Published in final edited form as:

Microchem J. 2012 November 1; 105: 32–38. doi:10.1016/j.microc.2012.02.011.

Mass Spectrometry-Based Multiplexing for the Analysis of Biomarkers in Drug Development and Clinical Diagnostics- How Much is too Much?

Uwe Christians, Jacek Klepacki, Touraj Shokati, Jost Klawitter, and Jelena Klawitter
iC42 Clinical Research & Development, Department of Anesthesiology, University of Colorado, Aurora, Colorado 80045, United States

Abstract

Biomarkers, or more specifically molecular markers, can detect biochemical changes associated with disease processes and drug effects before histopathological and pathophysiological changes occur. Multiplexing technologies such as high-performance liquid chromatography/mass spectrometry (LC-MS) allow for the measurement of molecular marker patterns that confer significantly more information than the measurement of a single parameter alone. The use of multiplexing assays for drug development, and as diagnostic tools, is attractive but will require regulatory review and approval and thus requires validation following regulatory guidances. Multiplexing assays always constitute a compromise. The number of analytes that can reasonably be included in a mass spectrometry-based multiplexing assay depend on the physico-chemical properties of the analytes and their integration into a single assay in terms of extraction, HPLC separation, ionization conditions and mass spectrometry detection. Another aspect includes biomedical considerations such as the differences in physiological concentrations of analytes, the required concentration range, and how much variability is acceptable before the clinical utility of a marker is negatively affected. Regulatory considerations include validation and quality control during sample analysis. Current bioanalytical regulatory guidelines have mostly been developed for single drug compounds and are not always adequate for multiplexing molecular marker assays that often quantify endogenous compounds. Specific guidances for multiplexing assays should be developed. Even if it is possible to integrate a wide variety and large number of analytes into a multiplexing assay, it should always be taken into consideration that a set of shorter, more specialized assays, may offer a more manageable and efficient alternative.

Keywords

liquid chromatography- mass spectrometry; multiplexing; biomarkers; metabolomics; regulatory aspects; validation

© 2012 Elsevier B.V. All rights reserved.

Address for Correspondence: Uwe Christians, iC42 Integrated Solutions in Systems Biology for Clinical Research and Development, University of Colorado, Bioscience East, Suite 100, 1999 North Fitzsimons Parkway, Aurora, Colorado 80045-7503, United States, Phone: +1 303 724 5665, Fax: +1 303 724 5662, uwe.christians@ucdenver.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Molecular Markers in Clinical Diagnostics and Drug Development

Cells either directly or indirectly (*via* extracellular fluid) communicate with body fluids. Cell metabolites, peptides and proteins are released from cells or are taken up from body fluids *via* normal excretion, trans-membrane diffusion or transport, and throughout the death process during which cells release all of their contents. Thus at least to a certain extent, biochemical and protein changes in cells and organs are reflected in body fluids. While tissue samples, biopsies, and certain fluids such as urine (kidney), bile (liver), and cerebrospinal fluid (CNS) mainly reflect changes in specific organs and thus are considered “proximal matrices”; plasma samples reflect systemic changes that often cannot be traced back to a certain organ. [1] Such changes of metabolites, peptides and proteins in body fluids, if mechanistically linked to disease processes and drug effects in tissues and organs, have the potential to serve as surrogate markers or biomarkers. A biomarker is defined as “a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes or pharmacological responses to therapeutic intervention”. [2] Based on this definition biomarkers have been used all along in clinical diagnostics ranging from the measurement of clinical symptoms such as blood pressure, ECG, over imaging technologies to modern high throughput gene arrays. Since here we discuss technologies that directly or indirectly assess molecular mechanisms, the more focused term “molecular marker” will be used instead of “biomarker”. A molecular marker can consist of the measurement of single molecular entity but it can also be a set of several molecular entities, a molecular pattern or fingerprint.

Today’s clinical diagnostics is typically based on a limited set of biomarkers, often only one parameter that is associated with the functional aspect of an organ or a specific disease process (for example creatinine in serum as a marker of kidney function). However, there is no single molecular entity marker, and there will never be one that captures the function of organs such as the kidney, liver or vascular endothelium in all its complexity.

Modern analytical technologies allow for the identification of patterns that confer significantly more information than the measurement of a single parameter, just like a bar code contains more information than a single number. Well-qualified molecular marker patterns yield more detailed and mechanistically relevant information translating into good specificity. The better the specificity of a molecular marker pattern, the more this will reduce non-specific background noise and reduced background noise, usually, results in better sensitivity. Thus it is reasonable to expect that diagnostic strategies based on molecular panels will provide the key to significant improvements such as enabling individualized and promoting preventive medicine.

