in square brackets [ ] are keyed to the reference list.

ST/NAR/26

Introduction

Background

Over the last decade there has been an enormous increase not only in the production and supply of illicit drugs, reflected by the huge and mounting quantities of drugs seized by national and international authorities, but also in the rate of drug abuse, i.e. the illicit demand for drugs. The drugs being seized include not only those already under national and international control but also unexpected new illicit drugs or combinations of drugs prepared by chemists working in clandestine laboratories. At the same time there are reports of the growing misuse or abuse of drugs used for medical purposes, such as barbiturates and benzodiazepines.

What was traditionally a problem of developed countries is no longer confined to those countries. Drug abuse is now a global problem affecting developed and developing countries alike: today, no nation is free from this threat.

The extent and diversity of abuse have placed increasing demands on nations to intensify their regulatory efforts; in some cases, they have introduced stringent legislation that may have serious consequences for an individual charged with a drug offence. Ultimately, the outcome of the legal proceedings would rest on the results of laboratory tests. This has placed greater pressure on national laboratories, which are now required not only to identify seized materials but also to detect drug abuse by analysing biological specimens. In addition, while in the past these laboratories were required to perform only qualitative analyses, they are now required to produce highly reliable quantitative results as well.

In the field of drug abuse, laboratories now have to be able to deal with more substances and to use methods of detection and analysis that are faster yet more accurate and specific. The analysis of biological specimens like urine and blood presents additional challenges because of the need to separate target substances from interferences in blood and urine, which are complex biological matrices.

The international nature of the drug abuse problem requires the speedy exchange of analytical data between laboratories and between law enforcement agencies and laboratories at the national and international levels. The development of internationally acceptable methods of detection and assay is expediting this exchange. The United Nations, through the United Nations International Drug Control Programme (UNDCP) and other bodies of the system, has been instrumental in achieving this aim. Two series of manuals, one dealing with methods for the analysis of seized materials and the other dealing with biological specimens, have been written for use by national laboratories.

Because analytical accuracy is so important, it is increasingly recommended that laboratories should implement effective quality assurance (QA) procedures, stimulate the use of good laboratory practices (GLP) and participate in proficiency

testing programmes. These issues were addressed in three UNDCP meetings,* the first of which made 14 recommendations. Two of the recommendations were that an international drug proficiency testing programme should be set up and that a glossary of quality assurance terms should be prepared.

Purpose of the Glossary

The field of quality assurance is beset by an extensive and sometimes confusing terminology that has not been universally agreed on or accepted by laboratories or professional bodies. In recent years several organizations have begun to standardize the use of certain terms.** This Glossary is the product of the three above-mentioned meetings dealing with quality assurance. By providing internationally acceptable definitions for some of the most common expressions in the field, it aims to promote and harmonize national efforts. It is intended to help analysts with implementing quality assurance programmes and good laboratory practices, and it is being provided to laboratories as an educational document within the training remit of UNDCP, to encourage them to collaborate and participate with the United Nations in quality assurance matters.

Use of the Glossary

The Glossary is intended to be used in conjunction with the United Nations *Recommended Guidelines for Quality Assurance and Good Laboratory Practices* (ST/NAR/25). It should be used as a supplement to textbooks on quality assurance, including reference books on statistical procedures used in quality assurance, not as a substitute for them. Furthermore, the analyst is expected to keep abreast of developments in the field by following current literature on the subject. The Technical Services Branch would welcome observations on the contents and usefulness of this Glossary. They may be addressed to:

United Nations International Drug Control Programme
Laboratory Operations
Research and Scientific Section
Technical Services Branch
United Nations Office at Vienna
Vienna International Centre
P.O. Box 500
A-1400 Vienna, Austria

*Consultative Meeting on Quality Assurance and Good Laboratory Practices, Glasgow, 1992; Consultative Meeting on a Glossary of Quality Assurance Terms, Vienna, 1993; and Consultative Meeting on an International Drug Proficiency Testing Programme, Vienna, 1993.

**Two of the resulting publications are International Organization for Standardization (ISO)/ Development Information System (DIS), *Quality Management and Quality Assurance Vocabulary,* publication 8402 (Geneva, 1991), and ISO, *International Vocabulary of Basic and General Terms Used in Metrology* (Geneva, 1984).

GLOSSARY

absolute error: See *error.*

acceptance criteria for specimens or **samples:** Procedures for acceptance or rejection of *specimens* or *samples* arriving at the analytical laboratory. Such procedures are focused on assessing the adequacy of the chain of custody [1].

accreditation: Procedure by which an *accreditation body* formally recognizes that a laboratory or person is competent to carry out specific tasks.

accreditation body: Independent science-based organization that has the authority to grant *accreditation.*

accuracy: Ability to get the true result [2]. For quantitative tests the accuracy expresses the closeness of agreement between the true value and the value obtained by applying the test procedure a number of times. It is affected by *systematic* and *random errors.*

action limit: See *limit.*

aliquot: Portion of a liquid *sample* or *solution.*

alternative hypothesis: See *hypothesis testing.*

analysis: See *test.*

analysis of variance (ANOVA): Statistical technique that can be used to separate and estimate the different causes of variation [3].

analyte or **target analyte:** Substance to be identified or measured [4].

