FDA Public Workshop on Liquid Chromatography-Mass Spectrometry (LC-MS) in the Clinic

> *The information and questions contained in this document are not binding and do not create new requirements or expectations for affected parties, nor is this document meant to convey FDA's recommended approaches or guidance. Rather, the information contained in this document offers background and the basis for discussions at the Public Workshop.*

The goal of this public workshop is to facilitate discussion with the clinical LC-MS community around specific analytical questions pertaining to the validation of LC-MS-based *in vitro* diagnostic products (IVDs) that identify and/or quantify specific proteins and peptides. This paper is designed to guide the discussion at the upcoming FDA Public Workshop on May 2, 2016 and is intended to enhance engagement with academia, government, industry, clinical laboratories, and other stakeholders regarding the questions presented below.

A. Background

Innovations in LC-MS have dramatically improved the specificity and scope of clinical diagnostic tests. The FDA has cleared and approved several MS-based devices; these include IVDs for screening newborns for metabolic diseases, for identifying microbes from human cultures, and for measuring the concentrations of therapeutic drugs in blood. As of this workshop, no LC-MS–based IVD that measures proteins and peptides has been cleared or approved by the FDA.

The purpose of this discussion paper is to enhance engagement on questions related to the analytical validation expectations for LC-MS-based devices and to solicit feedback from the community as to how these elements of validation are applied routinely in the clinical lab. The discussion is not intended to address questions regarding devices for which special controls already exist[1].

Clinical validation and reference range studies are out of scope for this discussion paper. While these studies are critical to demonstrating the safe and effective use of many IVDs, the design and execution of these studies are not different for devices that use LC-MS technology than they are for other devices. In addition, discussion of laboratory developed test (LDT) policy is outside the scope of the scientific discussions under consideration in the workshop.

B. General Principles of Analytical Validation

FDA regulates medical devices, including IVDs, with the objective of assuring that the devices are reasonably safe and effective for their intended uses. The intended use and indications for use of an IVD are critical to appropriately validate test design and performance to support marketing claims. For instance, analytical (and clinical) validation studies in complete submissions that FDA has received have been performed in the intended population and setting to ensure that results from these studies, and the subsequent marketing claims made based on these study results, appropriately reflect performance of the IVD in that population and setting. In the context of LC-MS-based IVDs, the indications for use typically have included the following critical elements: Information in an IU includes (a) a description of which analytes are measured or detected by the IVD; (b) the specimen or sample type that is needed to run the IVD; (c) the underlying-technology of the IVD ; (d) the type of measurement performed (i.e., qualitative, quantitative, semi-quantitative); (e) the intended population with which the test is designed to be used; (d) the clinical purpose for test result, including the disease/condition for which the device is offered (e.g., to aid in the diagnosis of [insert disease/condition]); (f) the setting in which the test is to be performed; and, (g) the setting in which the test result is most likely to be used.

[1] http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm077311.htm

Analytical validation studies evaluate a key subset of the performance metrics of an IVD, including precision and reproducibility, linearity, analytical measuring range, stability and traceability of calibrators and control materials, stability of reagents, detection limit/analytical sensitivity, analytical specificity, carryover and recovery, and matrix comparisons. There are considerations within these categories that are unique to LC-MS devices, as discussed in further detail below.

C. Key Components of Analytical Validation and Associated Discussion Points

Here we present FDA's experiences with and considerations for analytical studies for protein- and peptide-based LC-MS IVDs and points of discussion for which we are seeking input. We strongly encourage device manufacturers to contact FDA through the Pre-Submission[2] process to discuss study designs *prior* to starting validation studies. The guidelines and consensus standards referred to below have been recognized by the FDA as of the date of the workshop.

Precision and Reproducibility

Precision and reproducibility are determined by repeated measurement of samples that cover the claimed measuring range of the device. Useful general advice for designing and analyzing precision studies for IVDs can be found in the CLSI standard EP05-A3 (*Evaluation of Precision of Quantitative Measurement Procedures; Approved Guideline—Third Edition*). For LC-MS IVDs that include complex workflows, it is important to consider all steps in the workflow, from sample procurement through data output and analysis, and with acceptance criteria developed with all of the specific details of the assay in mind.

Results obtained from measuring several analytes in one MS injection can be utilized as a multiplexed assay or by combining them into one actionable result. In a multiplexed assay, the analytes are all detected in one test but each is evaluated and reported independently of the others. Alternatively, the levels of several different analytes, in this case peptides from one or many proteins in a patient sample, can be combined as individual elements of an algorithm to provide a final, actionable result, such as a score that predicts risk of disease. In these assays, FDA has historically evaluated the analytical validation of the individual components independently as well as with respect to their contribution to the score. For example, FDA generally has asked manufacturers to provide evidence of the contribution of each analyte to the final score alone and in various combinations of concentrations of each component.

Linearity and Recovery

General guidelines for designing and analyzing linearity studies for IVDs can be found in CLSI standard EP6-A (*Evaluation of the Linearity of Quantitative Measurement Procedures: A Statistical Approach; Approved Guideline*). For LC-MS IVDs that detect multiple analytes, either independently (as in a multiplexed panel) or combined (as in an IVDMIA), manufacturers have evaluated the linearity and recovery of each analyte independently from the other analytes.