Multiplexing analytical technologies, including but not limited to, arrays, bead immunoassays and mass spectrometry, allow for the assessment of molecular marker panels ideally in a single run. Among these technologies mass spectrometry is attractive due to its sensitivity, specificity and flexibility, and it allows for absolute quantification. Whereas antibodies are derived from biological sources, mass spectrometry-based multiplexing assays lack the manufacturing and batch-to-batch reproducibility challenges of many antibody-based assays. While in antibody-based assays the type and number of compounds detected is always limited by the antibodies included in the assay, mass spectrometry provides the basis for non-targeted approaches and therefore the means for detection of an open number of previously undefined signals. The goal of a non-targeted assay is to capture as much information as possible. [3] Non-targeted assays constitute the extreme of multiplexing assays and in an ideal world would capture the complete human metabolome or proteome. However, current technologies allow for capturing only a fraction of the metabolites and proteins in the body. [4] The major difference between a non-targeted and

targeted assay is that, as aforementioned, in a truly non-targeted analysis, the compounds underlying the signals are not known. Thus, such assays are semi-quantitative at best and can only be validated to a limited extent. In contrast to non-targeted assays, targeted multiplexing assays simultaneously measure several well-defined compounds, are validated and are quantitative. Targeted multiplexing assays avoid some of the analytical and statistical uncertainties associated with completely non-targeted data sets in terms of the quality of data, quantification, interferences as well as false negatives and positives.

2. Regulatory Aspects

Multiplexing assay strategies have been and can be used to identify unknown molecular mechanisms, can be used for drug discovery, pre-clinical and clinical drug development and as clinical diagnostic tools. This makes molecular marker strategies and associated multiplexing assays interesting and powerful tools for drug discovery and development (Figure 1). The utilization of molecular marker strategies to improve drug development efficiency and safety has been encouraged by drug regulatory agencies such as the United States Food and Drug Administration (FDA). [5]

Regulatory agencies have established review structures as well as guidances that outline the biomarker qualification process [6–10] (Figure 1). Based on these guidances, the biomarker development process can be divided into three stages: Discovery, Verification and Qualification. Verification mechanistically links the molecular marker to the biochemical process underlying a disease or drug effect. The qualification process bridges the results of molecular marker measurements, symptomatic drug effects and disease outcomes. Qualification has been defined as “a graded, fit-for purpose evidentiary process linking a biomarker with biology and clinical endpoints.” [11] Qualification has been differentiated from validation, which focuses on the reliability and performance characteristics of the analytical assay used to measure molecular markers. [6,12,13] Translation of a molecular marker from the discovery stage into pre-clinical testing and clinical development greatly depends on the availability of robust, precise and sensitive assays for the measurement of a larger number of samples. [14] While during the discovery phase a partial validation following “fit-for-purpose” principles will be sufficient, the validation must become more stringent if drug development strategies, regulatory approval and clinical decisions will depend on such a molecular marker (*vide infra*). During later clinical development and especially when developed as a clinical diagnostic tool, which typically involves comparison of individual results with normal values, absolute quantification of the analytes and complete validation following applicable regulatory guidances is critical. In the United States, the use of molecular marker data in regulatory review and decisions is currently based on the FDA guidance “Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products.” [15] If appropriately qualified and based on adequately validated assays, molecular markers can support primary outcomes; they may help to understand and monitor mechanisms of toxicity, drug interactions, disease-drug interactions and the effects of genotypes, gender and age (Figure 1). They can be used to stratify patient populations, guide subgroup analyses to bridge safety and efficacy data between different populations such as adults to pediatric patients, and among different ethnic groups.

Molecular markers may also be developed into clinical diagnostic tests. In many countries, for such an *in vitro* diagnostic device to enter the market, it must comply with a set of rules and regulations [16], such as 510(k) premarketing clearance or premarket approval (PMA) oversight by the FDA in the United States (CFR807, 814) [17,18], the Pharmaceutical Affairs Law (PAL) and Market Authorization Holder (MAH) oversight by the Pharmaceutical and Medical Devices Agency (PMDA) in Japan [19], and *in vitro* diagnostic device directive 98/97/EC oversight by regulatory authorities of the member states of the

European Union [20]. Again, the submission of adequate validation data for the analytical assays is critical for regulatory approval.