> **surrogate analyte:** Well-characterized substance that is taken as representative of the analyte [5].

analytical batch: See *batch.*

analytical method: See *test.*

archive: Collection of documents and records purposefully stored for a defined period of time [6].

arithmetic mean or **average:** Sum of the individual values in a set divided by the number of values [7].

assay: Quantitative measurement of an *analyte.*

assigned value: See *value.*

average: See *arithmetic mean.*

batch or **analytical batch:** Group of one or more *specimens* or *samples* that are analysed under conditions approaching *repeatability.* Usually it should contain *calibrators* and *quality control specimens* or *samples* in addition to the samples to be analysed.

best fit: See *goodness-of-fit.*

bias: Difference between the expectation of the *test* result and an accepted reference value. There may be one or more *systematic error* components contributing to the bias [5].

binomial distribution: See *distribution.*

blank: *Specimen* or *sample* not containing the *analyte.*

blind specimen or **sample:** *Specimen* or *sample* that is analysed by an operator who is unaware at the time of the analysis that the sample is for control purposes (based on [8]).

blunder: Big mistake, especially one that seems to be the result of carelessness or stupidity [6]. See *outlier.*

calibration: Set of operations that establishes, under specified conditions, the relationship between values indicated by a measuring instrument or measuring system, or values represented by a material measure, and the corresponding known values of a measurand [9].

calibration curve: Relationship between the signal response of the instrument and various *concentrations* of *analyte* in a suitable solvent or *matrix* [7].

calibration laboratory: *Laboratory* that performs *calibrations* [9].

calibration method: Defined technical *procedure* for performing a *calibration* [9].

calibrator: Pure *analyte* in a suitable solvent or *matrix,* used to prepare the *calibration curve.*

certification: *Procedure* by which a *certifying body* formally recognizes that a body, person or product complies with given *specifications.*

certified reference material (CRM): A reference material one or more of whose property values have been certified by a technical *procedure,* accompanied by or *traceable* to a certificate or other documentation that has been issued by a *certifying body* [9].

certifying body: Independent science-based *organization* that has the competence to grant *certifications.*

chain of custody: Procedures and documents that account for the integrity of a *specimen* or *sample* by tracking its handling and storage from its point of collection to its final disposition (adapted from [1]).

chirality: Property of a molecule not superimposable on its mirror image. Due to asymmetry in their structures, chiral molecules can exist as different isomers and will have special optical and biological properties.

chi-square distribution: See *distribution.*

clerical error: See *error.*

Cochran test: See *outlier.*

coefficient of variation or **relative standard deviation:** Measure used to compare the dispersion or *variation* in groups of measurements. It is the ratio of the *standard deviation* to the *mean,* multiplied by 100 to convert it to a percentage of the *average* [7], [10].

collaborative studies or **interlaboratory test comparisons:** Organization, performance and evaluation of *tests* on the same or similar items or materials by two or more different *laboratories* in accordance with predetermined conditions. The main purpose is *validation* of analytical methods or establishment of reference methods [11].

comparison-of-means test: See *significance test.*

concentration: Amount of a substance, expressed in mass or molar units, in a unit volume of fluid.

confidence coefficient: See *confidence level.*

confidence interval: Range of *values* that contains the *true value* at a given level of *probability*. This level of probability is called the *confidence level.*

confidence level or **confidence coefficient:** Measure of *probability,* α, associated with a *confidence interval,* expressing the probability of the truth of a statement that the interval will include the parameter *value* [12].

confidence limits: The extreme *values* or end values in a *confidence interval.* See *limit.*

confirmatory test: Second *test* by an alternative chemical method for unambiguous identification of a drug or *metabolite* (adapted from [2]).

consensus value: See *value.*

contractor: *Organization* that provides a service under contractual conditions. It should be ensured that the contractor provides services in line with specified criteria of competence.

control chart: *Plot* of *test* results with respect to time or sequence of measurements, with *limits* drawn within which results are expected to lie when the analytical scheme is in a state of *statistical control* [5].

> **cusum chart:** In a cusum chart each result is compared with a reference, usually the *assigned* or *target value*. The differences from the reference are then accumulated, respecting the sign, to give a cumulative sum of differences from the standard. The cusum chart has the advantage of identifying small persistent changes in the analytical scheme faster than the *Shewhart chart* [7].
>
> **Shewhart chart:** chart where the variable of interest is plotted against *batch* or time. The observed *values* are compared with the *expected* or *true value.* Lines corresponding to the *mean* value obtained from *replicate analysis* of *reference material* and *warning* and *action limits* are inserted to provide objective criteria for the interpretation of the chart [5].

control limit: See *limit.*

controls: *Specimens* or *samples* used to determine the validity of the *calibration,* that is, the linearity and stability of a *quantitative test* or determination over time. Controls are either prepared from the *reference material* (separately from the *calibrators*, that is, weighed or measured separately), purchased, or obtained from a pool of previously analysed specimens or samples. Where possible, controls should be *matrix*-matched to specimens or samples and calibrators [1].

> **positive control:** Control that contains the *analyte* at a *concentration* above a specified *limit.*
>
> **negative control:** Control that contains the *analyte* at a *concentration* below a specified *limit.* Usually a drug-free specimen or sample (*blank*) is used as a negative control.

corrective action: Action taken to eliminate the causes of an existing *deviation,* defect or other undesirable situation in order to prevent recurrence.

correlation coefficient: Number showing the degree to which two *variables* are related. Correlation coefficients range from 0 (no correlation) to –1 or +1 (perfect correlation).

cross-reacting substance: In immunoassays, a substance that reacts with antiserum produced for the target *analyte* [2].

cusum chart: See *control chart.*

cut-off concentration: *Concentration* of a drug in a *specimen* or *sample* used to determine whether the specimen or sample is considered *positive* or *negative* [13]. In some circumstances it is recommended that the cut-off concentration should be set equal to the *limit of detection.* See *threshold.*

degrees of freedom: Number of independent comparisons that can be made between the members of a *sample* [12].

detection limit: See *limit.*

deviation: Departure from what is considered normal [6]. See *standard deviation.*

discrimination: Ability to recognize and understand the differences between two things [6].

distribution: A ranking, from lowest to highest, of the values of a *variable* and the resulting pattern of measures or scores when they are plotted on a graph (adapted from [10]). A frequency distribution, for example, gives the possible values of a parameter versus the number of times each value occurred in the *sample* or *population.* In many instances it refers to the spread of the individual values of a sample or population around the *mean.*

binomial distribution: Based on the idea that if only one of two possible outcomes can occur on any one occasion, then the theoretical distribution of the different combinations of outcomes that could occur can be worked out if the number of occasions is known. One characteristic of this distribution is that it consists of a limited or finite number of events, n. When n becomes very large, tending to infinity, the binomial distribution becomes the *normal distribution.*

chi-square distribution: This distribution may be considered as that of the sum of squares of ν independent normal *variates* in standard form. The *parameter* ν is known as the number of *degrees of freedom* [12].

