

TDM Roundtable Recommendations: (Generic) Assay Validation Guidance

TDM Roundtable Participating Organizations:

American Association for Clinical Chemistry
American Medical Technologists
American Society for Clinical Laboratory Science
American Society for Clinical Pathology
College of American Pathologists
Food and Drug Administration

The concept of the Therapeutic Drug Management (TDM) Roundtable was developed by the AACCC TDM Renaissance Committee of the TDM-CT Division, in conjunction with the FDA, as means of fostering greater cooperation between the laboratory community and the agency to improve patient care. Special thanks to the TDM Roundtable workgroup that drafted the document.

*C. E. Pippenger, PhD, Co-chair
Dave Berry, PhD, Co-chair
Larry Broussard, PhD
Catherine Hammett-Stabler
Robert Murray, JD, PhD
Joseph Passarelli
Steven Wong, PhD*

TDM Roundtable Recommendations: (Generic) Assay Validation Guidance

Table of Contents

1. Introduction	3
2. Background	3
3. Risks to Health	3
4. Performance Characteristics	4
a. General Study Recommendations	5
b. Specific Performance Characteristics	5
(1) Precision	5
(2) Recovery	6
(3) Linearity	6
(4) Sensitivity	7
(5) Specificity for parent compound	8
(6) Interference	9
(7) Specimen collection and handling conditions	11
(8) Method comparison	11
(9) Studies at external sites	14
(10) Calibrators	14
5. References	14
6. Appendices	14

TDM Roundtable Recommendations: (Generic) Assay Validation Guidance

1. Introduction

This guidance document is intended to serve as a set of recommendations to researchers and manufacturers to facilitate the development and validation of therapeutic drug management (TDM) assays. Although the phrase “therapeutic drug monitoring (TDM)” has been used for many years and in many places to refer to quantitative measurement of therapeutic drugs in serum or plasma in order to assist a care provider to ensure that a patient is treated with optimal concentration of the drug in question, we have replaced “management” for “monitoring” in order to emphasize the purpose of the testing. “Management” implies that the laboratory measurement is an essential part of the treatment of the patient, whereas “monitoring” is focused on the analytical process, without reference to the clinical implications. The abbreviation “TDM” is retained throughout this document, but it is intended to refer to “management” and not to “monitoring.” As such it will establish scientifically sound expectations which are useful to document analytical performance of new testing devices or methods.

The remaining sections of this document describe the information generally needed in an FDA application for a TDM assay. Specific information related to some TDM assays is provided in the Annexes. Before undertaking development of any new TDM assay, the manufacture is strongly encouraged to contact the FDA, Office of In Vitro Diagnostics to discuss their validation strategy for FDA clearance or approval.

2. Background

Therapeutic Drug Management (TDM) assays are quantitative measures of a specific drug concentration in plasma or serum, and serve to aid in the management of a patient’s drug therapy. As analytical techniques, they are expected to accurately measure the concentration of the target drug, with defined precision, sensitivity, and specificity. While the typical specimen is plasma or serum, it is possible to clinically and analytically validate the assay to be used to test drug concentration in other biological samples e.g. whole blood, saliva, urine, or milk. Typically the metabolism and pharmacokinetics of the drug, which is the subject of the proposed assay will have been established and published in the scientific literature before a TDM assay is developed. The pharmacokinetics information, for the various matrices for which the test is intended and biological variations thereon, should be included in the information submitted to FDA, and in the information (package insert) provided to the user of the assay. Typically there is also information in the literature regarding optimal ranges. This information should also be presented.

3. Risks to Health

Frequently there are no known *direct* risks to patient health associated with the drug assay. Potential *indirect* risks exist, associated with the clinical consequences of an erroneous TDM result, which may

cause inappropriate patient dosing, leading to an inadequate or ineffective drug level in the patient, or to a toxic concentration.

Appendix I lists examples of toxicity associated with various drugs and drug classes for which there are approved (or cleared) TDM assays.

If a drug has been identified with specific risks, the corresponding recommended mitigation factor should be identified. The risks associated with antiepileptic drug assays are included as an example at Appendix II.

