A Bioassay Glossary

Robert Singer, David M. Lansky, and Walter W. Hauck

In the biopharmaceutical industry, white papers, standards, and regulatory documents are often published in draft form to elicit comments from industry over a designed review period. Occasionally, *BioProcess International* is given an opportunity to widen that forum by re- or copublishing one of these documents. In this case, because a large number of our readers are concerned with bioassay development and methods, we felt that it was important to invite them into this discussion.

This glossary came about when the United States Pharmacopeia (USP) empaneled biological and statistical experts to develop chapters pertaining to the practice of biological assays. The chapters include a rewrite of General Chapter *Design and Analysis of Biological Assays* <111> and the development of two new General Chapters, *Design of Biological Assays* <1032> and *Validation of Biological Assays* <1033>. In their discussions about the three chapters — and indeed, throughout all portions of *USP–NF* that relate to biological assays — the experts came to realize a need for clear and uniform terminology in their discussions. Accordingly, they developed a glossary pertinent to biological assays.

The nascent glossary was published for the USP audience in its bimonthly journal of standards development and compendial review, the *Pharmacopeial Forum* (*PF*), the means through which USP develops and revises standards by a process of public review and comment. (All changes and additions to *USP–NF* are first proposed in the *PF* to invite public comment). The authors hope to resolve and/or clarify differences in measurement terminology for well- (small molecules) and poorly characterized (biologicals and biotechnological) moieties, hereafter referred to collectively as drug substances.

WWW.PHOTOS.COM

In many cases the terms developed herein have common usages or appear elsewhere, for example in contemporary references by the International Conference on Harmonization (**1**), the Food and Drug Administration (FDA), and USP. For some of terms below, the derivation may be clear. Rather than claim originality, the authors seek to associate with this work a compendial perspective that will provide clarity broadly going forward; consistency with previous authoritative usage; and a useful focus on the bioassay context.

Comments on this proposed glossary can be sent to Tina Morris, PhD, Director, Biologics and Biotechnology, Standards Development, USP, e-mail tsm@usp.org.

PRODUCT FOCUS: ALL BIOPHARMACEUTICALS

PROCESS FOCUS: ASSAY, QA/QC, MANUFACTURING, AND ANALYSIS

WHO SHOULD READ: ANALYTICAL, QA/QC, PROCESS DEVELOPMENT, REGULATORY/COMPENDIAL AFFAIRS

KEYWORDS: US PHARMACOPEIA, BIOASSAY, GLOSSARY, PRECISION, VALIDATION, REGULATORY/COMPENDIAL AFFAIRS

LEVEL: INTERMEDIATE

GENERAL TERMS RELATED TO BIOASSAYS

ANALYTICAL PROCEDURE: detailed description of the steps necessary to perform the assay.

Notes: 1. The description may include but is not limited to the sample, the reference standard and the reagents preparations, use of the apparatus, generation of the standard curve, use of the formulae for the calculation, etc.

2. The FDA guidance provides a list of information that typically should be included in the description of an analytical procedure (**2**).

ASSAY: analysis (as of a drug) to determine the quantity of one or more components or the presence or absence of one or more components.

Notes: 1. *Assay* often is used as a verb synonymous with *to determine*, as in, "I will assay the material for impurities." In this glossary, *assay* is a noun and is synonymous with the *analytic procedure* (*protocol*) 2. The phrase, "to run the assay" means to perform the analytical procedure(s) as specified.

ASSAY DATA SET: the set of data used to determine a single potency or relative potency for all samples included in the bioassay.

Notes: 1. The definition of an assay data set can be subject to interpretation as necessarily a *minimal* set. It is important to understand that it may be possible to determine a potency or

relative potency from a set of data, but not to do this well. It is *not* the intent of this definition to mean that an assay data set is the *minimal* set of data that can be used to determine a relative potency. In practice, an assay data set should include, at least, sufficient data to assess similarity (qv). It also may include sufficient data to assess other assumptions.