F-distribution: Theoretical distribution used to study *population variances.* It is the distribution of the ratio of two independent variables each of which has been divided by its *degrees of freedom* [10].

normal distribution: Purely theoretical continuous *probability distribution* in which the horizontal axis represents all possible values of a *variable* and the vertical axis represents the *probability* of those values occurring. The scores on the variable are clustered around the *mean* in a symmetrical, unimodal pattern known as the bell-shaped (normal) curve. In a normal distribution, the *mean, median* and *mode* are all the same [10]. The normal distribution is obtained when the number of events in the *binomial distribution, n,* becomes very large, tending to infinity.

probability distribution: Distribution giving the *probability* of a value of x as a function of x or, more generally, the probability of joint occurrence of a set of *variates* $x_1, \ldots, x_p$ as a function of those quantities [12].

t-distribution: *Theoretical probability distribution* used in *hypothesis testing.* Like the *normal distribution,* the t-distribution is unimodal, symmetrical and bell-shaped [10].

theoretical probability distribution: Number of times it can be expected to get a particular number of successes in a large number of trials [10]. Important

theoretical probability distributions are the *normal, t-, chi-square* and *F-distributions.*

z-distribution: *Normal distribution* in which the scores are the z-scores [10].

distribution function: The distribution function $F(x)$ of a *variate* x is the total frequency of members with variate values less than or equal to x. As a general rule, the total frequency is taken to be unity, in which case the distribution function is the proportion of members bearing values less than or equal to x [12].

Dixon test: See *outlier.*

double blind procedure: Means of reducing *bias* in an experiment. In the clinical context, for example, such a procedure ensures that both those who administer a treatment and those who receive it do not know (are blind to) which subjects are in the control group and which are in the experimental group, that is, who is and is not receiving the treatment (adapted from [10]).

duplicate samples or **specimens:** Two *aliquots* of a *sample* or *specimen* analysed at the same time.

dynamic range: Range over which a relationship exists between *analyte concentration* and *assay* response [14].

error: Something done that is considered to be incorrect or wrong [6].

absolute error: Difference between the analytical result and the *true value* [15].

clerical error: Mistake made during routine jobs in an office or laboratory, e.g. a transcription error, a *specimen* misidentification or a filing error.

maximum tolerable error: Extreme values of an error permitted by *specifications*, regulations etc. for a given determination [16].

random error: Component of the *total error* of a measurement that varies in an unpredictable way. This causes the individual results to fall on both sides of the *average* value [3], [17].

relative error: *Absolute error* of a measurement divided by the *assigned value* of the analyte [16]. See *coefficient of variation* and *relative standard deviation.*

systematic error: Component of the *total error* of a measurement that varies in a constant way. This causes all the results to be in error in the same sense [3], [16].

total error: Sum of *random* and *systematic errors.*

type I error: Error made by wrongly rejecting a true *null hypothesis* [10]. If the null hypothesis is that the sample should be *negative*, a type I error will generate a *false positive* result.

type II error: Error made by wrongly accepting a false *null hypothesis* [10]. If the null hypothesis is that the sample should be *negative*, a type II error will generate a *false negative* result.

estimate value: See *value.*

evaluation: Systematic examination of the extent to which a product, process or service fulfils specified requirements [11].

expert witness: Knowledgeable person, for example, a forensic scientist, familiar with the testing and the interpretation of test results and able to give an expert opinion based on scientific fact or evidence, e.g. in court or at a hearing (adapted from [18]).

expiration date: Date after which the specified characteristics of a reagent, *solution*, *specimen*, *control* etc. can no longer be guaranteed.

false negative: *Test* result that states that no drug or *metabolite* is present when, in fact, such a drug or metabolite is present in an amount greater than a *threshold* or designated *cut-off concentration* (adapted from [19]).

false positive: *Test* result that states that a drug or *metabolite* is present when, in fact, it is not present or is present in an amount less than a *threshold* or designated *cut-off concentration* (adapted from [19]).

F-distribution: See *distribution.*

F-test: See *significance test.*

geometric mean: See *mean.*

good laboratory practices (GLP): *Organizational* process and conditions under which *laboratory* studies are planned, performed, monitored, recorded and reported. Includes a system of protocols (*standard operating procedures*) recommended to be followed so as to avoid the production of unreliable and erroneous data (adapted from [20] and [21]).

goodness-of-fit: How well a model, a *theoretical distribution* or an equation matches actual data [10].

Grubbs test: See *outlier.*

harmonization: Bringing about agreement on terminology, concepts etc. so that different entities can interact based on the same terms of reference.

hypothesis test: See *significance test.*

hypothesis testing or **significance testing:** Process of assessing the statistical significance of a finding. It involves comparing empirically observed *sample* findings with theoretically expected findings, expected if the *null hypothesis* is true (see *significance test).* This comparison allows one to compute the *probability* that the observed outcomes could have been due to chance alone [22]. See *non-parametric test.*

> **alternative hypothesis:** Hypothesis that must be accepted if the *null hypothesis* is rejected [22].
>
> **null hypothesis (H_0):** Any hypothesis to be tested. The term null implies that there is no difference between the observed and known values other than that which can be attributed to random variation (adapted from [3] and [22]).

imprecision: See *precision.*

independent test result: *result* obtained in a manner not influenced by any previous results on the same or similar material [5].

influence quantity: Quantity, e.g. an environmental condition, that is not the subject of measurement but that influences the *result* (adapted from [16]).

in-house reference material: See *reference material.*

initial test: See *screening test.*

instrument linearity: Straight-line relationship between *concentrations* of *analyte* and instrument response, in which a change in concentration causes a proportional change in response [19].

interfering substance: Substance other than the *analyte* that gives a similar analytical response or alters the analytical result [2].

interlaboratory test comparisons: See *collaborative studies.*

internal standard: Addition of a fixed amount of a known substance that is not already present as a constituent of the *specimen* or *sample* in order to identify or quantify other components [18]. The physico-chemical characteristics of the internal standard should be as close as possible to those of the *analyte.*

interpretation: Explanation of what analytical results mean based on chemical, pharmacological, toxicological and statistical principles.

intralaboratory test comparisons: Organization, performance and evaluation of *tests* on the same or similar items or materials within the same *laboratory* in accordance with predetermined conditions [11].

laboratory: Facilities where analyses are performed by qualified personnel using adequate equipment.

