# Approaches for in-vitro bioassay analysis and reporting of results

Prerna Maheshwari[1], Rajesh Singh[2]
[1]*School of Applied Sciences, Department of Biotechnology, Shri Venkateshwara University, Amroha, U.P, India*
[2]*School of Applied Sciences, Shri Venkateshwara University, Amroha, U.P, India*

**Abstract- For, most biopharmaceuticals, potency is assessed in a bioassay by comparing the dose -response curve of the test material and a reference standard. Bioassay systems are complex and tend to be sensitive to a greater variety of factors that are most physicochemical techniques. Some factors like pipetting controls, temperature, critical incubation time, serial dilutions errors can be controlled, but cell response, cell clumps cannot be controlled. Variation in these factors affects the response of a bioassay system to a test product, so its potency measurement is not an absolute value. Bioassays are therefore comparative, with the biological activity of a test material measured relative to that of a reference preparation. If the reference preparation is very similar to the test product, then their measured biological dose–response relationships should be affected equally by any variation in the system. Relative potency should, therefore, remain constant even though the measured response may vary among assays.**

**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 may consist of multiple blocks or complete replicates needed depends on the assay's inherent accuracy and precision and on the intended use of the reported value. It is practical to improve the precision of the reported value by reporting the geometric mean potency from multiple assays.**

**The purpose of this paper is to provide guidance on various approaches to bioassay data analysis and reporting of results.**

**Keywords – Bioassay, Relative potency, Reportable value, Reporting value, Assay acceptance criteria, Sample acceptance criteria**

## I. INTRODUCTION

This paper is restricted to discussion of bioassay data analysis applied to the responses of a bioassay system to reference standards, controls and test samples obtained during the performance of a potency assay, thereby reporting of the results.

Assay acceptance criteria are based primarily on a comparison of dose– response curves of control samples with a reference standard, all of which should be well characterized in the assay system. At least one control sample should be known to behave similarly to the reference standard in the assay system. Both this control and the reference standard should be known to behave similarly to the expected behavior of test samples.

The name assay control sample for the control material that behaves similarly to the reference standard and the expected behavior of the test samples. The origin of the assay control sample should be as independent of the reference standard and test samples as is possible within the constraint that all behave similarly in the assay system.

For plate-based biological potency assays, we propose the following two separate sets of acceptance criteria: assay acceptance criteria (AAC) and sample acceptance criteria (SAC).

A two-level, sequential assessment of acceptance criteria. First, AAC are assessed. Failure means that the entire plate is invalid. There is no processing of test sample data. Passing AAC allows processing of test sample data. Second, each test sample potency determination is then subjected to its own SAC. If it passes, then that test sample potency measurement is valid. If it fails, then that particular test sample potency quantification fails. Other test sample determinations on the same plate may still be valid.

The similarity of dose–response curves of reference standards and assay control samples is an essential AAC. The similarity of dose-response curves of the reference standard and test sample is an essential SAC.

AAC and SAC applied to an assay should be demonstrated to be useful in judging the validity of the assay. Both the criteria and the limits set on their values should be reviewed — and modified, if appropriate — as more assays are performed and further data are accumulated.

The rest of the paper is organized as follows. Proposed embedding and extraction algorithms are explained in section II. Experimental results are presented in section III. Concluding remarks are given in section IV.

## II. PROPOSED APPROACHES

*2.1 Data evaluation systemic approach:*

Below are the steps to guide for the analysis of bioassay data and reporting results:

- Exclude data which is considered to be an obvious error like technical problems such as contaminated wells, non-monotonic concentration- response curves.
- The obvious error in the analysis can be excluded based on known error. As a general, outlier testing has to be performed on transformed response data. Outlier testing/handling has to be defined in each assay protocol.
- Assess the system suitability on an un-restricted model that allows standard and controls to have their own curve fit.
- Define the system suitability criteria based on (assuming normal distributions for the assay acceptance criteria) +/- 3 SD around the means of the acceptance criteria from the historical data.
- If an assay (or a run) fails system suitability, the entire assay (run) shall be invalid and further data analysis shall not be computed.
- Assess sample suitability: if a test sample fails sample suitability, results for that sample are reported as " Fails Sample suitability " relative potency for another test sample shall be reported if other test sample passes sample suitability For example: In one assay plate if two test samples are tested labeled as S1 and S2, if S1 suitability fails, S2 suitability pass. S2 sample can be reported. Examples of sample suitability: relative confidence limits, slope ratio, EC50 criteria.
- If a sample passes sample suitability including similarity assessment, the sample value shall be considered as a reportable potency and used as a part of the reported value. The reportable potency shall be considered for constructing a reported value.
- If a sample fails sample suitability including similarity assessment, the sample potency shall not be reported for constructing a reported value and the sample shall be invalid for further data analysis like equivalence testing and analysis shall be repeated as shown in the flowchart
- If an assay passes sample suitability and fails in similarity, the determinant/run potency shall not be reported and to be considered as sample failure.
- If an assay passes sample suitability including similarity assessment and fails in relative potency criteria, the sample shall be regarded as Fail.

Table -1 Flow of data analysis and action taken.

| Assay system suitability | Sample suitability | Similarity assessment | Relative Potency | Results | Action |
|---|---|---|---|---|---|
| Pass | Pass | Pass | Pass | Reportable | NA |
| Pass | Pass | Pass | Fail | Reportable | Sample fail |
| Pass | Fail | NA | NA | Not | Repeat testing as per |
| Pass | Pass | Fail | Not reported | Reportable non similarity | Repeat testing as per flowchart |

An assay reportable value shall be comparted to product specification. The minimum number of determinants to compute reportable value shall be described. Computation of the number of determinants shall be determined based on the development data with a minimum of three runs and maximum depends upon the assay variability. If an assay reportable value falls outside the product specification. Such incidences shall trigger out of specification investigation.

Compute the % GCV on reportable value. The % GCV shall provide the precision on reportable value. For each method, depending upon the method validation/qualification, limits shall be defined or relative confidence limits can be defined on to the reportable value.

If an assay reportable value precision/relative confidence limits fall outside the criteria. The analysis cannot be considered for reporting. Sample suitability failures can be repeated as shown below in Figure.1 and 2.