Carryover

The potential for carryover—material from a previous sample that contaminates a new sample— varies depending on the workflow of an assay, and it is important for manufacturers to identify and evaluate all steps within the workflow from which carryover may occur. In the past, FDA has accepted carryover studies that alternate high-positive samples with negative or low-positive samples in a series that encompasses the entire workflow of the device.

Detection Limit/Analytical Sensitivity: Limits of Blank, Detection, and Quantitation

LC-MS IVDs rely on the detection and quantitation of a peak either directly in the mass spectrometer or as an ion current measured over the timescale of an LC chromatogram. Either measurement depends on the detection of the peak above the chemical and electronic noise present in the system. It may not be

[2] http://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm311176.pdf

necessary or practical to determine the limit of blank (LoB) for some LC-MS assays due to the low background noise inherent in LC-MS technology and the lack of reliable signal in blank samples; however, if a clinical decision point is near the limits of sensitivity of the instrument, LoB may be important to evaluate. For LC-MS IVDs, the Limit of Detection (LoD) is calculated using a lower threshold for the signal-to-noise (S/N) ratio. An appropriate S/N ratio may vary by device and in the past complete submissions have provided a sufficient rationale to support the use of a particular S/N ratio in the determination of LoD. CLSI standard EP17-A2 (*Protocols for Determining the Limits of Detection and Limits of Quantitation, Approved Guideline—Second Edition*) provides useful advice for the determination Limit of Quantitation (LoQ, also known as Lower Limit of the Measuring Interval (LLMI)) for an LC-MS IVD.

Analytical Specificity and Interferences

For LC-MS IVDs, there are two types of interferents that may affect the detection and quantitation of the test analyte. First, isobaric or near-isobaric ions that are present in the sample and detected simultaneously in the MS can interfere with detection and quantitation of the correct ion. It is important to evaluate the presence or absence of these interferents and to assess whether the presence of these interferents compromises the safety and effectiveness of the device, including taking into consideration the mass resolution of the MS instrument under standard operating conditions.

Second, endogenous and exogenous interferents in a sample may affect the detection and quantitation of the analyte. Guidelines for designing and analyzing interference studies for IVDs can be found in CLSI standard EP07-A2 (*Interference Testing in Clinical Chemistry; Approved Guideline—Second Edition*). In addition to endogenous materials such as hemoglobin and exogenous materials such as drugs a patient may be taking, chemical components in blood collection tubes or artifacts from elution of analytes from dried blood spots may produce peaks that interfere with target analyte peaks during analysis and interpretation.

Analytical specificity and interference testing may depend on the workflow that is used for the LC-MS IVD. The complexity of the sample that ultimately reaches the MS can vary broadly across device designs and this can significantly affect the number and type of potential interferents. The specifics of this situation would be ideal for discussion with the review division in a Pre-Submission.

***Discussion Point*:**

- *When traditional study designs are impractical given the nature of the protein/peptide LC-MS assay workflow, what alternatives have proven successful to demonstrate robust performance of the assay? How do you approach validating multiplex versus single-plex assays?*

Sample preparation

Sample preparation for protein/peptide LC-MS devices can be very complex and is known to be a key contributor to the imprecision and irreproducibility of the assay. The ability to control these variables across different operators, different instruments, or different sites may raise questions about safety and effectiveness of the IVD. In the past, FDA has asked IVD manufacturers to identify and evaluate all the source of variability in their devices and provide data that demonstrate that this variability does not affect the safety and effectiveness of the device.

Discussion Points:

- *How have you evaluated, validated, and controlled sample collection and workflow procedures for protein/peptide-based LC-MS?*

Internal Standards

For an LC-MS device, especially one that requires substantial sample preparation, internal standards are included to identify and quantify the analyte and to monitor different steps in the workflow. In general, internal standards are added as early in the analytical procedure as possible, e.g., to correct for losses during sample preparation. Standards are also used to normalize analyte values across samples, calibrators, and controls. While the type of standards used is up to the discretion of the test developer, the sufficiency of the standards to detect and quantitate the proper ion and/or control the steps in the workflow is an important consideration. In addition, the chemical and isotopic purity of the standards has a significant impact on the performance of the device, so it important to establish acceptance criteria for these attributes of the internal standard prior to analytical validation.

Calibrators and Control Material

Calibrators serve as a reference to convert the raw output of the device into a concentration of analyte that is used to inform a clinical decision. Guidelines are available containing general information on measuring stability (CLSI standard EP25-A, *Evaluation of Stability of In Vitro Diagnostic Reagents; Approved Guideline*) and metrological traceability (CLSI standard EP32-R, *Metrological Traceability and Its Implementation; A Report*) of calibrators and controls.

It is important to consider that for LC-MS IVDs, in addition to a known concentration of analyte, control materials may also contain a fixed amount of internal standard to provide relative quantitation across the calibrators and patient samples. The use of internal standards for quantitation is discussed above.