Laboratory Information Management System: See *LIMS.*

least-squares: Statistical method of determining a *regression* equation, that is, the equation that best represents the relationship among the variables [10].

level of significance: *Probability* that a result would be produced by chance alone, i.e. the probability of incorrectly rejecting the *null hypothesis.* It is, therefore, the probability of making a *type I error.*

limit: Prescribed or specified maximum or minimum amount, quantity or number [23].

> **action limit:** Corresponds to a ±3 *standard deviation* from the *mean.* If an observed value falls outside the action limit, the cause must be identified immediately and remedial action taken.
>
> **confidence limit:** The limits of the *confidence interval.*
>
> **control limit:** The limits, on a *control chart,* that are used as criteria for action or for judging whether a set of data does or does not indicate lack of *statistical control.*
>
> **detection limit:** Smallest measured content from which it is possible to deduce the presence of the *analyte* with reasonable statistical certainty [17], [24].
>
> **quantitation limit:** The smallest measured content from which it is possible to quantitate the *analyte* with an acceptable level of *accuracy* and *precision.*
>
> **warning limit:** Corresponds to a ±2 *standard deviation* from the *mean.* Even if the method is under *statistical control,* approximately 5 per cent of results may be expected to fall outside the warning limits.

LIMS (Laboratory Information Management System): Software package for collating, calculating, controlling and disseminating analytical data. It can perform a variety of functions, from *specimen* or *sample* registration and tracking to processing captured data, *quality control,* financial control and report generation [25].

linear regression: Method of describing the relationship between two or more variables by calculating a *best-fitting* straight line or graph [10].

logbook: Book that records *laboratory* activities, e.g. instrumentation, maintenance of instrumentation, sample preparation and reagents.

maintenance: Activity of keeping something such as facilities, machines or instrumentation in good condition by regularly checking it and doing necessary repairs [6].

matrix: Material that contains the *analyte,* e.g. urine or blood.

maximum tolerable error: see *error.*

mean: When not otherwise specified, refers to *arithmetic mean* [10].

geometric mean: The *n*th root of the product of *n* individual values.

median: Middle value of a ranked set of data [7].

metabolite: Compound produced in the body as a result of biochemical processes.

method: Detailed (defined) *procedure* for performing an analysis. See *test procedure.*

method traceability: Property of a method whose measurements give results that can be related, with a given *uncertainty*, to a particular reference, usually a national or international standard, through an unbroken chain of comparisons (adapted from [5]).

mode: In statistics, the value or values occurring most frequently in a set of data [7].

negative: Indicates that the *analyte* is absent or below a designated *cut-off concentration.* "Not detected" is sometimes used as a synonym for negative, although this is not recommended.

negative control: See *control.*

none detected: Indicates the absence of an *analyte* within the specifications of the *test*(s) performed.

non-parametric test: Statistical method that makes no assumptions about the *distribution* of the *population* from which the *sample* data are taken (adapted from [3]).

one-tail test: *Hypothesis test* stated so that the chances of making a *type I error* are located entirely in one tail of a *probability distribution*, e.g. it is applicable if we wish to test only whether method A is more precise than method B and not whether method B is more precise than method A (adapted from [10]).

two-tail test: Statistical test in which the critical region (the region of rejection of the *null hypothesis*) is divided into two areas at the tails of the sampling distribution, e.g. it is applicable if we wish to test whether methods A and B differ in their precision (adapted from [10]).

normal distribution: See *distribution.*

not detected: The use of this term as a synonym for *negative* is not recommended.

null hypothesis: See *hypothesis testing.*

one-point calibration: Simplified *calibration* procedure using a single *calibrator* and a *blank.*

one-tail test: See *non-parametric test.*

organization: Company, corporation or institute (or part thereof, e.g. a *laboratory*), private or public, that has its own functions and administration. Some of the international organizations dealing with quality assurance are the International Association of Forensic Toxicologists (TIAFT), the International Federation of Clinical Chemistry (IFCC), the International Olympic Committee (IOC), the International Organization for Standardization (ISO), the International Programme on Chemical Safety (IPCS), the International Union of Pure and Applied Chemistry (IUPAC), and the Organisation for Economic Co-operation and Development (OECD).

outlier: Result that appears to differ unreasonably from the *population* of the other results. Tests for outliers include the following:

Cochran test: Compares the largest of a set of *variances* with the other variances in the set [26].

Dixon test: Compares the difference between a measurement and the one nearest to it in size with the difference between the highest and lowest measurements in the set [3].

Grubbs test: Now recommended to replace the *Dixon test* or to be used sequentially after the Dixon test [27]. The single Grubbs test statistic is calculated as the percentage decrease in the *standard deviation* of a set of results following the removal of either the highest or lowest value in the set, whichever gives the largest decrease in the standard deviation. The pair Grubbs test statistic is calculated in an analogous manner by removing the two highest, two lowest or else both the highest and the lowest values in the original set of results, whichever gives the lowest standard deviation. The presence of an outlier or a Grubbs outlier pair is indicated if the Grubbs statistic exceeds a critical value, which depends on the number of results in the set and which is given by a reference table [27], [28].

parametric test: Statistical techniques designed for use when data have certain characteristics, usually when they approximate a *normal distribution* and are measurable [10].

personnel: Persons qualified, by virtue of training and experience to carry out their assigned functions (adapted from [9]).

pharmacology: Study of the interactions of drugs with living systems [29].

plot: Representation of data on or by a graph.

population or **universe:** (Theoretical) entity defined as an entire group of people, things or events that have at least one trait in common [7].