2. It is also not an implication of this definition that assay data sets used together in determining a reportable value (qv) are necessarily independent from one another, although it may be desirable that they be so. When a run (qv) consists of multiple assay data sets, independence of assay sets within the run must be evaluated.

BIOASSAY, BIOLOGICAL ASSAY (THESE TERMS ARE INTERCHANGEABLE): analysis (as of a drug) to quantify the biological activity(ies) of one or more components by determining its capacity for producing an expected biological activity, expressed in terms of units.

Notes: 1. Typically a bioassay involves controlled administration of the drug substance to living matter, in vivo or in vitro, followed by observation and assessment of the extent to which the expected biological activity has been manifested. 2. The description of a bioassay includes the analytic procedure, which should include the statistical design for collecting data, and the method of statistical analysis that eventually yields the estimated potency or relative potency. 3. Bioassays can be either *direct* or *indirect*.

Direct bioassays measure the concentration of a substance that is required to elicit a specific response. For example, the potency of digitalis can be directly estimated from the concentration required to stop a cat's heart. In a direct assay, the response must be distinct and unambiguous. The substance must be administered in such a manner that the exact amount (threshold concentration) needed to elicit a response can be readily measured and recorded.

Indirect bioassays compare the magnitude of responses for nominally equal concentrations of reference and test preparations, rather than test and reference concentrations that are required to achieve a specified response. Most biological assays in USP are indirect assays that are based on either quantitative or quantal (yes/no) responses.

POTENCY: the specific ability or capacity of the product, as indicated by appropriate laboratory tests or by adequately controlled clinical data obtained through the administration of the product in the manner intended, to effect a given result.

Notes: 1. A wholly impotent sample has no capacity to produce the expected response, as a potent sample would. Equipotent samples produce equal responses at equal dosages. Potency is typically measured relative to a reference standard or preparation that has been assigned a single unique value (e.g., 100%) for the assay; see *relative potency*. At times, additional qualifiers are used to indicate the physical standard employed (e.g., "international units").

2. Some biological products have multiple uses and multiple assays. For such products there may be different reference lots that do not have consistently ordered responses across a collection of different relevant assays.

RELATIVE POTENCY: a measure obtained from the comparison of a test to a reference drug substance on the basis of capacity to produce the expected biological activity.

Notes: 1. A frequently invoked perspective is that relative potency is the degree to which the test preparation is diluted or concentrated relative to the standard. 2. Relative potency is unitless and is given definition, for any test material, solely in relation to the reference material and the assay.

REPORTABLE VALUE: the potency or relative potency estimate of record that is intended to achieve such measurement accuracy and precision as are required for use.

Notes: 1. The reportable value is the value that will be compared to a product specification. The specification may be in the USP monograph, or it may be set by the company, e.g., for product release.

2. The term *reportable value* is inextricably linked to the "intended use" of an analytical procedure. Tests are performed on samples in order to yield results that can be used to evaluate some parameter of the sample in some manner. One type of test may be configured in two different ways because the resulting data will be used for two different purposes (e.g., lot release vs. stability). The reportable value would likely be different even if the mechanics of the test itself were identical. Validation is required to support the properties of each type of reportable value. In practice there may be one physical document that is the analytical procedure used for more than one application, but each application must be detailed separately within that document. Alternatively, there may be two separate documents for the two applications.

3. When the inherent variability of a biological response, or that of the log potency, precludes a single assay data set's attaining a value sufficiently accurate and precise to meet an assay specification, the assay, or analysis data set, may consist of multiple assay data sets, as necessary. The number of assay data sets needed depends on the assay's accuracy and precision and on the intended use and hence the properties of the reported value and is influenced by factors such as the type and variability of the biological activity being studied.

RUN: that performance of the analytical procedure that can be accomplished by a laboratory team in a set time with a given unique set of assay factors (e.g., standard preparations).