3. Analytical Challenges of Mass Spectrometry-Based Multiplexing Assays

Traditionally in clinical diagnostics a specific assay has been used to measure a specific analyte. Multiplexing assays open up new opportunities by measuring combinatorial markers allowing for pattern analysis, but are also associated with greater challenges as they require the integration of analytes with often different physico-chemical properties into a single assay. This potentially affects all aspects of an assay from sample handling to mass spectrometry-based quantification. The major steps of most mass spectrometry-based analytical procedures are extraction, high-performance liquid chromatography (HPLC) and mass spectrometry analysis including ionization, in most cases electrospray ionization and atmospheric pressure chemical ionization (APCI) (Figure 2). Each of these steps contributes to the specificity and sensitivity of the assay. During the extraction procedure potentially interfering compounds are eliminated from the resulting extract. HPLC separates the compounds of interest from remaining interferences. The problem is that although mass spectrometry detection can be highly specific, using concepts such as multiple reaction monitoring and/or exact mass, the ionization procedure is not. If samples are analyzed by LC-MS, it should be kept in mind that especially in the case of multiple analytes that are quantified simultaneously, the likelihood of significant interactions during the ionization process resulting in ion suppression or enhancement is increased [21]. A good example is reference [22] showing that during profiling of high-energy phosphates using LC-MS, the analytes suppressed each other's ionization if not completely separated by HPLC. As of today, the ionization process has to be considered the Achilles' heel of mass spectrometry-based multiplexing. The extraction procedure and HPLC separation are critical to reduce the number of potentially interfering analytes that are in the ion source at any given time and thus are important for assay performance including sensitivity, specificity and quantification. Other strategies to reduce the negative consequences of matrix effects are the use of ion sources that are less prone to ion suppression such as atmospheric pressure chemical ionization (APCI) or nanospray sources. [23] The use of appropriate deuterated internal standards will not only help in compensating for variability during extraction and degradation of the analyte, but will also compensate for fluctuation in ionization efficiencies [16]. However, the development of multiplexing assays that quantify a range of analytes with a variety of chemico-physical properties requires compromises. Dependent on the number of analytes, it becomes increasingly difficult to find extraction, HPLC and ionization conditions that allow for the reduction of matrix effects while allowing for the simultaneous quantification of the analytes with sufficient sensitivity. While non-targeted metabolomics and proteomics assays usually cannot take such effects into consideration, this is critical for targeted validated quantitative assays.

Another potential problem that is often underestimated is maintenance of sample integrity. Sample integrity is defined as stability of the analyte(s) in the biological matrix throughout variable environments spanning from sample collection, storage, shipping and further storage up to the last sample analysis. [14] The analytical results and the conclusions drawn from the results can only be valid if the sample that reaches the laboratory at the moment of processing for analytical analysis is of sufficient quality. The greater number of analytes that are included in a multiplexing assay, the greater the difficulty in establishing sample handling protocols that will ensure the stability of all compounds of interest.

Overall, it can be expected that the number of analytes that can simultaneously be quantified by mass spectrometry-based multiplexing assays while meeting acceptance criteria during validation is limited.

4. Validation and Quality Control

During molecular marker development and after the molecular marker of interest has been identified, the next step is to establish a targeted and validated assay that is capable of quantifying these specific compounds with acceptable total imprecision and sensitivity.

If a molecular marker plays an important role during the pre-clinical development of a drug, quantification of the molecular marker may need to comply with the rules of good laboratory practices (GLP). Validation of analytical assays for the quantification of metabolic and protein molecular markers may have to follow applicable regulatory and other guidances and standards, including but not limited to, FDA [24], Clinical Laboratory and Standard Institute [25], and EMEA/ICH [26]. Thus assay validation should include determination of the lower limit of detection, the lower limit of quantitation, the range of reliable response, the intra- and interday accuracy and imprecision, absolute extraction recovery and, if applicable, of dilution integrity. Carry over, matrix interferences and ion suppression/enhancement should be excluded. The following stabilities should be established: storage stability, freeze-thaw stability, bench top stability, autosampler stability (processed sample stability) and stock solution stability. Although GLP is limited to pre-clinical drug development and does not apply for the measurement of clinical samples, similar standards have been recommended and generated for sample analysis for clinical trials. The British Association of Research Quality Assurance published a detailed quality system termed Good Clinical Laboratory Practice (GCLP) [27] and in the United States the standards for accredited analytical laboratories have to be followed, for example those of the College of American Pathologists. As for pre-clinical studies, the “fit-for-purpose” principle applies implying that the analysis of clinical samples should closely follow the same principles as established in GLP guidances.

Method validation should demonstrate that a particular assay is “reliable for the intended application” and, thus as aforementioned, the rigor of method validation increases, from the initial validation proposed mainly for exploratory purposes, to the more advanced validation dependent on the evidentiary status of the biomarker and/or the use of the results. [11]

It has been recommended to differentiate the following categories of molecular marker assays: qualitative, quasi-quantitative, relative quantitative and definitive quantitative [27]. Based on this proposal, qualitative assays will only require establishing sample stability, sensitivity and specificity. In addition, quasi-quantitative assays will require determination of the range of reliable response and imprecision. Relative and definitive quantitative assays will require a complete validation as detailed in applicable guidance documents. Markers used as clinical diagnostic tools will always require a complete validation.