population statistics: Statistical descriptors of the *population*, e.g. *mean*, *median*, *mode* or *standard deviation.*

positive: Indicates that the *analyte* is present at a level above a designated *cut-off concentration.*

positive control: See *control.*

power of test: *Probability* of rejecting the *null hypothesis* when it is false.

precision: Closeness of agreement between independent *test* results obtained under prescribed conditions [5]. It is generally dependent on *analyte concentration*, and this dependence should be determined and documented [25]. The measure of precision is usually expressed in terms of *imprecision* and computed as a *standard deviation* of the test results. Higher imprecision is reflected by a larger standard deviation. Independent test results refer to results obtained in a manner not influenced by any previous results on the same or similar material [5]. Precision covers *repeatability* and *reproducibility* [30].

presumptive: Describes things that are based on presumptions about what is probably true rather than on certainty [6].

presumptive positive: *Specimen* or *sample* that has been flagged as *positive* by *screening* but that has not yet been *confirmed* by an adequately sensitive alternate chemical method [2].

presumptive negative: *Specimen* or *sample* that has been flagged as *negative* by *screening*. Usually no further tests are carried out, so there is no certainty about its content.

presumptive test: See *screening test.*

preventive action: Action taken to eliminate the causes of a potential *deviation* or other undesirable situation in order to prevent occurrence [31].

probability: Mathematical measurement of how likely it is that something will happen, expressed as a fraction or percentage [19]. Values for statistical probability range from 1 or 100 per cent (always) to 0 or 0 per cent (never) [10]. The relative frequency obtained after a long run of measurements or results will give good approximations to the true probability [22]. It is also understood in other ways: as expressing in some undefinable way a "degree of belief", or as the limiting frequency of an occurrence in an infinite random series (adapted from [32]).

probability distribution: See *distribution.*

probability function: Function of a discrete *variate* that gives the *probability* that a specified value will occur.

procedure: Specified way to perform an activity. For *quality assurance* purposes, procedures should be written [31].

test procedure: Total operation necessary to perform the *analysis,* e.g. the preparation of the *specimen* or *sample*, of the *reference materials*, or of the reagents, the use of instruments and of formulas for the calculations (when the test is quantitative), the preparation and use of *calibration curves* and the determination of the number of *replicates.*

processed data: Raw data that have been acted upon to make them clearer or more readily usable.

proficiency testing: Ongoing process in which a series of proficiency *specimens* or *samples*, the characteristics of which are not known to the participants, are sent to *laboratories* on a regular basis. Each laboratory is tested for its *accuracy* in identifying the presence (or *concentration*) of the drug using its usual *procedures.* An *accreditation body* may specify participation in a particular proficiency testing scheme as a requirement of *accreditation.*

qualitative test: *Test* that determines the presence or absence of specific drugs or *metabolites* in the *specimen* or *sample* [19].

quality assessment: Overall system of activities whose purpose is to provide assurance that the overall quality control job is being done effectively. It involves a continuing evaluation of the products produced and of the performance of the production system [33].

quality assurance (QA): System of activities whose purpose is to provide to the producer or user of a product or a service the assurance that it meets defined standards of quality with a stated level of confidence [33].

quality assurance management: All activities of the overall management function that determine and implement *quality policy*, objectives and responsibilities (adapted from [31]).

quality assurance programme: Internal control system designed to ascertain that the *studies* are in compliance with the principles of *good laboratory practices* (GLP) [21].

quality audit: Systematic and independent examination to determine whether quality activities and related results comply with planned arrangements and whether these arrangements are implemented effectively and are suitable to achieve the objectives [31].

quality control: Overall system of activities whose purpose is to control the quality of a product or service so that it meets the needs of users. The aim is to provide quality that is satisfactory, adequate, dependable and economic [33].

> **external quality control:** See *proficiency testing.*
>
> **internal quality control:** Set of *procedures* undertaken by a *laboratory* for continuous monitoring of operations and results in order to decide whether the results are reliable enough to be released. Quality control of analytical data primarily monitors the batchwise *trueness* of results on quality control *specimens* or *samples* and *precision* on independent *replicate analysis* of test materials [5].

quality management: That aspect of the overall management function that determines and implements the quality policy [31].

quality manual: Document stating the general quality policies, procedures and practices of an organization [34].

quality policy: Statement by top management regarding the laboratory's adherence to principles of quality. It may set forth codes of practice or ethics.

quality system: The organizational structure, responsibilities, procedures, processes and resources for implementing *quality management* [31]. In a laboratory it refers to the total features and activities of a laboratory aimed at producing accurate work and a high-quality product [7].

quantitation limit: See *limit.*

quantitative test: *Test* to determine the quantity of drug or *metabolite* present in a *specimen* [19].

random error: See *error.*

random sample: *Sample* taken in such a way that all the members of the *population* have an equal chance of being included, that is, each is chosen entirely by chance [3].

range: *Concentration* interval for which acceptable *accuracy* and *precision* can be achieved (adapted from [13]). Statistically it is the difference between the minimum and the maximum values of a set of measurements [7].

raw data: Data that are in their original state and have not been processed [10].

record: Document that furnishes objective evidence of activities performed or results achieved [31].

recovery: Percentage of the drug, *metabolite* or *internal standard* originally in the *specimen* or *sample* that reaches the end of the *procedure* [35].

reference material: Material or substance one or more properties of which are sufficiently well established to be used for *calibrating* an apparatus, assessing a measurement method or assigning values to materials [9].