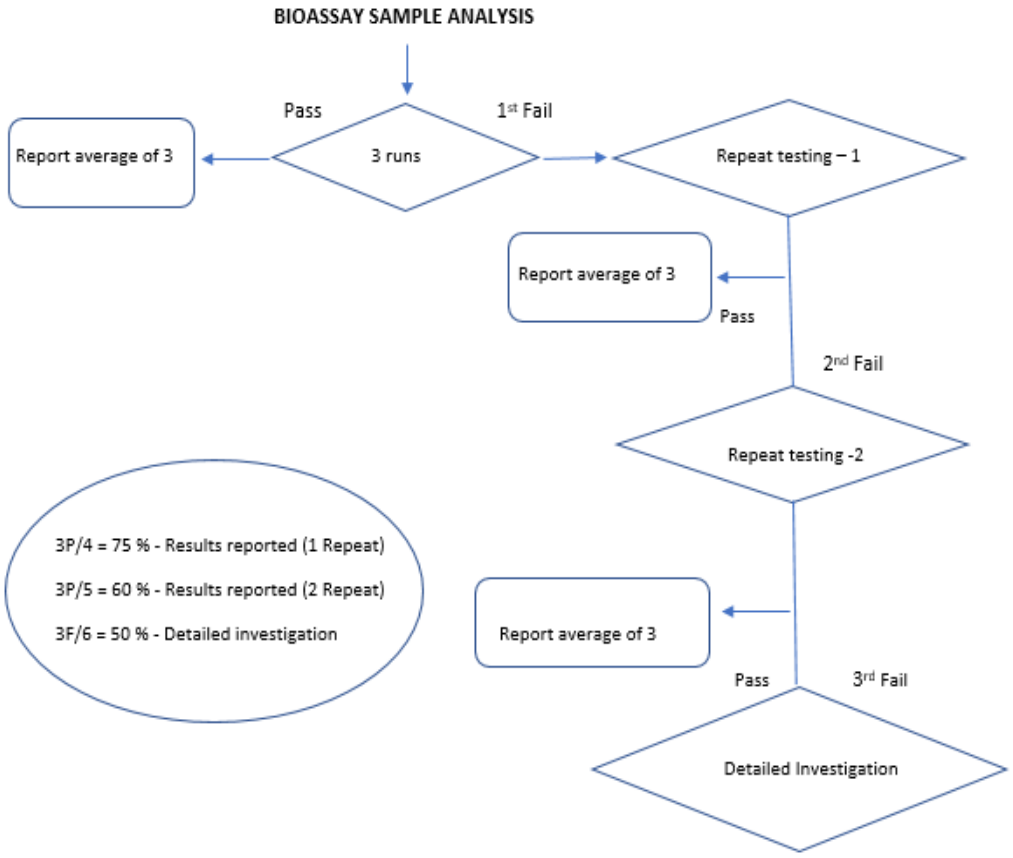


Figure 1. 3 runs analysis repeat approach

*2.2. Approach:2*

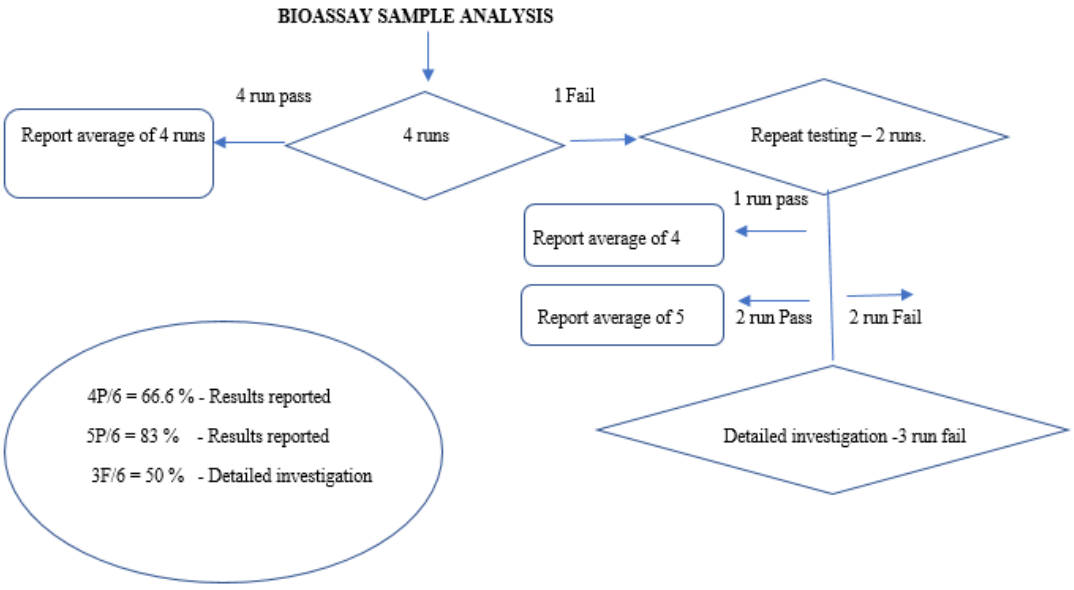


Figure 2 4 runs analysis repeat approach

2.3. Approach:3

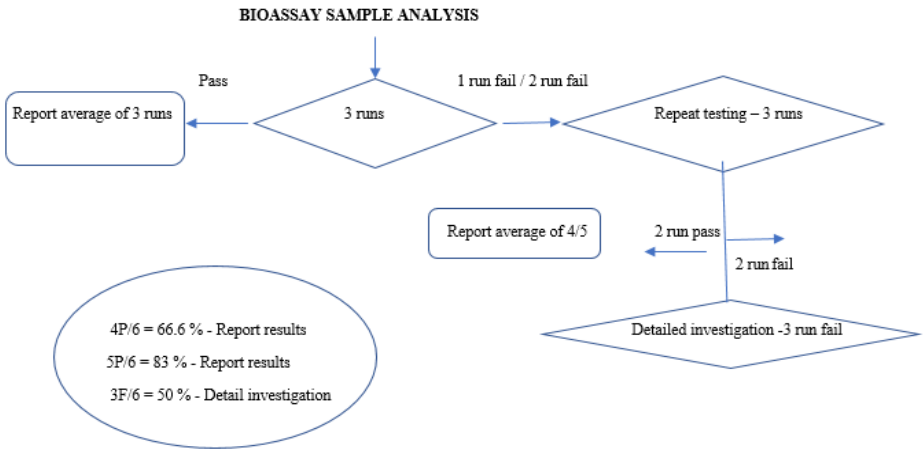


Figure 3. 3 runs analysis repeat approach

## III. DISCUSSION

Sample Acceptance: SAC are tests applied separately to each test sample. Their passing or failing is independent of other test samples. The similarity of the sample curve shape to that of the reference curve is a standard test. Other SAC (e.g., the variability of replicates) are usually similar to those for reference and assay control samples, but some may be different. For example, with some assay systems (particularly those that may not use the whole dose–response curve), the SAC may allow for fewer doses to be used in the curve fit than for the reference standard or assay control sample. Limits may be set for the ED50 of test samples to ensure that the dose–response curve falls within the validated dose range.

Reported Value: The reportable value shall be defined in the testing protocol. It is typically the mean of a defined minimum number of assay results from valid independent assays, with "independent" being defined in the protocol. Most protocols specify that failure of the assay to meet AAC or of a sample to meet SAC more than a specified maximum number or percentage of times shall trigger specified actions

The reported value is the fundamental unit of measure that is to be compared to the specification limit. It is a useful concept that captures all the details of the sampling design, the assay design and the estimation approach. It is the value that best represents the batch for the intended purpose of the measure.

Consider the difference between potency and content uniformity. Potency is a measure of the concentration or activity of the active ingredients in a given batch. Its focus is on the batch as a whole, not the individual units within the batch as such, potency should be represented by a reported value based on the average measure that represents the whole batch. This allows for aggregating across doses, units and replicates measure and, for the case of bioassay, even across replication of the