Notes: 1. There is no necessary relationship of *run* to *assay data set* (qv). The term *run* is laboratory specific; run relates to the physical capability of a team and its physical environment. An example of a run is given by one analyst's simultaneous assay of several samples in one day's bench work. During the course of a single run, it may be possible to determine multiple reportable values. Conversely, a single assay or reportable value may include data from multiple runs.

2. From a statistical viewpoint, a run is one realization of the factors associated with intermediate precision (qv). It is good practice to associate runs with factors that are significant sources of variation in the assay. For example, if cell passage number is an important source of variation in the assay response obtained, then each change in cell passage number initiates a new run. If the variance associated with all factors that could be assigned to runs is negligible, then the influence of runs

can be ignored in the analysis and the analysis can focus on combining independent analysis data sets.

3. When a run contains multiple assays, caution is required regarding the independence of the assay results. Factors that are typically associated with runs and that cause lack of independence include cell preparations, groups of animals, analyst, day, a common preparation of reference material, and analysis with other data from the same run. Even though a strict sense of independence may be violated because some elements are shared among the assay sets within a run, the degree to which independence is compromised *may* have negligible influence on the reportable values obtained. This would need to be verified and monitored.

Similar preparations (similarity): the property of two preparations such that each behaves as a dilution (or concentration) of the other.

Notes: Similarity is fundamental to methods for determination of relative potency. Bioassay similarity requires that the reference and test samples should be sufficiently similar for legitimate calculation of relative potency. Given demonstration of similarity, a relative potency can be calculated, reported, and interpreted. Relative potency is valuable in assessing consistency and also intra- and inter-manufacturer comparability in the presence of change. In the absence of similarity, a meaningful relative potency cannot be reported or interpreted. 2. The practical *consequence* of similarity is a comparable form of dose and/or concentration–response behavior. 3. Failure to statistically demonstrate dissimilarity between a reference and a test sample does not amount to demonstration of similarity. To assess similarity it is not sufficient to fail to find evidence that a reference and a test sample are not similar.

Terms Related to Running a Bioassay

Configuration, assay: the arrangement of experimental units (qv) by number, position, location, temporal treatment, etc. and the corresponding test, control, or reference sample dilution that will be applied to each.

Notes: 1. The assay configuration must be specified in the formalized assay protocol. 2. Assay configuration can include nested dimensions like plate design, multiple plates per day, single plates on multiple days, etc. The configuration will depend on what the variance analysis (performed during assay development) reveals regarding sources of variability on assay response.

Sample suitability: a sample is suitable (may be described as having a potency) if its response curve satisfies certain properties defined in the protocol.

Note: Most significant of these properties is that of similarity to the reference standard response curve. If this property of similarity is satisfied, then the sample is suitable for the assay and can be described via a relative potency estimate.

System suitability: the provision of assurance that the laboratory control procedure is capable of providing legitimate measurements as defined in the validation report.

Notes: 1. System suitability may be thought of as an assessment of current validity achieved at the time of assay performance. An example is provided by positive and negative controls giving values within their normal ranges, ensuring that the assay system is working properly. 2. As described in USP General Chapter *Validation of Compendial Methods* <1225> and ICH Q2B, system suitability testing is an integral part of many analytical procedures. The tests are based on the concept that the equipment, electronics, analytical operations, and samples to be analyzed constitute an integral system that can be evaluated as such. System suitability test parameters to be established for a particular procedure depend on the type of procedure being validated. *USP–NF* is a source of many system suitability tests.

Terms Related to Precision and Accuracy

Accuracy: an expression of the closeness of agreement between the value that is accepted either as a conventional true value or an accepted reference value and the value found.

Notes: 1. ICH and ISO give the same definition of accuracy. However, ISO specifically regards accuracy as having two components, bias and precision (**3**). That is, to be accurate as used by ISO, a measurement must be both "on target" (have low bias) and precise. In contrast, ICH Q2A says that accuracy is sometimes termed *trueness* but does not define *trueness*. ISO defines *trueness* as the "closeness of agreement between the average value obtained from a large series of test results and an accepted reference value" and indicates that "*trueness* is usually expressed in terms of bias."