It is also important to note that in this context *control material* refers to material that is used to verify total system performance, i.e., positive and negative controls. This should not be confused with in-process controls or internal standards described above that may be used in an LC-MS IVD workflow.

Discussion Point:

- *Given the variety of proteins measured and the work-flows involved, please discuss appropriate calibrators and control materials for quantitative protein/peptide LC-MS/MS devices, and what commonalities can be applied to all protein/peptide LC-MS IVDs. Please discuss when external vs. internal calibration is preferred.*

Comparator methods

FDA typically considers "gold standards" to be international reference materials, e.g., WHO and NIBSC standard preparations of specific analytes. Standard reference methods refer to the best available method for establishing the presence or absence of the target condition.

Although a test with a different technology has in many instances been used as a predicate when it has the same intended use, including detecting the same analyte, because the LC-MS device output is different from that of an immunoassay, a straightforward comparison of the two devices can be difficult. To resolve this difficulty and when a gold standard or reference method has not been available, FDA has sometimes accepted that device performance may be demonstrated by a comparison of the device performance to the clinical "truth" (i.e., clinician's full evaluation of a patient's status, based on current standard of care).

Discussion point:

- *What performance is required and how should a new protein/peptide LC-MS device be compared with another if the analyte composition is subtly different, e.g., LC-detectable isoforms are the analyte?*

Selection of peptides

Most peptides are selected as surrogates for intact proteins, with the exception of some small bioactive peptides that can be easily detected without proteolysis. Manufacturers may select whatever peptides they believe are most suitable for their test. The biophysical properties of the peptide may affect its performance. These include its susceptibility to post-translational modification either in vivo or during the workflow, the rate and efficiency of trypsinization, and the peptide's stability in the sample and throughout the workflow.

Peak Selection

Selection and evaluation of peaks is critical to the performance of LC-MS devices. Factors that contribute to the resolution and reproducibility of the peak eluting from the LC include choice of LC column, buffers, flow-rate, and temperature. Depending on the workflow being used, FDA has historically asked test developers to provide data demonstrating that resolution and reproducibility of LC peaks is maintained under the operating conditions that may be encountered in the clinical laboratory. As discussed above, internal standards can be used to confirm the identity of LC peaks and ions in the mass spectrometer. In the absence of an internal standard, FDA has historically asked manufacturers to provide data that confirms that the correct peaks and ions are being detected and quantified.

The software used to select and evaluate the correct LC peak is of critical importance. It is important that manufacturers understand and evaluate the accuracy and reproducibility of the peak-picking software, particularly in samples with levels of analyte near the lower limit of quantitation.

Critical IVD Components

A number of consumable reagents and components are critical to the performance of LC-MS IVDs, e.g., LC columns, source and type of trypsin, immunoaffinity antibodies, immunodepletion columns. Based upon FDA's previous experience, a change in a critical component may warrant validation of the change to the one component of the assay or re-validation of the entire assay.

Discussion Point*:*

- *Discuss the methods that may be used to validate a change to one component of the assay, e.g., LC-column or source of antibody, and if/how these are validated with or without re-validating the entire assay.*

Relevant Data

For many LC-MS devices, the relevant line data for demonstrating performance would be the calculated AUC and/or the ratio of the AUC to the internal standard. FDA has historically requested representative visualizations of peaks and the calculated S/N at key decision points.

Discussion Point:

- *Please discuss how raw data and processed data are managed and evaluated for the quality of the results and validation of the assay. Include a discussion of the number of peaks/chromatograms reviewed, and how the data are managed. How should in silico analyses be presented?*

D. Summary of Topics and Questions for Discussion

In view of the multiple challenges regarding design and validation of LC-MS/MS devices, we hope to have panel discussion on the concerns raised above, and in particular on the following concerns and questions specific to LC/M/MS devices for proteins and peptides, and LC-MS devices generally:

1. *Basic Principles*
   When traditional study designs are impractical given the nature of the protein/peptide LC-MS assay workflow, what alternatives have proven successful in resulting in equivalently robust assurance regarding the performance of the assay? How do you approach validating multi-plex versus single-plex assays?

2. *Sample Collection, storage and preparation:*
   How have you evaluated, validated and controlled sample collection and handling procedures for use with protein/peptide-based LC-MS?

3. *Standards and Internal controls*
   Given the variety of proteins measured and the work-flows involved, please discuss appropriate calibrators and control materials for quantitative protein/peptide LC-MS/MS devices, and what commonalities can be applied to all protein/peptide LC-MS IVDs. Please discuss when external vs. internal calibration is preferred.

4. *Comparator methods*
   What performance is required and how should a new protein/peptide LC-MS device be compared with a predicate if the analyte composition is subtly different, e.g., LC-detectable isoforms are the analyte?

5. *Analytical specificity, harmonization and critical components*
   Discuss the methods that are used to validate a change to one component of the assay, e.g., LC-column or source of antibody, and if/how these are validated with or without re-validating the entire assay.

6. *Relevant data*
   Please discuss how raw data and processed data are managed and evaluated for the quality of the results and validation of the assay. Include a discussion of the number of peaks/chromatograms reviewed, and how the data are managed. How should in silico analyses be presented?