entire bioassay. Content uniformity, on the other hand, is focused on measures from the unit of the individual dose, because its primary goal is to ensure that the variation in the potency from doses unit to doses unit is sufficiently small. Thus, for content uniformity, the reportable value is the value obtained from the individual doses unit.

Some have argued that every assay replicate is by definition are reported value regardless of the purpose of the measure and thus should be compared to product specification limits.

Essentially these approaches use specification limits to control the quality of all individual measures. This restriction is too limiting. Specification limit generally does not reflect the short term capability of the processes- they are a poor choice for monitoring quality. Secondly, this approach forces constraints upon how much data can be collected from any batch, since too much testing against limits by definition will lead to failures. The greater need is to be able to use the power of averaging to achieve the necessary level of precision for measure.

Assay Acceptance criteria: Reference standard and assay control sample dilution curves should be included on every plate to permit the AAC to be applied to each plate. In the past, some assay designs have included no assay control sample. In such cases, assay acceptance would be based solely on the reference standard dilution curve. That is a less rigorous test of assay validity, so this design should be avoided. In some designs, several plates are treated as

one unit, which passes or fails the AAC as a single unit. Usually, this is done to reduce the proportion of wells required for reference standard and assay control sample and to increase the number of wells available for test samples. Treating several plates as one unit is not recommended, however: Plate-to-plate variability can limit the use of this approach.
As shown above, there are various approaches to repeat sample suitability failures and construct a reported value based on the valid reportable value from individual determinants. The above approaches are widely used in various quality control units of biopharmaceutical industries to construct a reported value for the bioassay results.