The 2001 FDA guidance on Bioanalytical Method Validation defines accuracy in terms of "closeness of *mean* test results" (emphasis added) and is thus consistent with the ICH usage. This glossary adopts the ICH approach. That is, it uses the phrase "accurate and precise" to indicate low bias (accurate) and low variability (precise).

2. Considerable caution is needed when using or reading the term *accuracy*. In addition to the inconsistency between ICH and ISO, common usage is not consistent.

Error, types of: two sources of uncertainty that affect the results of a biological assay are *systematic* and *random error*.

A ***systematic error*** is one that happens with similar magnitude and consistent direction repeatedly. This introduces a *bias* in the determination. Effective experimental design, including randomization and/or blocking, can reduce systematic error.

A ***random error*** is one whose magnitude and direction vary without pattern. Random error is an inherent variability or uncertainty of the determination. Transformation of systematic into random error will increase the robustness of a biological assay and allow a comparatively simple analysis of assay data.

Intermediate precision: expresses within-laboratory precision associated with changes in operating conditions.

Notes: 1. Factors contributing to intermediate precision involve anything that can change within a given laboratory and that may affect the assay, including different days, different analysts, different equipment, etc. Intermediate precision is thus "intermediate" in scope between the extremes of repeatability and reproducibility. 2. Any statement of intermediate precision should include clarification about which factors varied.

For example, "The intermediate precision associated with changing equipment and operators is . . . "

3. There can also be value in separately identifying the precision associated with each source; e.g., inter-analyst precision. This may be part of assay development and validation, when there is value in identifying which are the important contributors to intermediate precision.

4. When reporting intermediate precision, particularly for individual sources, analysts should take care to distinguish between intermediate precision variance and components of that variance. The variance includes repeatability and thus must be necessarily at least as large as the repeatability variance. A variance component, e.g., for analyst, is also a part of the intermediate precision variance for analyst, but it could be negligible and need not be larger in magnitude than the repeatability variance.

Precision: the closeness of agreement (degree of scatter) between a series of measurements obtained from multiple sampling of the same homogeneous sample under the prescribed conditions.

Notes: 1. Precision may be considered at three levels: repeatability (qv), intermediate precision (qv), and reproducibility (qv). 2. Precision should be investigated using homogeneous, authentic samples. However, if it is not possible to obtain a homogeneous sample, precision may be investigated using artificially prepared samples or a sample solution. 3. Precision is usually expressed as the variance, standard deviation, or coefficient of variation.

Repeatability: the expression of the precision under the same operating conditions over a short interval of time.

Notes: 1. ICH Q2A says that repeatability is also termed "intra-assay" precision. In the bioassay context, the better term is "intra-run," and a "short interval of time" is meant to connote "within-run." 2. The idea of a "short interval of time" can be problematic with bioassay. If a run takes multiple weeks and consists of a single assay set, then intra-run precision cannot be determined. Alternatively, if a run consists of two assay data sets and a run can be done in a single day, repeatability of the relative potency determination can be assessed. 3. Operating conditions will include, but not be limited to, equipment and analyst.

Reproducibility: expresses the precision between laboratories.

Notes: 1. Reproducibility includes contributions from repeatability and all factors contributing to intermediate precision as well as any additional contributions from inter-laboratory differences. 2. Reproducibility will apply to collaborative studies, such as those for standardization or portability of methodology. Depending on the design of the collaborative study, it may be possible to separately describe variance components associated with intra- and inter-laboratory sources of variability.

Specificity: the ability to assess unequivocally the analyte in the presence of components that may be expected to be present.

Note: Typically these components may include impurities, degradants, matrix, etc.

Terms Related to Characterization and Validation

Detection limit: the lowest amount of analyte in a sample that can be detected but not necessarily quantified or quantified to any given level of precision and accuracy.