Validation of more complex molecular markers that are based on multiplex assays can be challenging. In most cases, such markers are endogenous compounds so an appropriate blank matrix may not be available. Sometimes charcoal stripping, the use of corresponding matrices from other species or artificial surrogate matrices, may provide a solution. In the case that blank matrices are not available, samples from healthy individuals or animals, depending on the species relevant for molecular marker testing and preferably with low concentrations of the compound of interest, have to be enriched with the reference compounds and the endogenous signal has later to be subtracted. In this case the result is influenced by two measurements, the endogenous signal and the signal of the added compound.

Based on generally accepted guidances, an assay is only acceptable if, except at the lower limit of quantitation, inter-day precision is $\leq 15\%$ and inter-day accuracy is within $\pm 15\%$ of the nominal value. Accordingly, an analytical run of study samples is accepted if at least 2/3

of the quality control samples fall within 15% of their nominal value. For the quantification of macromolecules using immuno-based assays regulatory agencies may accept limits of ± 25 and $\pm 30\%$. [27]

It has to be realized that current regulatory guidances have been written mostly with the quantification of single drug compounds in mind and may be too rigid for emerging multiplexing technologies. The challenge with multiplexing assays is that several compounds are quantified simultaneously and that, as aforementioned, it is not possible to optimize the assay for each compound to the extent that is possible for analysis of single compounds. The major differences between the quantification of drugs and molecular markers are listed in more detail (Table 3). In addition, the larger the number of simultaneously measured compounds, the higher the statistical probability that one accidentally fails to meet acceptance criteria.

5. Discussion

It has been suggested that molecular marker multiplexing assays should not be classified as ‘GLP’ assays, nor should they be validated by the same guiding principles developed for drug analysis by HPLC/UV or LC–MS. [14,27].

On the other hand, there is substantial benefit in the additional information conveyed by molecular marker multiplexing assays so that the risk-benefit-ratio has to be evaluated using different criteria than with standard assays designed to measure single drug compounds. This has been recognized by regulatory agencies. FDA guidelines have suggested that ‘further research is needed to establish the validity of available tests and determine whether improvements in biomarkers predict clinical benefit’. [28]

Based on the discussion above, it becomes obvious that there is (A) a clear need for validating multiplexing assay to realize their full potential in drug development and clinical diagnostics and (B) that the standard acceptance criteria that have mostly been developed with single drug analyte assays in mind may not be appropriate for multiplexing assays. It seems necessary to develop novel concepts for quality control and acceptance criteria that are considered fit for GLP studies and clinical diagnostic tools taking the characteristics of multiplexing assays into account. It seems reasonable to assume that the lack of clear regulatory guidance is one of the reasons preventing a more widespread development and utilization of multiplexing assays in drug development and clinical diagnostics.

One of the paradigms that will need to be emphasized over the currently widely used concept of applying standard acceptance criteria across a wide variety of bioanalytical assays is the “fit-for-purpose” principle. Today, it is common practice that all analytes measured in the same assay have to meet the same acceptance criteria and that if one analyte fails to meet acceptance, the whole analytical run fails even if all other analytes are within predefined quality specifications.

As discussed above, due to the potentially relatively wide range of physico-chemical properties of analytes included in a multiplexing assay, the performance of specific analytes may vary and some analytes may be more likely to fail standard acceptance criteria than others. Therefore, it has been suggested to widen the acceptance criteria from 15 to 25% or even 30% for molecular marker multiplexing assays. [27] Although this may be acceptable for some analytes, this may be too liberal for others. A “one fits all” approach to assay acceptance may not be a reasonable strategy for multiplexing assays. This may depend on the extent of change of a specific parameter that is considered clinically relevant. In many cases a clinically significant change is assumed, when a clinical marker changes by 20%. In such a case, an assay with an analytical imprecision of 30% would result in a molecular

marker with poor sensitivity that will not be able to pick up smaller but potentially clinically relevant differences. On the other hand, if a parameter has at least to triple for the change to be considered clinically relevant, an analytical imprecision of even 50% may still be acceptable. Therefore, it seems a more reasonable approach to allow for the definition of different acceptance criteria for individual analytes measured within a multiplexing assay. Such acceptance criteria should be defined based on the pre-study validation results in the context of mechanistic as well as clinical considerations for each individual component of a molecular marker multiplexing assay.

Today, an assay is either considered quantitative and all analytes meet pre-defined performance and quality criteria or it is not. However, for molecular marker multiplexing assays it may be of value to simultaneously analyze quantitative and qualitative markers. Even the inclusion of one or more exploratory markers that have not been further validated may add value if there are compelling scientific and clinical reasons for inclusion.

