Mass spectrometry—based in vitro diagnostic devices that measure proteins and peptides are underutilized in clinical practice, and none has been cleared or approved by the Food and Drug Administration (FDA) for marketing or for use in clinical trials. One way to increase their utilization is through enhanced interactions between the FDA and the clinical mass spectrometry community to improve the validation and regulatory review of these devices. As a reference point from which to develop these interactions, this article surveys the FDA’s regulation of mass spectrometry—based devices, explains how the FDA uses guidance documents and standards in the review process, and describes the FDA’s previous outreach to stakeholders. Here we also discuss how further communication and collaboration with the clinical mass spectrometry communities can identify opportunities for the FDA to provide help in the development of mass spectrometry—based devices and enhance their entry into the clinic.

© 2015 American Association for Clinical Chemistry

Technological innovations in mass spectrometry (MS) have revolutionized analysis of the human proteome and dramatically improved clinical assay throughput and precision. The US Food and Drug Administration (FDA) has cleared and approved MS-based devices as diagnostic tests for screening newborns for metabolic problems, identifying microbes from human cultures, and measuring the concentrations of therapeutic drugs in blood. However, no HPLC-MS–based device for measuring proteins and peptides in in vitro diagnostic devices (IVDs) has yet been cleared or approved by the FDA for any use, including in clinical trials. Many factors contribute to this current absence (1–3), among which may be a lack of experience of many test developers with the FDA’s regulatory processes or concerns about FDA expectations for device performance and validation. What follows is an overview of the FDA’s experience regulating MS-based devices, outreach history, and where we would like to go with the MS community to promote entry of these devices into the clinic.

FDA Experience with MS Devices

OVERVIEW OF FDA REGULATION OF IVDs

IVDs are tests that measure disease or the condition of a patient, and are considered regulated medical devices under the Food Drug and Cosmetic Act § 321(h)(2) (4) and 21 CFR § 809.3 (5). IVDs include tests run by a clinical testing laboratory, performed by a physician’s office or laboratory, or purchased by consumers over the counter. IVDs are reviewed primarily by the Office of In Vitro Diagnostics and Radiological Health (OIR) at the Center for Devices and Radiological Health (CDRH) at the FDA. IVDs include devices used to screen for, diagnose, or monitor a disease, those that predict treatment response or prognosticate outcomes, and those used in clinical trials.

The regulatory overview of IVDs is entirely dependent on the risk of the test to the population in which it is meant to be used, formally called its Intended Use (IU) or Indications for Use (IFU). Risk in the context of most IVDs means the risk to a patient stemming from actions taken based on a false-positive or a false-negative result (e.g., unnecessary surgery, treatment delay). The FDA classifies IVDs into 1 of 3 risk-based classes: classes I, II, and III, with Class I devices posing the lowest risk (e.g., control materials), Class II devices posing moderate risk (e.g., some disease-monitoring devices or tests that help diagnose a condition), and Class III devices posing the...
highest risk (e.g., many cancer screening devices). The scope of data required to demonstrate the probability or severity of a risk therefore is dependent on the device classification. A Class III device is approved through the premarket approval (PMA) pathway, with substantial clinical, analytical, and manufacturing review before it can be legally marketed. A Class II device is cleared through the 510(k) pathway, which is based on demonstrating that performance of a new device is substantially equivalent to an FDA-cleared, marketed device with the same IU/IFU, (i.e., a predicate device) (6). Details of CDRH’s regulatory processes have been discussed at length (7–9), and information describing the definition of IU/IFU, the required analytical and clinical data, the device labeling, and how to perform benefit/risk analysis is publicly available on CDRH’s website (6, 10, 11).

REGULATORY STATUS OF CLINICAL HPLC AND MS INSTRUMENTS

Currently, HPLCs and mass spectrometers designed for clinical use are Class I exempt devices under the respective regulations 21 CFR § 862.2260 and 21 CFR § 862.2860 (12–14). This means that the manufacturer does not have to obtain FDA approval or clearance prior to marketing, although instruments must be manufactured under quality systems regulations [Quality System Regulation, 21 CFR Part 820 (15)] and labeled for IVD use [Section 809.10(b)(16)]. The MS and the manufacturer need to be registered and listed. It is important to note, however, that the Class I status applies only to the instrument itself—it does not apply to an assay that is performed using the instrument unless the assay itself also is Class I exempt. When an instrument is used to measure a specific analyte, the instrument plus the associated reagents are classified as a test system and the system as a whole is reviewed on the basis of the risk of the assay. For example, a Class III HPLC-MS cancer screening IVD will require that the HPLC-MS instrumentation also be reviewed as a Class III device and its use in the assay is subject to PMA. Moreover, for both Class II and Class III devices, subsequent modifications to the assay, instrumentation, or software require additional submissions for FDA review.