Linearity, bioassay: the ability (within a given range) of a bioassay to obtain log relative potencies that are directly proportional to the log relative potency of the sample.

Notes: 1. Bioassay linearity, sometimes called *dilutional linearity*, is demonstrated across a range of known relative potency values by considering a plot of true log potency vs. observed log potency. If that plot yields an essentially straight line with a y-intercept of 0 and a slope of 1, the assay has direct proportionality. If that plot yields an essentially straight line but either the y-intercept is not 0 or the slope is not 1 (or both), the assay has a proportional linear response.

2. To assess whether the slope is (near) 1.0 requires an a priori equivalence or indifference interval. It is not proper statistical practice to test the null hypothesis that the slope is 1.0 against the alternative that it is not 1.0 and conclude a slope of 1.0 if this is not rejected. Assay linearity is separate from consideration of the shape of the concentration–response curve. Linearity of concentration–response is *not* a requirement of assay linearity. Linearity as discussed in ICH Guideline Q2B is concentration–response linearity

Quantitation limits: the limits of true relative potencies between which the assay has suitable precision and accuracy.

Note: This applies to assay results (log potency) rather than the reportable value.

Range: the interval between the upper and lower relative potencies for which the bioassay is demonstrated to have a suitable level of precision, accuracy, and assay linearity.

Note: This applies to assay results (log potency) rather than the reportable value.

Robustness: a measure of an analytical procedure's capacity to remain unaffected by small but deliberate variations in method parameters.

Notes: 1. Robustness is an indication of a bioassay's reliability during normal usage. For example, a cell culture assay system that is robust to the passage number of the cells would provide potency values with equivalent accuracy and precision across a consistent range of passage numbers. 2. ICH Q2B states:

> the evaluation of robustness should be considered during the development phase and depends on the type of procedure under study. It should show the reliability of an analysis with respect to deliberate variations in method parameters. If measurements are susceptible to variations in analytical conditions, the analytical conditions should be suitably controlled, or a precautionary statement should be included in the procedure. One consequence of the evaluation of robustness should be that a series of system suitability [qv] parameters is established to ensure that the validity of the analytical procedure is maintained whenever used.

Validation, assay: a formal, archived demonstration of the analytical capacity of an assay that provides justification for use of the assay for an intended purpose and a range of acceptable potency values.

Note: Formal validations are

conducted prospectively according to a written, approved plan.

Terms Related to Statistical Design and Analysis

Blocking: the grouping of related experimental units in experimental designs.

Notes: 1. Blocking is often used to reduce the variability of a measure of interest. 2. Blocks may consist of groups of animals (a cage, a litter, or a shipment), individual 96-well plates, sections of 96-well plates, or whole 96-well plates grouped by analyst, day, or batch of cells. 3. The goal is to isolate a systemic effect, such as cage, so that it does not obscure the effects of interest.

A ***complete block design*** occurs when all levels of a treatment factor (in a bioassay, the primary treatment factors are sample and concentration) can be applied to experimental units for that factor within a single block. Note that the two treatment factors, sample and concentration, may have different experimental units. For example, if the animals within a cage are all assigned the same concentration but are assigned unique samples, then the experimental unit for concentration is cage and the experimental unit for sample is animal; cage is a blocking factor for sample.

An ***incomplete block design*** occurs when the number of levels of a treatment factor exceeds the number of experimental units for that factor within the block.

Confounded design: two factors are confounded if their levels vary together (they are not crossed).

Notes: 1. For example, in a bioassay validation experiment in which one analyst performs assays on a set of samples for three days using cells from one passage number, then another analyst performs assays on the same set of samples for another three days using cells from a different passage number, the passage number of the cells and the analysts are confounded. [Also note that days are nested (qv) within analyst and cell passage number.] When factors are confounded one cannot tell which of the factors has caused an observed experimental difference. 2. Fractional factorial designs (qv), in which factors are only partially crossed, also are partially confounded. A full factorial design also can be confounded if the number of treatment combinations (sample and concentration) is greater than the block size.

