FDA perspectives on potential microarray-based clinical diagnostics

Živana Težak,[1] Daya Ranamukhaarachchi,[2] Estelle Russek-Cohen[3] and Steven I. Gutman[1]*

[1]Office of *In Vitro* Diagnostic Device Evaluation and Safety, Center for Devices and Radiological Health, US Food and Drug Administration, Rockville, MD 20850, USA
[2]Office of Science and Engineering Laboratories, Center for Devices and Radiological Health, US Food and Drug Administration, Silver Spring, MD 20903, USA
[3]Office of Surveillance and Biometrics, Center for Devices and Radiological Health, US Food and Drug Administration, Rockville, MD 20850, USA
**Correspondence to*: Tel: +1 240 276 0597; Fax: +1 240 276 0652; E-mail: zivana.tezak@fda.hhhs.gov.

Date received: 2nd September 2005

Abstract

The US Food and Drug Administration (FDA) encourages the development of new technologies such as microarrays which may improve and streamline assessments of safety and the effectiveness of medical products for the benefit of public health. The FDA anticipates that these new technologies may offer the potential for more effective approaches to medical treatment and disease prevention and management. This paper discusses issues associated with the translation of nucleic acid microarray-based devices from basic research and target discovery to *in vitro* clinical diagnostic use, which the Office of *In Vitro* Diagnostic Device Evaluation and Safety in the Center for Devices and Radiological Health foresees will be important for assurance of safety and effectiveness of these types of devices. General technological points, assessment of potential concerns for transitioning microarrays into clinical diagnostic use and approaches for evaluating the performance of these types of devices will be discussed.

Keywords: *microarray,* in vitro *diagnostics, FDA, gene expression, genotyping, array comparative genomic hybridisation (array CGH)*

Introduction

The question of whether high-throughput gene platforms like nucleic acid microarrays are ready to be used as clinical diagnostic devices[1] is the focus of many ongoing discussions and efforts in the scientific and medical communities. Specific concerns include standardisation, reproducibility and accuracy, appropriate statistical approaches for data analysis and validation and the availability of control materials to ensure that these platforms can be reliably used in clinical diagnostics.

In this paper, we will attempt to assess areas of concern in transitioning microarrays from research into the clinical diagnostic arena as *in vitro* diagnostic devices (IVDs),[2] and provide information on the scientific issues that need to be considered to demonstrate whether these devices are safe and effective as clinical diagnostic tools.

'The Critical Path to New Medical Products' of the US Food and Drug Administration (FDA) has identified genomic technologies as being crucial in advancing medical product development and personalised medicine.[3] The microarray community and regulatory agencies are coordinating efforts to develop standards to support microarray data for regulatory decision making. The FDA has participated in a series of public workshops,[4–6] drafted several guidances[7,8] related to genomics and pharmacogenomics and released a concept paper addressing the co-development of pharmacogenomic-based therapeutic agents and diagnostic tests necessary for therapeutic decision making.[9] Recently, a joint FDA/Johns Hopkins University/Pharmaceutical Research and Manufacturers of America (PhRMA) workshop entitled 'Microarrays in Transcriptional Profiling' focused on expression profiling issues as they may relate to nucleic acid microarrays for diagnostic use.[10] Issues that were addressed included microarray controls, variation between laboratories and platforms, data normalisation, microarray scanner performance and clinical validation.

Technology overview and potential concerns for translation to clinical diagnostics

We will limit the scope of this paper to nucleic acid microarrays and will not address tissue and protein microarrays. We begin with a brief technology overview, emphasising

 © HENRY STEWART PUBLICATIONS 1473–9542.

potential areas of concern in transitioning microarrays into the clinical IVD arena for three types of nucleic acid arrays: (i) gene expression arrays, (ii) genotyping arrays and (iii) comparative genomic hybridisation arrays (array-CGH).