If there are other patient management risks, these should be addressed in the product labeling.

4. Performance Characteristics

a. General Study Recommendations

Patient samples or sample pools, derived from the intended use population (i.e., patients being treated with the drug in question) should be included in the analytical protocols described below. Spiked samples may be used under some circumstances, but at a minimum, samples from patients taking the target drug, must be included in the precision and recovery studies, as well as method comparison studies. This is important because patient samples reflect the relevant proportions of free and bound drug, metabolites, and other drugs commonly co-administered to the type of patients who require the target drug; therefore this is essential to demonstrate the robustness of the assay.

Spiked samples can be used to supplement the studies; however caution must be exercised against using spiked samples as the only matrix in the evaluations, because spiked samples, which may or may not contain metabolites of the target drug, provide a less complete assessment of the performance characteristics.

The effect of freezing/thawing samples, variables in collection and storages, should also be thoroughly investigated

All analytical protocols should be performed according to the procedures specified by the manufacturer in the testing program. The package insert will subsequently be developed from the standards and reflect a level of performance that can be achieved when the assay is performed according to the package insert. Therefore, each pre-analytical and analytical step must be specified and included in each of the analytical studies; preanalytical pretreatment steps, for example, should be included for individual replicates in a precision study and for individual dilutions in a linearity study. All of the manufacturer's recommended quality control and calibration procedures must be followed.

Appropriate specifics concerning protocols should be provided so that results can be interpreted properly and duplicated, if necessary. These specifics are also necessary to aid users in interpreting information in the labeling. For example, when referring to National Committee for Clinical Laboratory Standards (NCCLS) evaluation protocols or guidelines, indicate which specific aspect of the protocols or guidelines were followed.

In studies using spiked samples, information should be provided to document the purity of drugs, metabolites, or potential interferents, as well as the type of sample that the drug is spiked into.

Serum/plasma is the matrix recommended for most TDM assays, and equivalence must be demonstrated using the commonly employed anticoagulants and collection devices. In cases where whole blood or other biological matrices are to be analyzed, this should be clearly stated and appropriate correlations (comparison to serum or plasma assays, and comparisons among different anticoagulants as well as collection tubes) must be provided.

b. Specific Performance Characteristics

The following performance characteristics should be assessed in order to document performance and properly label the device in conformance with 21 CFR 809.10(b)(12).

(1) Precision

Within-run, and total precision should be characterized according to guidelines provided in “Evaluation of Precision Performance of Clinical Chemistry Devices; Approved Guideline” (1999) National Committee for Clinical Laboratory Standards (NCCLS), Document EP05-A². That document includes guidelines for experimental design, computations, and format for statement of claims.

Precision should be evaluated for at least three concentrations spanning most of the assay range. Typically these concentrations are chosen to represent (a) sub-optimal range or near low end of the reportable range (b) concentrations considered to be within the optimal range and (c) near high end of reportable range or toxic range.

Whenever possible, precision studies should be performed utilizing patient specimens. If patient specimens are not readily available at the time initial precision studies are performed, then as soon as possible during assay development, precision utilizing patient specimens should be evaluated to confirm that other compounds present in the patients’ biological fluids do not affect the TDM assay precision. When interpreting the significance of precision values it is important to recognize that the smallest coefficient of variation is the goal. However, it is equally important to recognize that clinical decision points associated with the interpretation of TDM values are generally reflected by a 20% change. Thus, a 10% CD for precision may be acceptable.

The description of the protocol and results should include the items listed below:

- sample types (e.g., pooled patient samples, spiked serum/plasma)
- point estimates of the concentration
- standard deviations of within-run and total precision
- sites at which precision protocol was run
- number of days, runs, and observations.

- calibration curve stability (if stored)

The factors that were held constant and which were varied during the evaluation (e.g., instrument calibration, reagent lots, and operators) should also be identified.

Computational methods, if they differ from those described in NCCLS EP05-A, should also be identified.