Crossed (and partially crossed): two factors are crossed (or fully crossed) if each level of each factor appears with each level of the other factor. Two factors are partially crossed when they are not fully crossed but multiple levels of one factor appear with a common level of the other factor.

Notes: 1. For example, in a bioassay in which all samples appear at all dilutions, samples and dilutions are (fully) crossed. In a bioassay validation experiment in which two of four analysts each perform assays on the same set of samples on each of six days and a different pair of analysts is used on each day the analysts are partially crossed with days. 2. Each factor may be applied to different experimental units, and the factors may be both fully crossed and nested (qv), creating a split-unit or split-plot design (qv). 3. Experiments with factors that are partially crossed require particular care for proper analysis. 4. A randomized complete block design (RCBD) (qv) is a design in which the block factor (which often is treated as a random effect) is crossed with the treatment factor (which is usually treated as a fixed effect).

Experimental design: the structure of assigning treatments to experimental units.

Note: Blocking (qv), randomization (qv), replication (qv), and specific choice of design (to be covered in the planned General Chapter *Design of Biological Assays* <1032> are some aspects of experimental design. Important components of experimental design include the number of samples, the number of concentrations, and how samples and concentrations are assigned to experimental units and are grouped into blocks.

Experimental unit: the smallest unit to which a distinct level of a treatment is randomly allocated.

Notes: 1. Randomization of treatment factors to experimental units is essential in bioassays. 2. An experimental unit needs to be distinguished from a sampling unit, the smallest unit on which a distinct measurement is recorded (e.g., a well). Because the sampling unit is often smaller than the experimental unit, it is an easy mistake to treat sampling units as if they are experimental units; this mistake is called pseudoreplication (qv). 3. Different treatment factors can be applied to different experimental units. For example, samples may be assigned to rows on a 96-well plate while dilutions are assigned to columns on the plate. In this case, rows are the experimental units for samples, columns are the experimental units for concentrations, and wells are the experimental units for the interaction of sample and concentration.

Factor: an assay design element that may affect assay response and that varies in an experiment.

Note: In a bioassay there will be at least two treatment factors: sample and concentration.

A fixed factor is a factor that is deliberately set at specific levels in an experiment; inference is made only to the levels used in the experiment. In a bioassay, sample and concentration are both fixed factors.

A random factor is one for which its levels represent a sample of ways in which that factor might vary. In a bioassay, the test organisms, plate, and day often are considered random factors.

Factorial design: one in which there are multiple factors and the factors are partially or fully crossed.

In a *full factorial design,* each level of a factor appears with each combination of levels of all other factors. For example, if factors are sample (test and reference), concentration, and analyst, for a full factorial design each analyst must analyze all combinations of sample and concentration.

A *fractional factorial design* is one in which some factors are deliberately partially confounded with interactions associated with other combinations of factors.

Independence: For two measurements or observations *A* and *B* (raw data, assay sets, or relative potencies) to be independent, values for *A* must be unaffected by *B*'s responses and vice versa.

Note: In practice this means that if two potency or relative potency measurements share a common analyst, cell preparation, incubator, group of animals, or aliquot of reference sample,

then the assumption must be that they can not be assumed to be independent. The same holds true if the two potency or relative potency measurements are estimated together from the same model or are in any way associated. In some bioassays ongoing evidence from data can be used to show that it is reasonable to treat potency measures as independent even if they share a common level of a factor such as cell preparation.

INTERACTION: two factors are said to interact if the effect of one factor depends on the level of the other factor.

LEVEL: a location on the scale of measurement of a factor.

Notes: 1. Factors have two or more distinct levels. For example, if a bioassay contains two samples, test and reference, then there are two levels for the factor sample. 2. Levels of a factor in a bioassay may be quantitative, such as concentration, or categorical, such as sample (i.e., test and reference).

MODELING, STATISTICAL: the mathematical specification of the concentration–response relationship and important sources of variation in the bioassay.