Overall, the answer to the question how many analytes can reasonably be included in an LC-MS-based multiplexing assay depends on several factors:

- A. *The analytical assay:* Sample stability and handling, extraction, HPLC, ionization conditions and mass spectrometry detection (e.g. positive or negative mode and number of ion transitions monitored) limit the number of compounds that can reasonably be included in the same assay. The range of physico-chemical properties is often determined by the nature of the molecular marker panel that needs to be assessed in a multiplexing assay. These can roughly be classified into two categories:
- **Functional multiplex:** These are molecular marker assays that capture different functional aspects of a cell or an organ by assessing different pathways. This usually requires the quantification of molecules that are structurally unrelated. An example is an assay established and validated in the authors' laboratory that assesses vascular endothelial dysfunction in plasma and includes arginine, ornithine, citrulline, asymmetric and symmetric dimethyl arginine (ADMA, SDMA), cysteine, homocysteine and glutathione.
 - **Structural multiplex:** These are assays that quantify structurally related analytes. Examples established in the authors' laboratory are the profiling of prostaglandins in plasma and sputum and of fatty acids in very low-density lipoprotein (VLDL) particles.

Due to the structural similarities of the analytes, it is easier to integrate a larger number of analytes in structural multiplexes than into functional multiplexes.

- B. *The required sensitivity:* A frequent limitation is the differences in physiological concentrations of molecular markers in the matrix of interest that can differ by several orders of magnitude. This may preclude markers from being combined in the same assay. An example is the above-mentioned assay to assess vascular endothelial function. Although the inclusion of 15-F_{2t}-isoprostane as a stable marker of oxidative stress would have been attractive, 15-F_{2t}-isoprostane requires a lower limit of quantification of 2.5 pg/mL in plasma that can only be achieved in a highly specific and sensitive assay that specifically addresses the challenges of this analyte. [29,30] On the other hand, a highly abundant analyte may require dilution of the samples and this may result in insufficient sensitivity for other compounds of interest.

- C. *The required throughput:* During clinical drug development and in clinical diagnostics high sample throughput is often desirable and most LC-MS assays have run times of less than 10 minutes. However, the separation and elution of analytes with a wide range of hydrophilicity/lipophilicity may require relatively long chromatography times. In such cases, it may be advisable to split the assay into two or more shorter assays with more specialized and shorter chromatography for specific groups of analytes.
- D. *Purpose of the assay and required validation:* As mentioned above, during clinical drug development and as a clinical diagnostic tool, different levels of validation are required. Non-targeted mass spectrometry assays represent the extreme of multiplexing assays. These are purely exploratory and cannot be validated. The number of analytes that can reasonably be included in a multiplexing assay is reduced by the required extent of validation and the stringency of the acceptance criteria. If compliance with acceptance criteria as recommended by current applicable regulatory guidances (*vide supra*) are required, the authors' experience is that assays with more than 10 analytes often become challenging, but again this will be determined by the diversity and structural relationship of the analytes included. Instead of being challenged with frequent rejection of analytical runs, two or more shorter and more specialized assays may present a more efficient alternative.

In conclusion, multiplexing assays always constitute a compromise and in most cases will not perform as well as assays specifically developed for individual analytes. The number of analytes that can reasonably be included in a mass spectrometry-based multiplexing assay depend on several factors such as stability and physico-chemical properties of the analytes and how efficiently these can be integrated in a single assay considering sample handling and stability, extraction, HPLC separation, ionization conditions and mass spectrometry detection. Another aspect includes biomedical considerations such as the differences in physiological concentrations of analytes, the required concentration range and how much variability is acceptable before the clinical utility of a marker is negatively affected. A third category of limitations are regulatory considerations including validation and quality control during the analysis of study samples or the analysis of clinical samples. In this regards, it is a challenge that current bioanalytical guidelines have mostly been developed for single analyte assays for drug compounds and do not seem adequate for multiplexing molecular marker assays that in most cases quantify endogenous compounds. Guidances for multiplexing assays should be developed. When developing a complex multiplexing assay, even if it is possible to integrate a wide variety and large number of analytes, it should always be kept in mind that a set of shorter and more specialized assays may offer a more manageable and efficient alternative.

Acknowledgments

This work was supported by the United States National Institutes of Health grants NIDDK P30 DK048520 and NICHD R01 HD070511.