(2) Recovery

As a measure of accuracy, the percent recovery of the target drug should be characterized. Typically, these studies involve spiking known amounts of the pure drug into samples that are either negative for this drug or contain known drug concentrations. Spiking into samples from patients taking the target drug should be included as part of the study. Final concentrations of the spiked samples should span a significant part of the reportable range and include potential medical decision levels.

Recovery should be determined at both sub optimal and toxic concentrations to verify consistent performance across the assay range.

Replicates of each concentration or sample should be evaluated and the number of replicates chosen to ensure that any clinically significant differences observed will be statistically significant. Description of the study protocol should include:

- sample types and concentrations
- statement of how target concentrations were determined
- materials used for spiking
- number of replicates
- definition or method of calculating recovery.

When reporting results, the range of recoveries for each concentration evaluated should be indicated since this approach is more informative than describing mean recoveries at each concentration level.

(3) Linearity

The linear range of the assay response should be characterized by evaluating samples whose concentration levels are known relative to one another. When practical, the linearity of the assay should be characterized using dilutions of patient samples containing elevated lamotrigine concentrations. Spiked serum/plasma may be used when patient samples are not available (e.g., at very high drug concentrations). If patient specimens are diluted, they should be diluted with the same biological fluid to maintain the physiological dynamics of the system.

A graphic display or table of the known concentration vs. the observed concentration should be included. The sample concentrations should be evenly distributed across the reportable range of the assay. The appropriate number of replicates and concentration levels depends on the reportable range of the assay. Diluted patient sample pools are appropriate samples for the study. "Evaluation of the Linearity of Quantitative Analytical Methods; Approved Guideline" (2003) NCCLS Document EP06-A3 describes a protocol for sample preparation, value assignment, appropriate analyte range and concentrations to test, as well as statistical design and analysis methods, and a format for statement of claims.

Some immunoassays may exhibit a "high dose hook effect," in which there is a decrease in response of the assay at high concentrations. Whenever appropriate (e.g., for two-site or sandwich immunoassays), the linearity studies should be extended beyond the reportable range to the highest concentrations that may be encountered in clinical settings in order to evaluate whether the device exhibits a high dose hook effect.

The protocol description should include sample types and preparation, concentrations, number of replicates and statistical methods used. When practical, the linearity of the assay should be characterized using dilutions of patient samples containing an elevated drug concentration. Spiked serum/plasma may be used when patient samples are not available, (for example at very high drug concentrations). The description of results should include the acceptable maximal differences from linearity or the measured maximal differences (including confidence intervals) from linearity and the range of linearity, as described in NCCLS EP06-A. Data from the high-dose hook evaluation, should be included

Information on how to treat samples with concentrations outside the reportable range should be provided. If users are recommended to dilute samples that are above the reportable range, a specific protocol for dilution and a validation of that protocol should be provided. It is also necessary to clarify how samples with concentrations outside the range of linearity are reported to the user.

A validated protocol recommending how to dilute patient specimens without changing the assay's performance is an essential component of every TDM assay.

(4) Sensitivity

The functional sensitivity (lower limit of quantification) of the assay is defined as the lowest drug concentration for which acceptable assay precision and accuracy are observed, and this should be characterized and reported. This is generally considered to be the concentration at which the intra-assay coefficient of variation is not greater than 10%. The acceptance criteria for sensitivity of a TDM assay should take into account the expected serum/plasma concentrations at the lower limits of therapeutic dose and any possible patient non-compliance issues. The accuracy at the level of sensitivity (LLOQ) should also be described, based on samples with known drug concentrations.

The description of the sensitivity evaluation should include sample type, definition of the measures of sensitivity and results. Clarify how measurements below the level of sensitivity are reported to the user. The sensitivity and CV may vary depending upon the sensitivity of the analytical techniques utilized.

(5) Specificity for parent compound

As a measure of assay specificity, cross-reactivity with metabolites of the target drug should be characterised. Primary known metabolites should be included whenever possible. These may be obtained from the pharmaceutical manufacturer. Obtaining the parent compound is usually much less of a challenge than obtaining a full range of metabolites. For those drugs with many metabolites, the manufacture may have limited supplies of the major metabolites but not of metabolites present in low concentrations in biological fluids. Thus the pharmaceutical manufacturer may elect not to synthesize minor metabolites because they have no pharmacological activity. While developing a TDM assay, the developer is encouraged to establish a close working relationship with the pharmacology and drug metabolism division of the pharmaceutical manufacturer.