OIR-CLEARED MS ASSAYS

Several types of IVDs that use HPLC-MS or MS alone as an analytical platform have been cleared by OIR under different device-specific regulations (Table 1). For example, an HPLC-MS/MS device was approved in 2008 with the IU of detecting inborn errors of metabolism as part of newborn screening programs [21 CFR 862.1055, Newborn screening test system for amino acids, free carnitine, and acylcarnitines using MS/MS (17)]. Metabolites extracted from dried blood spots are analyzed using an HPLC coupled to an MS/MS system. In 2006, a quantitative HPLC-MS/MS version of a previously cleared device for monitoring therapeutic concentrations of tacrolimus, an immunosuppressive drug [21 CFR 862.1678 Tacrolimus Test System (18)] was cleared. In this case, device performance was compared to 2 immunoassays for the same analyte. A device utilizing MALDI-TOF MS was approved by OIR in 2012 for the identification of microorganisms cultured from human samples. With the use of this device, the identities of microorganisms in test samples are determined by comparison of the spectrum of the unknown to a database of reference spectra of known microorganisms.

Each device described above uses MS for detection, but the technological differences between HPLC-MS/MS and MALDI-TOF MS are significant. These differences, and the specifics of each assay and system (e.g., sample type, sample preparation, analyte) determined the analytical study designs and the variables that were evaluated during review (see Table 1). These review processes are relevant to future devices with the same IU; devices that measure novel analytes, or use an algorithm to combine results from a panel of analytes to produce a single result, may require additional validation compared with devices that measure previously cleared analytes.

It is critical to appreciate that the FDA’s definition of an assay or device includes not just the HPLC and the MS elements but the entire device work flow, from sample collection and preparation to generation and reporting of the assay result, including software, algorithms, and databases. For example, the effect of culture medium and microorganism culture age on spectrum matching were key variables evaluated during review of microbial identification devices. All steps of a new assay need to be evaluated for their contribution to the risk of generating incorrect results.

Guidances and Standards

FDA GUIDANCE DOCUMENTS

FDA guidance documents are one of the methods used to inform device developers and manufacturers about regulatory policies, and these devices reflect the FDA’s thinking on a particular device at the time of publication. They are intended to identify the most critical challenges in
gaining clearance or approval of a general category of devices and define solutions to those challenges. Guidance documents are written and finalized in response to comments received from the public. Guidance topics range from the very general [e.g., how to submit a 510(k) (19)] to the very specific [e.g., special controls that lay out the requirement for using MS/MS in newborn screening (21)].

An FDA guidance focused on validation of HPLC-MS bioanalytical methods [e.g., as tools for drug development (22)] is available, but at present no OIR guidance document exists to address the validation of HPLC-MS–based IVDs for measuring proteins and peptides. The scope of the studies required to validate IVDs compared to those required to validate analytical methods can be significant; therefore, we believe device developers and manufacturers would benefit from more detailed information about regulatory requirements for using MS/MS in newborn screening (21).

An FDA guidance focused on validation of HPLC-MS bioanalytical methods [e.g., as tools for drug development (22)] is available, but at present no OIR guidance document exists to address the validation of HPLC-MS–based IVDs for measuring proteins and peptides. The scope of the studies required to validate IVDs compared to those required to validate analytical methods can be significant; therefore, we believe device developers and manufacturers would benefit from more detailed information about regulatory requirements for using MS/MS in newborn screening (21).