References

1. Oresic M, Vidal-Puig A, Hänninen V. Metabolomic approaches to phenotype characterization and applications to complex diseases. *Expert Rev Mol Diagn.* 2006; 6:575–585. [PubMed: 16824031]
2. Biomarkers Definition Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther.* 2001; 69:89–95. [PubMed: 11240971]
3. Nobeli I, Thornton JM. A bioinformatician's view of the metabolome. *Bioessays.* 2006; 28:534–45. [PubMed: 16615085]

4. Christians U, Klawitter J, Hornberger A, Klawitter J. How unbiased is non-targeted metabolomics and is targeted pathway screening the solution? *Curr Pharm Biotechnol.* 2011; 12:1053–1066. [PubMed: 21466457]
5. U.S. Department of Health and Human Services, Food and Drug Administration. [accessed 11/4/2011] Challenge and opportunity on the critical path to new medical products. 2004. http://www.nipte.org/docs/Critical_Path.pdf
6. Müller PY, Dieterle F. Tissue-specific, non-invasive toxicity biomarkers: translation from preclinical safety assessment to clinical safety monitoring. *Expert Opin Drug Metab Toxicol.* 2009; 5:1023–1038. [PubMed: 19611406]
7. United States Food and Drug Administration, Center for Drug Evaluation and Research. [accessed 11-4-2011] Guidance for the Industry. Pharmacogenomic data submissions. Mar. 2005 <http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm126957.pdf>
8. Goodsaid F, Frueh F. Biomarker qualification pilot process at the US Food and Drug Administration. *AAPS J.* 2007; 23:E105–108. [PubMed: 17408233]
9. Goodsaid F, Frueh F. Implementing the US FDA guidance on pharmacogenomic data submissions. *Environ Mol Mutagen.* 2007; 48:354–358. [PubMed: 17567848]
10. Goodsaid F, Frueh F. Process map proposal for the validation of genomic biomarkers. *Pharmacogenomics.* 2006; 7:773–782. [PubMed: 16886901]
11. Wagner JA, Williams SA, Webster CJ. Biomarkers and surrogate end points for fit-for-purpose development and regulatory evaluation of new drugs. *Clin Pharmacol Ther.* 2007; 81:104–107. [PubMed: 17186007]
12. Lee JW, Figeys D, Vasilescu J. Biomarker assay translation from discovery to clinical studies in cancer drug development: quantification of emerging protein biomarkers. *Adv Cancer Res.* 2007; 96:269–298. [PubMed: 17161683]
13. Burckart GJ, Amur S, Goodsaid FM, Lesko LJ, Frueh FW, Huang SM, Cavaille-Coll MW. Qualification of biomarkers for drug development in organ transplantation. *Am J Transplant.* 2008; 8:267–270. [PubMed: 18093264]
14. Lee JW, Weiner RS, Sailstad JM, Bowsher RR, Knuth DW, O'Brien PJ, Fourcroy JL, Dixit R, Pandite L, Pietrusko RG, Soares HD, Quarmby V, Vesterqvist OL, Potter DM, Witliff JL, Fritche HA, O'Leary T, Perlee L, Kadam S, Wagner JA. Method validation and measurement of biomarkers in nonclinical and clinical samples in drug development: a conference report. *Pharm Res.* 2005; 22:499–511. [PubMed: 15846456]
15. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research. [accessed 11-4-2011] Guidance for Industry: Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products. 1998. Version May 1998. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm078749.pdf>
16. Rifai N, Gillette MA, Carr SA. Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nat Biotechnol.* 2006; 24:971–983. [PubMed: 16900146]
17. [accessed 11-4-2011] Code Federal Regulations, Title 21, Food and drugs chapter i- -Food and Drug Administration, Department of Health and Human Services. Subchapter h, Medical Devices. Part 807, Establishment registration and device listing for manufacturers and initial importers of devices. <http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?CFRPart=807>
18. [accessed 11-4-2011] Code Federal Regulations, Title 21, Food and drugs chapter i- -Food and Drug Administration, Department of Health and Human Services. Subchapter h, Medical Devices. Part 814, Premarket approval of medical devices. <http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?CFRPart=814>
19. Ministerial ordinance on standards for manufacturing control and quality control for medical devices and in-vitro diagnostic reagents. Pharmaceuticals and Medical Devices Agency; Tokyo: http://meddevmirq.com/Documents/MiRQ%20Overview%20of%20Japanese%20Regulatory%20System_PAL%20Summary%202009_07_18%20Sec%202.pdf
20. Dati F. The new European directive on in vitro diagnostics. *Clin Chem Lab Med.* 2003; 41:1289–1298. [PubMed: 14580154]