> **in-house reference material:** Material whose composition has been established by the user *laboratory* by several means, by a *reference method* or in collaboration with other laboratories [5].

reference method or **standard consensus method:** Method developed by *organizations* or groups that use *collaborative studies* or similar approaches to *validate* it. Its value depends on the authority of the organizations that sponsor it (adapted from [7]).

reference preparation: Processed *reference material.*

reference standard: A standard, generally of the highest quality available at a given location, from which measurements made at that location are derived [9].

regression analysis: Method of explaining or predicting the *variability* of a dependent *variable* using information about one or more independent variables. Also, techniques for establishing regression equations [10].

regression curve: Curve that comes closest to approximating a distribution of points in a *scatter diagram* [10].

relative error: See *error.*

relative frequency: Number calculated by dividing the number of values with a certain characteristic by the total number of values [10]. Also, the frequency of an event that would occur in the long run given the *probability* of the event [22].

relative standard deviation: See *coefficient of variation.*

reliability: Extent to which an experiment, *test* or measuring *procedure* yields *accurate* results in repeated trials [23].

repeatability: Closeness of the agreement between the results of successive measurements of the same *analyte* made under repeatable conditions, e.g. same method, same material, same operator and same *laboratory* and carried out in a narrow time period (adapted from [16], [25]). Results should be expressed in terms of the repeatability *standard deviation*, the repeatability *coefficient of variation* or the *confidence interval* of the *mean* value.

replicate analysis: Multiple analysis of separate portions of a *test* material using the same test method under the same conditions, e.g. same operator, same apparatus, same *laboratory* [18].

report: Document containing a formal statement of results of *tests* carried out by a *laboratory.* It should include the information necessary for the *interpretation* of the test results [9].

representative sample: Statistically, a *sample* that is similar to the *population* from which it was drawn. When a sample is representative, it can be used to make inferences about the population. The most effective way to get a representative sample is to use random methods to draw it [10]. Analytically, it is a *specimen* or *sample* that is a portion of the original material selected in such a way that it is possible to relate the analytical results obtained from it to the properties of the original material [25].

reproducibility: Closeness of the agreement between the results of successive measurements of the same *analyte* in identical material made by the same method under different conditions, e.g. different operators and different laboratories and considerably separated in time (adapted from [16], [25]). Results should be expressed in terms of the reproducibility *standard deviation*, the reproducibility *coefficient of variation* or the *confidence interval* of the *mean* value.

resolution: Ability to distinguish meaningfully between closely adjacent values [16].

result: Information obtained from a *test* or series of tests. Usually it refers to *processed data.*

retention sample or **specimen**: Amount of material equivalent in quantity to the *assay specimen* or *sample* and taken from the consignment in a manner similar to that used to assay the sample or specimen. It should be stored under specified conditions [36].

review: Evaluation of *laboratory* results to ensure that they have been correctly *interpreted.*

robustness or **ruggedness:** Capacity of a test to remain unaffected by small variations in the *procedures.* It is measured by deliberately introducing small changes to the method and examining the consequences [25]. See *influence quantity.*

sample: Analytically equivalent to *specimen,* it is a representative portion of the whole material to be tested. Statistically, it is a set of data obtained from a *population* [7].

sample statistics: Statistical descriptors of the *sample,* e.g. *mean, median, mode, standard deviation, range* or size.

sampling: Analytically, the whole set of operations needed to obtain a *sample* or *specimen,* including planning, collecting, recording, labelling, sealing, shipping etc. Statistically it is the process of determining properties of the whole *population* by collecting and analysing data from a representative segment of it [23].

scatter diagram or **scatter plot** or **scattergram:** Pattern of points that results from plotting two variables on a graph. Each point or dot represents one subject or unit of analysis and is formed by the intersection of the values of the two variables [10].

screening test or **initial test** or **presumptive test:** First test carried out on a *specimen* or *sample* for the purpose of determining a presumption of a *positive* or *negative* assay. Usually, presumptive positives are followed by a *confirmatory test.*

selectivity: Extent to which a method can determine particular *analyte*(s) in a complex mixture without interference from the other components in the mixture [25]. A method that is perfectly selective for an analyte or group of analytes is said to be *specific.*

sensitivity: Difference in *analyte concentration* corresponding to the smallest detectable difference in the response of the method. It is represented by the slope of the *calibration curve.* Sometimes it is used, erroneously, to mean *detection limit.*

Shewhart chart: See *control chart.*

significance test or **hypothesis test:** Statistical *test* whose purpose is to draw a conclusion about a *population* using data from a *sample.* It is used to determine the likelihood that observed characteristics of samples have occurred by chance alone in the population from which the samples were selected. Frequently used significance tests include the following:

> **comparison-of-means test** or **t-test:** Compares the *mean* of the results from one sample taken from a given population with the mean of the results from a second sample taken from the same population, with the two sets of results having, for example, been produced by different analytical methods. It answers the question, "Are the two means significantly different?" The *null hypothesis* is that the means are not significantly different and that the samples are therefore part of the same population. It is assumed that the *variances* of the samples are the same and that the samples are representative of the whole population. The larger the number of results for each sample, the more likely

this is to be true. Statistical comparison of the means will indicate whether any differences between the samples could have arisen by chance alone. The *t-test* is used under particular circumstances, for example, when the size of the samples is small (usually less than 20) or when a single sample is taken from a population for which the variance is unknown.

variance-ratio test or **F-test:** In the *comparison-of-means test* it is assumed that the *variance* of each sample is the same. A variance-ratio test is used to check if this assumption is reasonable.

significant figures: Number of figures that are consistent with the *precision* of the test.

skewness: Said of measures or scores that are bunched on one side of a central tendency parameter (*mean, median, mode*) and trail out on the other. The more skewness in a *distribution*, the more variability in the scores [10]. Also used to refer to asymmetry in, for example, a chromatographic peak shape ("tailing" and "fronting").

solution: Liquid in which a solid substance or a gas has been dissolved [6].

specification: Statement of requirements, usually in written form.

specificity: See *selectivity.*

specimen: Analytically, equivalent to *sample.* In the context of this Glossary, any biological material for examination, study or analysis.

spiked sample: A test material containing a known addition of *analyte* (adapted from [5]).

split-level model: Statistical model that splits the study sample according to a predetermined assumption so that only a portion of the cases falls into the category of interest, e.g. so that only some of the specimens that were positive for group A of drugs will be positive for group B.

split specimen or **sample:** Practice of dividing a *specimen* or *sample.* A urine specimen, for example, may be divided into two portions, one of which may be submitted for analysis and the other preserved by freezing for *confirmatory analysis* or reanalysis (adapted from [19]).

stability: Resistance to decomposition or other chemical changes, or to physical disintegration [23].