Table 1. Comparison of regulatory criteria of 3 cleared, Class II MS-based assays reviewed.a

<table>
<thead>
<tr>
<th>Scope of assay reviewed</th>
<th>Sample preparation to output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regulation</td>
<td>21 CFR § 862.1678</td>
</tr>
<tr>
<td>Procedure</td>
<td>Whole blood, protein precipitated, supernatant injected, solid-phase cleanup, calibration curve of spiked, freeze-dried whole blood</td>
</tr>
<tr>
<td>Analytical factors</td>
<td>Comparison of traces of different transitions in different clinical samples, evaluation of metabolites, stability of signal intensity</td>
</tr>
<tr>
<td>Clinical study</td>
<td>&gt;50 samples from each demographic group, multiple sites</td>
</tr>
<tr>
<td>Software/instrumentation</td>
<td>Detailed review of documentation for device software, including, standalone software applications and hardware-based devices that incorporate software</td>
</tr>
<tr>
<td>HPLC-MS/MS tacrolimus assay</td>
<td>HPLC-MS/MS for newborn screening</td>
</tr>
<tr>
<td>MALDI-TOF for microbial identification</td>
<td></td>
</tr>
</tbody>
</table>

*The assays are quite different technologically, but the requirements for validation are similar in scope, and are based on the specifics of the assays (21 CFR § 862.1678 (18), 21 CFR § 862.1055 (17), 21 CFR § 866.3361 (35)).

**Outreach to the Clinical MS Community**

The FDA engages with device developers and manufacturers in many ways, including direct communication through regulatory submissions, participation in seminars at national and international scientific meetings, hosting scientists for seminars at the FDA, organization of public meetings and workshops, and publications in scientific journals. In recent years, 2 key public work-
shops were held that focused on proteomic devices: the Interagency Oncology Task Force Molecular Diagnostics workshop in 2008 and the CDRH Proteomics in the Clinic public workshop in 2014. An overview of these workshops is presented below.

INTERAGENCY ONCOLOGY TASK FORCE MOLECULAR DIAGNOSTICS WORKSHOP

On October 30, 2008, the National Cancer Institute and the FDA co-organized a workshop to discuss the regulatory requirements for analytical validation of proteomic devices, i.e., multianalyte HPLC-MS-based devices and immunoaffinity array devices. A detailed summary of the outcomes from this meeting has been published (24).

An outcome of this workshop was the preparation of 2 “mock 510(k) presubmission” documents (25). These documents presented putative validation studies for a multiplex immunoaffinity MS platform for protein quantification [Supplement 1 of (25)] and an immunological array platform quantifying glycoprotein isoforms [Supplement 2 of (25)]. OIR scientific, clinical, and statistical staff reviewed and commented on the proposed design, function, and performance testing of the devices. Although no data were submitted with these mock 510(k) presubmission documents, the collaboration between OIR and the proteomics community helped educate both the workshop participants and OIR about the regulatory challenges facing device developers and manufacturers.

PROTEOMICS IN THE CLINIC PUBLIC WORKSHOP

On June 13, 2014, CDRH held a public workshop to solicit insights from a broad range of clinical MS stakeholders. OIR invited participants from academia, industry, not-for-profit interest groups, large testing labs, the public, and government agencies, including the NIH and the NIST. The agenda, complete transcript and recorded webcast are publicly available (26).

A major purpose of the meeting was to encourage discussion of critical challenges in analytical validation of HPLC-MS devices faced by clinical laboratories. Speakers pointed out that the scope of work involved in analytical validation can be daunting, for example, if validation requires sample numbers in excess of those recommended by CLSI to evaluate performance with all demographic groups with which the device is intended to be used. Even so, appropriate validation is fundamental to ensuring device performance will be clinically acceptable when offered for patient care.

The FDA has discussed common missteps in device validation made by novice and experienced device developers. From the point of view of validating performance, there is nothing unique about HPLC-MS–based devices: the device must be demonstrated to be safe and effective, e.g., have characteristics such as acceptable precision, lack of interferences, and adequate sensitivity. On the other hand, HPLC-MS–based devices may present novel technological challenges not seen with other devices, e.g., selection of controls and calibrators and effects of sample preparation differences on reproducibility.

The process that led to the approval and clearance of the MALDI-TOF devices for microbial identification was presented by OIR [Table 1 and (26)]. These approvals were the direct result of intensive communication and preparation by both the sponsors and OIR. The sponsors made use of the presubmission process (27) that CDRH has implemented to encourage early discussions between the manufacturers and the FDA about their technology and validation approaches. As the result of a series of presubmissions, both the device manufacturers and OIR had a clearer understanding of performance testing that was necessary and how OIR would evaluate the results. The result of the interactions was that the review process was already established when the devices were submitted.