21. Annesley TM. Ion suppression in mass spectrometry. *Clin Chem.* 2003; 49:1041–1044. [PubMed: 12816898]
22. Klawitter J, Schmitz V, Klawitter J, Leibfritz D, Christians U. Development and validation of an assay for the quantification of 11 nucleotides using LC/LC-electrospray ionization-MS. *Anal Biochem.* 2007; 365:230–239. [PubMed: 17475198]
23. Wickremesinha ER, Ackermann BL, Chaudhary AK. Validating regulatory-compliant wide dynamic range bioanalytical assays using chip-based nanoelectrospray tandem mass spectrometry. *Rapid Commun Mass Spectrom.* 2005; 19:47–56. [PubMed: 15570573]
24. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research and Center for Veterinary Medicine. [accessed 11-4-2011] Guidance for the Industry. Bioanalytical Method Validation. 2001. Version May 2001. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM070107.pdf>
25. Clinical Laboratory and Standard Institute. [accessed 11-4-2011] www.clsi.org
26. European Agency for the Evaluation of Medicinal Products. [accessed 11-4-2011] ICH topic Q2 (R1). Validation of analytical procedures: text and methodology. Note for guidance on analytical procedures: Text and methodology. 1995. CPMP ICH/381/95. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500002662.pdf
27. Cummings J, Ward TH, Greystoke A, Ranson M, Dive C. Biomarker method validation in anticancer drug development. *Br J Pharmacol.* 2008; 153:646–656. [PubMed: 17876307]
28. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research. [assessed 11-4-2011] Guidance for Industry: Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics. 2007. Version May 2007. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm071590.pdf>
29. Haschke M, Zhang YL, Kahle C, Klawitter J, Korecka M, Shaw LM, Christians U. Quantification of 15-F_{2t}-isoprostane in human urine and plasma using high-performance liquid chromatography – atmospheric pressure chemical ionization-tandem mass spectrometry. *Clin Chem.* 2007; 53:489–497. [PubMed: 17259231]
30. Klawitter M, Haschke T, Shokat J, Klawitter U, Christians U. Quantification of 15-F_{2t}-isoprostane in human plasma and urine: ELISA and LC-MS/MS results cannot be compared. *Rapid Commun Mass Spectrom.* 2011; 25:463–468. [PubMed: 21259353]
31. Goodsaid FM, Frueh FW, Mattes W. Strategic paths for biomarker qualification. *Toxicology.* 2008; 245:219–223. [PubMed: 18280028]

Highlights

- Molecular marker panels have the potential to revolutionize clinical diagnostics and drug development
- mass spectrometry-based multiplexing is a powerful tool to measure molecular marker panels in one run
- the number of compounds that can reasonably be integrated in one assay is limited by the physicochemical properties of the compounds, extraction, HPLC, mass spectrometry detection, required throughput, differences in physiological concentrations, required sensitivity and ability to meet acceptable validation and assay performance.

Figure 1. Development and Regulatory Review of Drugs and Context- Dependent Molecular Markers in Parallel to Drug Development

The regulatory review process is based on the United States Federal Drug Administration pilot process for biomarker qualification [31]. In general, the review processes proposed or established by other regulatory agencies follow a similar flow. The FDA biomarker review process includes the following steps [31]: (A) Submission of an initial letter defining the biomarker, its context and data sources for its qualification, and review by an Initial Proposal Review Group (IPRG), (B) Decision to proceed to a full qualification by the Biomarker Qualification Review Team (BQRT), (C) Full submission of qualification data for review by the BQRT. (C) Voluntary eXploratory Data Submission (VXDS) meeting to go over the qualification data and to identify potential information gaps before a full review can be completed for the qualification package. (D) Review drafted by BQRT, (E) internal review at FDA, and (F) communication of decision to sponsor.

Figure 2. Achievement of Specificity During HPLC-MS Analysis Workflow

HPLC-MS analysis usually comprises four steps: extraction, HPLC separation, ionization and mass-spectrometry detection. During these steps specificity of detection is achieved using different concepts. During extraction interfering substances are eliminated, during HPLC they are separated from the compounds of interest, and during mass spectrometry compounds of interest are selectively detected and all others are ignored. The contribution of ionization to the selectivity of an assay is limited. Ionization can potentially become a major cause of matrix effects leading to ion suppression and/or enhancement.