The description of the evaluation should include description of types of samples used for spiking, number of replicates, concentration of metabolite, computation or definition of cross-reactivity used and percent cross-reactivity for each metabolite.

With few exceptions, target drug metabolites are not routinely assayed in the same biological fluids, as routine TDM assays. Most but not all drug metabolites are present in concentrations significantly less than those of the parent compound at any given moment in plasma or serum. Identification of a drug metabolite cross reactivity during the development of a new TDM assay is challenging, but important. A simple approach to establish whether or not drug metabolite cross reactivity is present is to spike the parent compound into a drug free matrix and the same concentration into the appropriate biological fluid of a patient who has taken the drug. Another approach is to obtain assay results performed on patient specimens, particularly from patients with compromised renal function, and to compare such results with the results of a highly specific assay, such as mass spectrometry. It may be helpful to consult with FDA prior to undertaking this alternative type of study.

Whenever possible biological fluid specimens from patients in renal failure who are taking the drug should be retrieved and the reported concentration compared to that of specimens spiked to the same reported concentration in the same drug free biological fluid. If the observed results are significantly different between the two specimens, a metabolite cross reactivity problem should be suspected. The major metabolite of many drugs is a pharmacologically inactive glucuronide, which may cross react but under normal circumstances does not produce an elevated result with the TDM assay because plasma/serum concentrations are extremely low. However, patients in renal failure have extremely high glucuronide concentrations which may produce a cross reactivity with the TDM assay. This is the reason for quantitating biological fluids from patients in renal failure.

(6) Interference

The effects of potential interferents on assay performance should be characterised. Potential sources of interference that you should test include, but are not limited to, the following:

Endogenous compounds, particularly those listed below; at the suggested concentrations. The object of these studies is to confirm that elevated analyte concentrations of naturally occurring compounds that may occasionally be elevated do not interfere with the TDM assay.

- bilirubin (60 mg/dL)
- triglycerides (1500 mg/dL)
- cholesterol (500 mg/dL)
- uric acid (20 mg/dL)
- rheumatoid factor (500 IU/ml)
- hematocrit (15-60%)
- albumin (12 g/dL)
- gamma globulin (12 g/dL)
- human anti-mouse antibodies, HAMA
- hemoglobin (20-2000mg/dl, due to hemolysis)
- blood substitutes

Commonly co-administered drugs including, but not limited to those listed below. Drugs commonly co administered to treat a specific disease should also be evaluated for potential TDM assay interferences; the list of specific drugs to be checked is dependent upon the TDM assay under development.

- all available antiepileptic drugs together with relevant metabolites
- all available antipsychotic and antidepressant drugs
- Common tranquilizers and hypnotics

- commonly prescribed antibiotics
- common over-the-counter drugs

Anticoagulants or preservatives with which the sample is likely to come into contact, such as EDTA, fluoride/oxalate and heparin, various types of gels contained in serum separator blood collection tubes, and different collection and storage tube materials, such as plastic and glass. When testing these interferents, the concentrations of the target drug in the sample should be adjusted to medical decision levels. Typically, interference studies involve adding potential interferents to the sample containing the drug and determining any bias in the drug recovery, relative to a control sample (to which no interferents have been added). In addition to anticoagulants and endogenous substances it is essential that the various specialized biological fluid collection devices (e.g., serum gel separators, filter paper, and ultra filtrations membranes) also be evaluated.

Recommended guidelines for interference testing are described in detail in “Interference Testing in Clinical Chemistry; Approved Guideline” (2002) NCCLS Document EP07-A4. This document includes guidelines for setting decision criteria as well as for protocol designs, statistical methods, evaluating interference using patient specimens and establishing validating and verifying interference claims. The following considerations should included when interferent testing is being planned:

- For endogenous substances, test at the highest concentration expected based on experience with the intended population. Interference studies using samples naturally high in the endogenous compound being tested can be informative and this approach should be considered when such samples are available.
- For drug levels, test to levels 3 times the highest acute peak concentration reported following therapeutic dosage.
- For specimen additives, test up to levels five times the recommended concentration.