standard addition: The addition of a known amount of a pure component supposed to be present as a constituent of the specimen or sample in order to verify and quantitate this component [18]. Operationally, a measurement is made on the specimen or sample, a known amount of the desired constituent is added, the modified specimen or sample is remeasured, and the amount of the constituent originally present is determined by proportionation [37].

standard consensus method: See *reference method.*

standard deviation: A statistic that shows the spread or dispersion of scores in a *distribution* of scores. It is calculated by taking the square root of the *variance* [10]. It is applicable to all kinds of repeated measurements, e.g. between-*batch*, within-batch, *repeatability* and *reproducibility.*

standard operating procedures (SOP): Written *procedures* that describe how to perform certain *laboratory* activities [21].

standard solution: Solution of known *concentration* prepared from characterized material.

statistical control: A *procedure* is in statistical control when results consistently fall within established *control limits*, that is, when they have constant *mean* and *variance* (adapted from [7]). Statistical control should be monitored graphically with *control charts* [5].

statistical correlation: Extent to which two or more things are related to one another. This is usually expressed as a *correlation coefficient* [10].

statistical significance: Said of a value or measure of a *variable* when it is larger or smaller than would be expected by chance alone. Statistical significance does not necessarily imply practical significance [10].

stock solution: Concentrated *standard solution* used to prepare *calibrators.*

study: Experiment or set of experiments performed to obtain information on a particular subject.

surrogate analyte: See *analyte.*

surveillance: Monitoring of certain activities to ensure that specified requirements have been fulfilled (adapted from [31]).

survey: Study conducted among *organizations* to collect information on their activities or performance.

systematic error: See *error.*

target analyte: See *analyte.*

target value: See *value.*

t-distribution: See *distribution.*

test: Technical operation to determine one or more characteristics of or to evaluate the performance of a given product, material, equipment, organism, physical phenomenon, process or service according to a specified procedure (adapted from [9]).

test linearity: Ability within a given range to obtain *test* results directly proportional to the *concentration* (amount) of *analyte* in the *specimen* or *sample* [38].

test procedure: See *procedure.*

theoretical probability distribution: See *distribution.*

threshold: A particular, significant amount, level or *limit*, at which something begins to happen or take effect. See *cut-off concentration.*

total error: See *error.*

traceability: Ability to trace the history, application or location of an entity by means of recorded identification [31]. See also *chain of custody.*

trueness: Closeness of agreement between the *average* value obtained from a large series of *test* results and an accepted reference or *true value.*

true value: See *value.*

t-test: See *significance test.*

type I errors and **type II errors:** See *error.*

uncertainty: A parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the *analyte.*

universe: See *population.*

validated method: Method whose performance characteristics meet the specifications required by the intended use of the analytical results (adapted from [25]). Some of the performance characteristics to be evaluated are *limit of detection*, *limit of quantitation*, *linearity*, *precision*, *range*, *ruggedness*, *selectivity* and *specificity*, and *trueness*.

validation: Confirmation by examination and provision of objective evidence that the particular requirements for a specific intended use are fulfilled [31].

value: The expression of a quantity in terms of a number and an appropriate unit of measurement [16].

> **assigned value:** Best available estimate of the *true value* [39].
>
> **consensus value:** Value produced by a group of experts or referee *laboratories* using the best possible methods. It is an estimate of the *true value*.
>
> **estimate value (statistical):** Value(s) of *population* characteristic(s) obtained from *sample* data.
>
> **target value:** Numerical value of a measurement result that has been designated as a goal for measurement quality [40].
>
> **true value:** Value that characterizes a quantity perfectly defined in the conditions that exist when that quantity is considered. The true value of a quantity is an ideal concept and, in general, cannot be known exactly [16].

variability: Spread or dispersion of scores in a group of scores; the tendency of each score to be unlike the others [10].

variable: Generally, any quantity that varies. More precisely, a quantity that may take any one of a specified set of *values* [12].

variance: Statistic that shows the spread or dispersion of scores in a distribution of scores. It is calculated as the sum of the squares of the differences between the individual *values* of a set and the *arithmetic mean* of the set, divided by one less than the number of values [7], [10].

variance-ratio test: See *significance test*.

variate: In contradistinction to a *variable*, a variate is a quantity that may take any of the *values* of a specified set with a specified *relative frequency* or *probability*. It is often known as a random variable [12].

verification: Confirmation by examination and provision of objective evidence that specified requirements have been fulfilled [31].

warning limit: See *limit*.

working standard solutions: *Standard solutions* prepared by diluting the *stock solution* containing the *concentrations* used to establish the *calibration curve*.

z-distribution: See *distribution*.

z-score: Number of *standard derivation* units that separate a *value* from its *mean* (based on [41]).