Notes: 1. Modeling includes methods to capture the dependence of the response on the samples, concentration, and groups or blocking factors in the assay configuration. 2. Modeling of bioassay data includes making many choices, some of which are driven by data. With *continuous data* there is a choice between linear and nonlinear models. With *discrete data* there is a choice among logit/log models within a larger family of generalized linear models. In limiting dilution assays there is published literature advocating Poisson models and Markov chain binomial models. One can use either fixed-effects models or mixed-effects models for bioassay data. The fixed-effects models are more widely available in software and are somewhat less demanding for statisticians to set up.

On the other hand, mixed models have advantages over fixed ones. The former are more accommodating of missing data and, more importantly, can allow each block to have different slopes, asymptotes, median effective concentrations required to induce a 50% effect (EC50s), or relative potencies. Particularly when the analyst is using straight-line models fit to nonlinear responses or in assay systems in which the concentration–response curve varies from block to block, the mixed model captures the behavior of the assay system in a much more realistic and interpretable way.

3. It is essential that any modeling approach for bioassay data use all available data simultaneously to estimate the variation (or, in a mixed model, each of several sources of variation). It may be necessary to transform the observations before this modeling; to include a variance model; or to fit a "means" model (in which there is a predicted effect for each combination of sample and concentration) to get pooled estimate(s) of variation.

NESTED: a factor *A* is nested within another factor *B* if the levels of *A* are different for every level of *B*.

Notes: 1. For example, in a bioassay validation experiment two analysts may perform assays on the same set of samples on each of six days when no analyst performs the assay on more than one day (this requires 12 analysts who are qualified to perform the assay); these analysts are nested within days.

2. Nested factors have a hierarchical relationship. 3. For two factors to be nested they must satisfy the following: a) be applied to different-sized experiment units; b) the larger experimental unit contains more than one of the smaller experimental units; and c) the factor applied to the smaller experimental unit is not fully crossed with the factor applied to the larger experimental unit. When conditions (a) and (b) are satisfied and the factors are partially crossed, then the experiment is partially crossed and partially nested. Experiments with this structure require particular care for proper analysis.

PARALLELISM (OF CONCENTRATION–RESPONSE CURVES): the concentration–response curves of the test and standard are identical in shape and differ only in a constant horizontal difference.

Notes: 1. When test and reference preparations are similar (qv) and assay responses are plotted against log concentrations, the resulting curve for the test preparation will be the same as that for the standard but shifted horizontally by an amount that is the logarithm of the relative potency. Because of this relationship, similarity (qv) is generally referred to as *parallelism*. Note that similarity is the primary concept and that parallelism is not necessary for similarity; see slope ratio models in General Chapter *Design and Analysis of Biological Assays* <111>, in which samples with similar concentration–response relationships have a common (or nearly common) *y*-intercept, but may differ in their slopes.

2. In practice, it is not possible to demonstrate that the shapes of two curves are exactly the same. Instead, the two curves are shown to be sufficiently similar (equivalent) in shape. Note that *similar* should be interpreted as "we have evidence that the two values are close enough" rather than "we don't have evidence that the two values are different."

3. The assessment of parallelism depends on the type of function used to fit the response curve. Parallelism for a nonlinear assay using a four-parameter logistic fit means that: a) the slopes of the rapidly changing parts of the sample and reference standard curves (that is, slope at tangent to the curve, where the first derivative is at a maximum) should be similar; and b) the upper and lower asymptotes of the response curves (plateaus) should be similar. For straight-line analysis, the slopes of the lines should be similar.

RANDOMIZATION: a process of assignment of treatment to experimental units based on chance so that all equal-sized groups of units have an equal chance of receiving a given treatment.

Notes: 1. The chance mechanism may be an unbiased physical process (rolling unbiased dice, flipping coins, drawing from a well-mixed urn), random-number tables, or computer-generated randomized numbers. Care must be taken in the choice and use of method. Good practice is to use a validated computerized random-number generator.