*3.1 Common cause of failure:*
One of the major causes of frequent problems is to acceptance being made too dependent on an individual data point. A protocol illustrating this point involves standard, assay control sample and test samples tested in duplicate dose-response curves on a plate. Upper asymptotes are determined as the mean of duplicate responses at maximum dose with a maximum acceptable percentage difference between upper asymptotes on one plate. The reportable value is the mean of the potencies determined on two replicate plates, with a maximum acceptable percentage difference between the two potencies. One aberrant well at the maximum dose of the assay control sample on one plate results in the upper asymptote of the assay control sample being aberrant compared with the reference standard so the plate is rejected. No reportable value is obtained because of one aberrant well out of 96.
Another common cause of problems is the failure to ensure that each of the acceptance criteria is valuable in judging the validity of the assay as established during the assay characterization and continued monitoring. The inclusion of unnecessary criteria may result in the rejection of assays that are fit for purpose. For example, in some systems, ED50 may vary widely between assays which are fit for purpose, so setting tight limits on an absolute value for ED50 could result in pointless rejection of an assay.

*3.2 What is unique about bioassays?*
Bioassay differs from small molecule analytical methods in many details but as noted many times here a key difference is the high level of noise typically present. Regulatory agencies recognize these differences in part by allowing relative broader specification limits 80-125 % or broader are common. But even limits of this magnitude requires an ambitious level of precision for a bioassay. Imagine a bioassay with a precision of Sp of 10 %. Given this level can we define a reported value that supports 80-125 % specification? Imagine a situation where manufacturing and chemical processes are contributing little noise and further simplify the problem by assuming K and R are, thus fixed at 1. Thus we define reportable value based solely on n. Is there an n that achieve the objectives? Consider the problem with two perspectives:
How many assay would it take to ensure that LSL is greater than 80 %
How many assay would it take to have a meaningful release limits that lie outside the process control limits

Table 1 list shows LCL and LRL for different levels of N
Under the ideal condition where LCL=LSL, to achieve first goals require minimum 3 assays. To achieve the second require 6 assays
However, if target bioassay precision of Sp of 20 % the second goal cannot be achieved with up to 10 assays and the first goal is only achieved at n-10. These are expensive operating conditions. It may be that if an assay has a relatively large within the run precision, one may be able to achieve the run precision by increasing the number of replicates (r) but the assay load required to achieve the specification goal is substantial.

Table -2 lists show LCL and LRL for different levels of N

| N | Sp | Variance of the reportable value | LCL | LRL |
|---|---|---|---|---|
| 1 | 10.0 | 10 | 70 | 96 |
| 3 | | 5.7 | 82 | 90 |
| 6 | | 4.0 | 87 | 86 |
| 10 | | 3.2 | 90 | 85 |
| 1 | 20.0 | 20.0 | 40 | - |
| 3 | | 11.3 | 64 | - |
| 6 | | 8.2 | 75 | 84 |
| 10 | | 6.2 | 81 | 91 |

## IV. CONCLUSION

The appropriate design/assay configuration of potency assays is important if the results they provide are to be valid and useful. Attempts must be made to assess and reduce variability and possible bias.

Specifications for potency should only be derived following establishment of a well validated and controlled bioassay and require some measures of the variability of the estimate in addition to the actual potency estimate.

The geometric mean or weighted mean of three analysis runs are recommended to construct a concrete reported value to provide a level of precision to a reported value. This provides an assurance to regulatory bodies that the reported value is of good quality or, in other words, is sufficiently precise.

Obvious error in assay analysis is recommended to remove using statistical approaches.

AAC and SAC are the key component for any bioassay analysis and concluding the reportable results.

More than 50 % sample suitability failure triggers a detailed investigation of methods variability and controls shall be kept to control the variability of the assays.

A minimum of 66.6 % sample suitability passed results are recommended to ensure the level of precision and to determine the concrete reported value.

Based on the case study mentioned in the table.2, it is concluded as increasing the number of n decreases the variance of the reported value.

## V. REFERENCES

[1] Brown F,Mire-Sluis A (eds): The design and analysis of potency assays for biotechnology products DevBiol.Basel,Karger,2002,vol 106,pp 117-127

[2] Torbeck LD: Reportables value for out of specification test results,Pharmaceutical Technology February 1999.

[3] BurnJH :The errors of biological assays.Pysiological reviews 1:146-169

[4] Mire-Sluis AR: Progress in the use of biological assays during the development of biotechnology product pharm Res 2001

[5] USP guidance - <1034> Analysis of biological assays