Table
Comparison of bioanalytical assays for xenobiotics and endogenous molecular markers multiplexing assays

Current guidances for the validation of bioanalytical assays such as references [24–26] have primarily been written with the quantification of drug molecules and xenobiotics in mind. Molecular markers have different challenges since they are endogenous and in many cases these assays are multiplexing assays that measure several molecules simultaneously. While most drugs are small molecules, many molecular marker assays require the measurement of macromolecules such as proteins. The table is based on reference [27]

Parameter	Assay for Xenobiotics	Molecular Marker Assays
Specificity	Xenobiotics, by definition, are not present in body fluids; tissues and blank matrices are available. Blank matrices are essential for chromatography-based and LC-MS assays to evaluate the presence of interfering signals.	Molecular markers monitor the response of a biological system in response to xenobiotics and diseases. This requires the measurement of endogenous compounds that are usually present in the matrix of interest. Blank matrices are typically not available or matrices have to be modified to remove endogenous compounds. In many cases matrices also contain other compounds with very similar structures that may interfere with the assays. An example is arachidonic acid derivatives and metabolites. The analytes themselves may not be consistent. In the case of proteins this may be due to missed cleavages and the extent of post-translational modifications. The analysis of blank and zero samples as required by guidance documents for analytical assay validation may not be possible.
Matrix interferences	Assays typically include specific sample clean up procedures and chromatographic separation to minimize interferences.	Especially for multiplexing assays, options for sample clean up are often limited due to the different physical-chemical properties of the analytes. Also complete chromatographic separation is often impossible. This increases the likelihood of matrix interferences, such as ion suppression and enhancement in ESI-MS assays. For immuno-based multiplexing assays it is often not possible to test for all possible interferences for each antibody-analyte interaction.
Sensitivity	Detection for individual compounds can be optimized, including sample clean up, analytical columns, solvent composition, tuning of the MS instrument, <i>etc.</i>	For multiplexing assays, the optimization of LODs for individual analytes is often not possible.
Reference materials	Reference materials with certificates of analysis are available.	Purified and standardized reference materials are often not available.
Internal standards	Appropriate internal standards compensate for variability during extraction, stability, and ionization of compounds. Deuterated internal standards are considered ideal.	Appropriate internal standards are often not available.
Calibrators	Calibrators are generated by adding the analyte to the blank matrix.	Blank matrix is often not available; calibrators have to be generated by enriching a matrix that already contains the endogenous compounds of interest with known amounts of the analytes. The quality of the calibrators varies, <i>e.g.</i> , depending on the amount of endogenous compounds on the matrix. The use of blank surrogate matrices may not adequately reflect the situation of the study samples. The inclusion of blank and zero samples may not be possible.
Calibration curve fit	The best fit for the analyte is chosen.	If multiple compounds are analyzed, the fit for individual compounds may be different. Compounds may have limited dynamic ranges.
Range of reliable response	Is defined by the lower LOQ and the upper LOQ. The pharmacologically/toxicologically relevant concentration ranges are usually known.	For many endogenous compounds, such as proteins, baseline concentrations, and normal ranges are not known and thus, the target lower LOQ is unknown. The same applies to the required ranges of reliable response. For multiple analytes the concentration ranges may be very different and accommodating all analytes may limit the dynamic range of individual compounds.
QC samples	QC during study sample analysis typically includes blank samples spiked with known analyte concentrations and pooled samples of treated subjects.	Since in many cases blank samples are not available either pooled samples of healthy controls or patients containing different concentration ranges of the analytes or samples enriched with known amounts of the compounds have to be used. Blank surrogate matrices may be used, but may not adequately control all aspects of the analytical procedure. multiple analytes are measured

Parameter	Assay for Xenobiotics	Molecular Marker Assays
		simultaneously, there is a statistically greater chance that the QC values of one analyte may not meet acceptance criteria (= 2/3 of the QC samples within ± 15 % of the nominal concentration). This should not necessarily lead to rejection of a run and all analyte measurements.
Precision and accuracy	Assays are acceptable if, except at the lower LOQ, total imprecision is ≤ 15 % and accuracy within ± 15 % of the nominal concentration.	Measurement of macromolecules such as proteins is inherently more variable and larger ranges of total imprecision and accuracy than for small molecules may have to be considered acceptable. In multiplexing assays, one or a few compounds may not meet acceptance criteria. This does not necessarily mean that the whole assay is not valid; there may still be value to measuring these compounds albeit with lower precision and accuracy than the other compounds in the assay. Cross validation with established assays is an efficient strategy to ensure accuracy. For example the results of individual compounds obtained with a multiplexing assay can be compared with those of ELISA or LC-MS assays.
Robustness	Robustness describes the susceptibility of an assay to deliberate variations of its conditions, such as temperatures, pH, solvent compositions, and timing.	Due to the higher complexity of matrices, the measurement of inherently instable molecules and the measurement of multiple compounds with different physical-chemical properties, in many cases molecular marker assays are less robust than assays for xenobiotics. The robustness for the measurement of individual compounds may vary. This has to be taken into account during assay development and validation.
Stability	Stability of xenobiotics and drugs is often already known or is known early during development.	In the case of multiplexing assays, different compounds may have different stabilities under various conditions. Control of stability and sample validity often have to start with the sample collection procedure.