If interference is observed at the concentration levels tested, lower levels should be tested in order to determine the lowest concentration that could cause interference. Replicate samples should be tested in these protocols. In addition to anticoagulants and endogenous substances it is essential that the various specialized biological fluid collection devices e.g. serum gel separators, filter paper, and ultra filtration membranes also be evaluated.

The description of the evaluation should include the following items (if description of the protocol refers to NCCLS EP07-A, clarify which aspects of the guidelines were followed):

- names and concentrations of interferents tested

- sample type (e.g., spiked whole blood pools, samples naturally high in endogenous compounds)
- concentrations of target drug in the sample
- number of replicates tested
- definition or method of computing interference.

When reporting results, any observed trends in bias (i.e., negative or positive) across the concentration range of interferents tested must be identified. Include the standard error of the observed recoveries at each concentration or the range of observed recoveries at each concentration evaluated for a potential interferent. This approach is more informative than listing average recoveries alone.

For substances listed as non-interfering, state the criteria on which this is based, e.g., inaccuracies due to these substances are less than 10% at a given concentrations. If any compounds are known from the literature or other sources to interfere with the test system, these should be included among the information in the labeling. It may not be necessary to perform additional interference testing with these known interferents.

(7) Specimen collection and handling conditions

The labeled recommendations for specimen storage and transport must be substantiated, by assessing whether the device can maintain acceptable performance (e.g., precision and accuracy) over the storage times and temperatures (including freeze/thaw cycles), and across various anticoagulants (and, if applicable, the use of gel-containing serum separator tubes) recommended as acceptable by the manufacturer. An appropriate study includes analysis of sample aliquots stored under the conditions of time, temperature, or allowed number of freeze/thaw cycles recommended in the package insert. Storage conditions and freeze- thaw cycles are especially important for research studies where long storage specimen storage periods are required. Manufacturers' should update the package insert as new information on storage criteria becomes available.

(8) Method comparison

The new assay must be compared with a reference method, specific for the parent compound. Carefully validated chromatographic methods that specifically measure parent drug should be used as the comparator in such a study. If the discordance exceeds 10% relative to the reference procedure, the reasons for the discordance should be addressed. The steps to be taken to minimize the risk of patient mismanagement when based on the results of such tests must be described.

For any TDM assay it is essential that a comparison be made with the accepted published reference method. At the present time, ideally the technique should be compared to a GC/MS or LC/MS technique. Often, simpler LC, GC, immunoassay or other techniques are published. It behoves the TDM assay developer to compare the TDM assay to any analytical technique that may be routinely utilized in clinical

chemistry laboratories for drug analysis. Such initial comparisons allow the manufacturer to establish the performance of the TDM assay under various analytical conditions.

Guidelines provided in the document, “Method Comparison and Bias Estimation Using Patient Samples; Approved Guideline” (1995) NCCLA Document EP09-A5 concerning experimental guidelines and statement of claims should be followed. Patient samples with drug concentrations distributed across the reportable range of the assay, when used in applications for which the drug is approved, should be evaluated. Banked (retrospective) samples are appropriate for these studies as long as the information listed below concerning sample characterization is available. Samples from multiple geographic sites or clinical centers should be included.

Appropriate sample size depends on factors such as precision, interference, range, and other performance characteristics of the test. The number of patients should also be large enough so that inter-individual variation would be observed. A statistical justification to support the study sample size should be provided in the protocol description. It is expected that the sample size target, however supported, will include a minimum of 100 samples distributed fairly evenly over a minimum of 50 *individual patients*.

If multiple measurements from individual patients are included, the results should be summarized using appropriate statistical analyses such as Analysis of Variance, Generalized Estimating Equations, or Bootstrapping, to account for correlation of repeat measurements within patients in the study.