The following additional major workshop topics were included:

(a) Encouraging transparency around design and results of analytical and clinical validation. As noted above, without access to validation study designs, statistical analysis plans, and the computer code and algorithms used to calculate the results, a device user cannot judge if the device is reliable enough for use in patient management decisions. Although this information is publicly available for devices that are cleared or approved by FDA [PMA Summary of Safety and Effectiveness (28) and 510(k) databases (29)], there is no such transparency for in-house assays unless test developers voluntarily choose to publish the results. It was suggested in the meeting that this information might be included as relevant in peer-reviewed publications, for example, as supplementary material.

(b) Ensuring that correct samples are used for device development and validation. Sample collection study design remains a topic of considerable discussion. The probability of a device succeeding in validation is lower if it is developed on a population that differs from the intended use population; the probability of a device succeeding in validation is higher if sample collection strategies are developed to test a particular hypothesis or to support a specific IU.

(c) Addressing the need for targeted advice from the FDA. OIR participates in considerable outreach efforts to the IVD community, defining regulatory terms and explaining the general IVD review principles that ensure that devices are safe and effective for their IUs. This general advice does not address particular concerns of HPLC-MS device developers and manufacturers, which has led to uncertainty regarding how OIR’s review policies might be applied to novel technologies and analytes. For example, the regulatory review of Class II devices is
based on the concept that their safety and effectiveness can be adequately demonstrated by comparison to a predicate device. Many developers are unsure of the regulatory pathway if a predicate does not exist for a novel device with a unique analyte or analytes. This apparent conundrum is a fairly routine concept in OIR, which has substantial experience in developing de novo pathways to validate novel technologies that do not unduly burden the developer. This process was used to approve and clear the MALDI-TOF MS devices for identification of microorganisms, for which there were no predicates. Although device developers are strongly encouraged to directly consult with OIR through the presubmission process, both OIR and the clinical HPLC-MS community would benefit from continued and widespread dialogue and dissemination of OIR review policies.

**Next Steps**

OIR plans to build on our previous outreach efforts to address specific challenges that hamper entry of MS-based protein and peptide IVDs into the clinic. OIR will seek input from the clinical MS community on considerations related to clearance or approval of MS-based IVDs. Some of these issues may include the following: (a) Identification of HPLC- and MS-specific regulatory issues. Two CLSI documents focus on clinical MS: C50-A, *Mass Spectrometry in the Clinical Laboratory: General Principles and Guidance* (30) and C62-A, *Liquid Chromatography-Mass Spectrometry Methods* (31). Both documents seek to educate MS test developers, healthcare providers, instrument manufacturers, and regulatory agencies about clinical MS, and provide practical advice on the development and validation of clinical MS-based tests. They primarily address devices that measure drugs or metabolites rather than proteins and peptides. Both discuss different types of clinical HPLC-MS systems, including HPLCs, autosamplers, ionization sources, and mass analyzers, as well as instrument performance parameters that must be considered when developing a test. They provide information on the general testing required to analytically validate the performance of any device, (e.g., precision, linearity, analytical measuring range, analytical sensitivity, recovery) and reference other CLSI consensus documents on these topics. At the time of writing, CDRH has not recognized these 2 documents.

These documents may provide a useful baseline from which to begin the discussion of analytical challenges for validating MS-based peptide IVDs that use selected or multiple reaction monitoring (32–34). As discussed in the existing documents, considerable attention should be paid to instrument and reagent performance characteristics such as mass accuracy, mass resolution, MS tuning, HPLC peak shape and retention time, and choice of HPLC stationary and mobile phases. Also of importance are the factors that contribute to reproducible and accurate quantification of proteins and peptides, such as MS signal to noise, choice of ion transitions, and the effect of dwell time, points across the peak, and ion suppression on quantitation. Additionally, HPLC-MS-specific issues concerning sourcing and testing of reagents and the robustness of software used for data acquisition, data analysis, and results reporting merit attention. The CLSI documents do not address the specific challenges associated with detection and/or quantification of intact proteins or larger peptides not amenable to selected or multiple reaction monitoring, and it will be important to consider areas in which the development of MS-based peptide and protein IVDs that use different technologies differ.

Many of these topics have been discussed at length in the literature, but the development of common standards and methods has been difficult for the research and clinical communities. It is our hope that continued discussion of these topics in the context of IVDs could assist device developers and manufacturers to bring these IVDs to the clinic.