2. The use of randomization results in systematic error becoming random error not associated with particular samples or a dilution pattern but distributed throughout the assay. In 96-well bioassays, plate effects can be substantial and cause bias or trending,

particularly in assays involving long-term cell culturing or multiple addition and wash steps. In animal studies, a variety of factors associated with individual animals can influence responses. If extraneous factors that influence either plate assays or animal assays are not routinely demonstrated to have been eliminated or minimized so as to be negligible, randomization that removes the influence of the biasing factor is essential to obtain unbiased data required for the calculation of true potency. Randomization is central to the experimental design and analysis of data obtained from most biological assays.

Replication: a process in which multiple independent experimental units receive the same level of a treatment factor.

Notes: 1. The purpose of replication is to minimize the effects of uncontrollable sources of random variability.

2. Replication can occur either completely at random or across blocks. Generally, replication within blocks is *pseudoreplication* (see below).

True Replicates: samples based on independent experimental units.

Pseudoreplication: the identification of samples from experimental units as independent and thus true replicates when they are actually not independent.

1. Pseudoreplication results in wrong inferences and the appearance of more replicates than are actually present.

2. Pseudoreplication is dangerous because it is an easy mistake to make, it is easy to overlook, and the consequences can be serious. For example, pseudoreplicates commonly arise when analysts are making a dilution series for each sample in tubes (the dilution series can be made with serial dilutions, by single-point dilutions, or with any convenient dilution scheme). The analyst then transfers each dilution of each sample to several wells on one or more assay plates. The wells are then pseudoreplicates because they are simply aliquots of a single dilution process.

3. In general, pseudoreplication should be avoided because, unless it is properly addressed in the analysis, it leads to underestimation of replicate variance.

4. The simple way to analyze data from pseudoreplicates is to average over the pseudoreplicates (if a transformation of the observed data is used, the transformation should be applied before averaging over pseudoreplicates) before fitting any sort of concentration–response model. In many assay systems averaging over pseudoreplicates will leave the assay without any replication. A more complex way to use data containing pseudoreplicates is to use a mixed model that treats the pseudoreplicates as a separate random effect. The only case in which pseudoreplication is useful is when the pseudoreplicate (i.e., well-to-well) variation is very large compared to the variation associated with replicates and the cost of pseudoreplicates is much lower than the cost of replicates.

References

1 ICH. Q2(R1) Harmonised Tripartite Guideline. *Validation of Analytical Procedures: Text and Methodology.* Current Step 4 version of parent guideline of 27 October 1994 (*Complementary Guideline on Methodology* dated 6 November 1996 incorporated in November 2005); accessed 29 August 2006 from www.ich.org/LOB/media/MEDIA417.pdf.

2 FDA. *Guidance for Industry: Analytical Procedures and Methods Validation —Chemistry, Manufacturing, and Controls Documentation*. August 2000; accessed 29 August 2006 from www.fda.gov/cder/guidance/2396dft.pdf.

3 International Organization for Standardizagion (ISO). International Standard 5725-1: *Accuracy (Trueness and Precision) of Measurement Methods and Results, Part 1 — General Principles and Definitions*. Geneva, Switzerland,1994; accessed 29 August 2006 from www.iso.org/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=11833&ICS1=17&ICS2=20&ICS3=&scopelist=.

Robert Singer *(PDL Biopharma),* ***David M. Lansky*** *(Lansky Consulting), and* ***Walter W. Hauck*** *(Thomas Jefferson University) all serve on USP Biologics and Biotechnology Expert Committees. Please address correspondence to* ***Tina Morris,*** *director of the Complex Actives division in the Department of Standards Development at US Pharmacopeia, 12601 Twinbrook Parkway, Rockville, MD 20852-1790; 1-301-816-8397; tsm@usp.org.*

This glossary is reprinted with permission from Pharmacopeial Forum *32(4)2006: 1359–1365, United States Pharmacopeia, 2006.*