References

1. United Nations Drug Control Programme, "Report of the Consultative Meeting on Quality Assurance and Good Laboratory Practices", Glasgow, 2-6 November 1992.
2. National Institute for Drug Abuse, *Urine Testing for Drugs of Abuse,* Research Monograph 73 (Rockville, Maryland, Department of Health and Human Services, 1986).
3. J. C. Miller and J. N. Miller, *Statistics for Analytical Chemistry* (Chichester and New York, Ellis Horwood, 1984).
4. R. Wennig, *Practical Compendium for Health Professionals: Drugs of Abuse Currently Used in Europe,* Publication CEC/V/E/I/Lux 92 (Luxembourg, Commission of the European Communities, Health and Safety Directorate, 1992).
5. International Olympic Committee/Reference Materials Committee of the International Organization for Standardization, "Quality control of analytical data produced in chemical laboratories", Publication 271, draft protocol presented to the Fifth International Symposium on the Harmonization of Internal Quality Assurance Schemes for Analytical Laboratories, Washington, D.C., 23 July 1993.
6. *PONS English Language Dictionary* (London, Collins Cobuild., 1987).
7. F. M. Garfield, Quality Assurance Principles for Analytical Laboratories (Arlington, Virginia, Association of Official Analytical Chemists, 1991).
8. American Society of Crime Laboratory Directors, *Laboratory Accreditation Manual* (Norfolk, Virginia, 1992).
9. International Organization for Standardization/International Electrotechnical Commission, Guide 25: *General Requirements for the Competence of Calibration and Testing Laboratories* (Geneva, 1990).
10. W. P. Voight, *Dictionary of Statistics and Methodology: A Non-technical Guide for the Social Sciences* (Thousand Oaks, California, Sage Publications, 1993).
11. International Organization for Standardization/International Electrotechnical Commission, *Guide 2: General Terms and Their Definitions concerning Standardization and Related Activities* (Geneva, 1991).
12. M. G. Kendall and W. R. Buckland, *A Dictionary of Statistical Terms* (London, Longman Group, 1976).
13. C. Visher, *A Comparison of Urinalysis Technologies for Drug Testing in Criminal Justice* (Washington, D.C., National Institute of Justice and Bureau of Justice Assistance, 1991).
14. Society of Forensic Toxicologists Inc./American Academy of Forensic Sciences, *Forensic Toxicology Laboratory Guidelines,* 1991.
15. K. Eckschlager, *Errors, Measurement and Results in Chemical Analysis* (London, Van Nostrand Reinhold, 1969).
16. International Union of Pure and Applied Chemistry, *Spectrochemica Acta,* 1978.
17. International Organization for Standardization, *International Vocabulary of Basic and General Terms Used in Metrology* (Geneva, 1984).
18. National Measurement Accreditation Service, NIS 46, *Accreditation for Forensic Analysis and Examination* (Teddington, Middlesex, 1992).

19. B. S. Finkle, R. V. Blanke and J. M. Walsh, eds., *Technical, Scientific and Procedural Issues of Employee Drug Testing* (Rockville, Maryland, NIDA, 1990).
20. J. A. Timbrell, *Introduction to Toxicology* (London, Taylor and Francis, 1989).
21. Organisation for Economic Co-operation and Development, *The OECD Principles of Good Laboratory Practice,* Environmental Monograph No. 45, Paris, 1992.
22. D. McCormick and A. Roach, *Measurement, Statistics and Computation* (Chichester, Wiley, 1988).
23. *Webster's Third New International Dictionary of the English Language* (Springfield, Massachusetts, Merriam, 1971).
24. Food and Agriculture Organization of the United Nations/World Health Organization, Codex Alimentarius Commission, *Criteria to Limit the Number of False Positive and False Negative Results for Analytes Near the Limit of Determination,* CX/MAS 92/15, Budapest, 1992.
25. EURACHEM/WELAC Chemistry Working Group, *Accreditation for Chemical Laboratories: Guidance on the Interpretation of the EN45000 Series of Standards and ISO/IEC Guide 25* (Teddington, Middlesex. EURACHEM/WELAC, 1993).
26. R. Caulcutt and R. Boddy, *Statistics for Analytical Chemists* (New York, Chapman and Hall, 1983).
27. W. Horwitz, "Protocol for the design, conduct and interpretation of collaborative studies", *Pure and Applied Chemistry,* vol. 60 (1988), pp. 855-864.
28. P. C. Kelly, "Outlier detection in collaborative studies", *Journal of the Association of Official Analytical Chemists,* vol. 73 (1990), pp. 58-64.
29. B. G. Katzungt and A. J. Trevor, *Pharmacology* (London, Prentice Hall International, 1993).
30. European Community, *Guideline Criteria for Reference Methods,* BNL SP/Lab/div (92) 5 (1992), p. 27.
31. International Organization for Standardization/Development Information System 8402, *Quality Management and Quality Assurance Vocabulary* (Geneva, 1991).
32. F. H. C. Marriot and M. G. Kendall, *A Dictionary of Statistical Terms,* 5th ed. (Harlow, Essex and New York, Longman Scientific & Technical/Wiley, 1990).
33. J. K. Taylor, *Quality Assurance of Chemical Measurements* (Boca Raton, Florida, Lewis, 1987).
34. L. Stebbing, *Quality Assurance: the Route to Efficiency and Competitiveness,* 3rd ed. (Chichester and New York, Ellis Horwood, 1993).
35. A. C. Moffatt and M. D. Osselton, "Characterisation of quantitative methods for the analysis of drugs in blood in forensic toxicology", *Bulletin of the International Association of Forensic Toxicologists,* vol. 20, No. 3 (1990), pp. 36-41.
36. World Health Organization, *Good Laboratory Practices in Governmental Drug Control Laboratories,* WHO/PHARM/84.512/Rev.2, 1984.
37. I. M. Kolthoff, P. J. Elving and E. J. Meehan, eds., *Treatise on Analytical Chemistry,* 2nd ed. (New York, John Wiley, 1978).
38. Commission of the European Communities, Directorate III, *Addendum, Guidelines on the Quality, Safety and Efficacy of Medical Products for Human Use,* vol. III, Publication CEC/III/EN/9104/90 (Brussels, 1990).
39. Analytical Methods Committee, "Proficiency testing of analytical laboratories: organization and statistical assessment", *Analyst,* vol. 117 (1992), pp. 97-117.
40. International Organization for Standardization/REMCO, "The international harmonized protocol for the proficiency testing of (chemical) analytical laboratories", Publication 263, Geneva, 1992.
41. W. Horowitz and R. Albert, "Experience with the IUPAC-1987 Harmonized Protocol for Method-Performance Studies: suggestions for revision and application to internal quality control systems", *Quality Assurance for Analytical Laboratories,* M. Parkany, ed. (Cambridge, Royal Society of Chemistry, 1993).

Bibliography

Canada. Ministry of Health and Welfare. Guide to the drugs directorate laboratory activities quality assurance programme. 1991.

DeAngelis, G. G. Testing and screening for drugs of abuse. New York, Dekker, 1976.

International Organization for Standardization. Guide 58: calibration and testing laboratory accreditation systems — general requirements for operation and recognition. 1993.

International Union of Pure and Applied Chemistry. Glossary of terms used in toxicology. Forthcoming.

Jardin, A. K. S., J. D. McFarlane *and* C. S. Greensted. Statistical methods for quality control. Heinemann, 1975.