For the results to be properly interpreted all relevant information on the sample population should be provided in the package insert. Information on the sample population should include:

- the number of individual patients represented by the samples
- the number of data points
- the number of clinical sites
- information regarding the time of last dose

Any specific selection (inclusion or exclusion) criteria for samples should be stated together with an indication of whether samples were collected from patients with specific clinical outcomes, or from centers using atypical or novel drug regimens. Factors such as age range (e.g., adults), and time of blood draw with respect to drug administration (e.g., trough, peak) might influence drug-to-metabolite ratios and consequently, assay bias (although significant concentrations of the major metabolites would not be expected to accumulate sufficiently in most patients to cause bias). In general, pre-dose blood is the preferred sample for TDM, but for the purpose of a method comparison, any time of sampling would be acceptable.

Ideally one would like sufficient clinical information to be able to calculate Concentration Dose Response Ratios. Often, however, it is not practical to obtain dosing information and sample draw times for stored laboratory specimens. Storage conditions can affect the quantitation of specimens particularly if they have been stored for an extended period of time. If there is wide variance between the TDM assay and the reference method in stored specimens, it is suggested that the specimens be re-assayed utilizing the reference technique before comparing results with the new TDM assay. Such analysis compensates for storage changes that alter drug concentrations.

The comparator method used must be clarified, and references to validation of the procedure included. If samples evaluated in the study include both trough and other times of blood draw relative to drug administration, a separate statistical analysis for these groups should be conducted as well. When providing the results of the method comparison study, the following information should be included:

Scatterplots of the new assay versus the reference method. The plots should contain all data points, the estimated regression line and the line of identity. Data points in the plot should represent individual measurements.

A description of the method used to fit the regression line and results of regression analysis including the slope and intercept with their 95% confidence limits, the standard error of the estimate (calculated in the y direction), and correlation coefficient should be included. In cases where parameters are not consistent throughout the reportable range, estimates of more than a single range may be appropriate. If the comparator, as well as the new assay is subject to measurement error, a regression method such as the Deming method may be appropriate, rather than Least Squares.

- To illustrate the degree of inter-individual variations, include graphs of difference in measurements (i.e., new device minus reference method) versus the reference method. Appropriate representations include a bias plot of difference in measurements ($y - x$) versus the reference method (x), as recommended in NCCLS Document EP09-A, or versus the mean of y and x , as recommended by Bland and Altman (Bland, 1995).

The points above apply to any reference method. The more information that is available comparing the reference method to the new TDM assay, the easier it is for the reviewer to recognize the validity of the new assay. Providing the information initially in sufficient detail and clarity speeds the review process and we emphasize the importance of clear and frequent communication with the FDA Diagnostics Division during the development of any new TDM assay.

A variety of clinical circumstances can influence the interpretation of any drug concentration. The purpose of a TDM assay is to provide a tool that can be utilized in conjunction with other clinical parameters and diagnostic procedures to enhance any clinician's ability to provide optimal patient care through the use of Therapeutic Drug Management at any time. More efficient and better patient care will result from more readily available TDM assay.

(9) Studies at external sites

Performance at external laboratory sites in addition to that of the manufacturer's site should be evaluated. The assay should be evaluated in at least three external sites. This may be included as part of the method comparison study described above. Data from individual sites should initially be analyzed separately to evaluate any inter-site variation. Method comparison results from the individual sites can be pooled in the package insert, if it is demonstrated that there are no significant differences in results among sites.

(10) Calibrators

Provide the following information about the calibrators in the assay kit in the summary report:

- Protocol and acceptance criteria for real-time or accelerated stability studies for opened and unopened calibrators.
- Protocol and acceptance criteria for value assignment and validation, including any specific instrument applications or statistical analyses used.
- Identification of traceability to a domestic or international standard reference material.
- Protocol and acceptance criteria for the transfer of performance of a primary calibrator to a secondary calibrator.

For information about calibrators marketed separately as class II devices under 862.1150, see the guidance "Abbreviated 510k Submissions for *In Vitro* Diagnostic Calibrators."

5. References

6. Appendices