(b) QC materials and standards. Use of QC materials and calibrators should be addressed to ensure that tests are accurate and reproducible across clinical laboratories. This includes the use of isotope dilution methods and isotopically labeled internal controls for both quantification and normalization of peptide IVDs. These internal standards, which are unique to MS-based IVDs, provide opportunities to develop tests with excellent analytical parameters, but present a challenge to OIR and the device developers and manufacturers to define how and when these controls should be used. NIST has been very active in developing standard materials that may be useful for HPLC-MS IVDs, and the further development and proper inclusion of these in the HPLC-MS IVD workflow should be discussed.

(c) Devices with multiple analytes. HPLC-MS devices can be used to quantify many different proteins or peptides in a single analysis of 1 sample. Although OIR has cleared and approved a number of devices that use panels of analytes, the potential scale of multiplexing using MS can far exceed that of other devices. How OIR will apply its experience with multiplex validation requirements to this level of complexity for MS-based devices is a topic of considerable deliberation in the OIR.

(d) Comparator used to evaluate a test’s performance. Demonstration of IVD performance is based on the use of predicates, reference and control materials, and clinical diagnostic gold standards; however, these often are not available for the analytes and conditions that MS-IVDs are designed to test. As discussed above, the lack of predicates is a challenge that OIR faces regularly with novel devices, and we do not view this as a serious impediment.
to clearance or approval of MS-based IVDs. The lack of recognized international reference materials and control materials and the potential lack of clinical diagnostic gold standards, e.g., for devices that predict patients’ response to drugs, are complicated issues for all types of technologies, including MS, which will benefit from further discussion.

(e) Prespecification of acceptance criteria. A critical tenet of device validation is that a device developer should understand their test’s performance in the context of the IU population and must prespecify acceptance criteria for each of the various parameters (e.g., precision, linearity) prior to validation. This prespecification requirement ensures that the device meets those parameters in samples different from those on which the test was developed. It has often been requested that OIR set official acceptance criteria, e.g., that the CV for overall precision may not exceed 10%. This approach is not feasible because a single acceptance criterion is not appropriate for all devices, as performance is inherently related to the analyte(s) and the assay technology. For example, a fully automated immunoassay may be expected to have a CV of <4%, whereas some manual ELISAs may have a CV of 15% because in this case, <10% is too liberal for the automated ELISA and too restrictive for the manual ELISA. Acceptance criteria are dependent on the intended use, i.e., a standalone diagnostic vs a device that aids in the diagnosis of a disease or condition. We do not know what level of precision will be appropriate for various MS assays, some of which may include no or substantial sample pretreatment.

OIR also does not set official acceptance criteria because the biological variability of different diseases or IU populations can dictate the limits of acceptance criteria, such as the prevalence of coexisting liver disease in a test for hepatocellular carcinoma or age-based cutoffs. Device developers are generally in the best position to understand the nature of the device technology and the disease, and therefore OIR relies on them to propose and adhere to acceptance criteria in the context of their device.

OIR is keen to explore additional outstanding concerns with the clinical HPLC-MS community, and the steps that we can take to provide useful, relevant advice on how to bring these IVDs to the clinic, where they can fulfill their potential to help in patient management.

Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

Authors’ Disclosures or Potential Conflicts of Interest: Upon manuscript submission, all authors completed the author disclosure form. Disclosures and/or potential conflicts of interest:

Employment or Leadership: J.T. Lathrop, FDA; D.A. Jeffery, FDA; Y.R. Shea, FDA; P.F. Scholl, FDA.
Consultant or Advisory Role: None declared.
Stock Ownership: None declared.
Honoraria: None declared.
Research Funding: None declared.
Expert Testimony: None declared.
Patents: None declared.
Role of Sponsor: No sponsor was declared.

Acknowledgments: The authors gratefully acknowledge the expert assistance of Michail Alterman, Michael Boyne, Sahar Dawisha, Jinong Li, David Litwack in planning the 2014 public workshop and the generous advice and thoughtful participation of the workshop speakers: Andrew Hooftnagle, Lisa McShane, Russell P. Grant, Emily Boja, Salvatore Sechi, Karen Phinney, Leigh Anderson, Abraham Tzou, and Marina Kondratovich.

References