Gene expression arrays

Gene expression profiling as a genomic technique is intended to determine the fraction of genes that are expressed — that is, actively transcribed into mRNA under specific circumstances in certain cells. Gene expression varies depending on factors such as cell lineage, stage of differentiation, extracellular stimuli or intracellular regulation. Most microarrays used in exploratory expression profiling measure the relative expression of tens of thousands of genes at a time, creating a molecular profile of the mRNA in a specific sample. It is believed that expression profiling will become a useful tool for evaluating disease susceptibility, to make earlier or more reliable diagnoses and to classify tumours by their molecular signature.[11,12] For example, we may expect a potential diagnostic microarray device to be used for class prediction among existing classes of disease types — that is, to determine in which category a patient belongs. In the case of diseases where the expression of multiple genes may be needed to distinguish between the disease states or types, such as certain types of cancer,[12] there is an implied need for a diagnostic platform that can reliably and quantitatively measure expression differences of a few hundred genes, mirroring clinically important diagnostic and prognostic differences.[12]

Although gene expression microarray analysis is a tremendously promising field, many problems remain to be addressed. Comparing gene expression levels using microarrays is a very complex process, highly prone to variation.[13,14] Specific concerns include the reproducibility of the results between different laboratories and operators using the same platform, as well as the ability to reproduce results generated for the same samples using a different platform. The number and type of probes on a microarray may alter its performance and hence its reproducibility.

Several recent publications have provided valuable insights and suggested ways of overcoming some of the sources of variation in microarray experiments.[15–21] These studies demonstrated that better reproducibility can be achieved between array platforms and among laboratories if experiments are carefully designed, controlled and executed.[16] Implementation of standardised protocols for both experimental and computational aspects of the comparison study led to a dramatic increase in reproducibility and also underlined the need to make raw data available, so that data normalisation can be understood and possibly reassessed using a more appropriate algorithm.[15] Many microarray normalisation methods currently used in research and discovery involve the use of data from all arrays within a given experiment.[21,22] This approach might not be practical for a commercial device, where data normalisation may need to be performed for a single patient array, possibly requiring a reference database.[23]

Concordance between two array platforms largely depends on the exact gene sequences measured by each platform and how the measured genes map to each other.[18,24] Discordant measurements between platforms may, in many cases, be a consequence of a significant level of cross-hybridisation[24] or differential use of splice variants inherent in the probe design. If quantitative reverse transcriptase polymerase chain reaction (QRT-PCR) is used to verify discordant measurements between the microarray platforms, this verification may depend on the region of the gene from which the QRT-PCR assay is designed. Thus, QRT-PCR may not be able to validate all microarray platform results, whether discordant or concordant among different microarray platforms.[17,19]

Microarray technology allows the analysis of large quantities of data without a starting hypothesis and the resulting patterns may have no clear biological meaning. Therefore, assessing the clinical and analytical performance of the potential diagnostic test may be even more important for RNA expression analysis than in other fields. Many of the published studies, however, have not appropriately demonstrated analytical reproducibility or clinical validity.[25,26] Often, RNA expression data may not be generalisable, such as when the method is developed on one group of patients but subsequently evaluated on another patient group with different characteristics.[27] Bias is often present in patient-selection and result interpretation. Bias may also be due to sample collection, processing and storage, or to factors related to operator performance or instrument use.[25]

Another statistical issue related to analysis of the complex datasets from expression microarray experiments is over-fitting of the data when using multivariate models for a large number of potential predictors while trying to discriminate using a relatively small number of patients.[26,28] While models based on large numbers of predictors appear to do well in a training set, this performance is often not duplicated when using a separate validation set. Training and validation (testing) sets of samples should be distinct and independent to avoid these pitfalls.[29] It should be noted that subdividing a group of patients in two — for example using half of the data from each of four hospitals rather than using data from distinct hospitals for training and validation parts — may not yield a reliably independent test of a device performance. If reproducibility of the study results has not been demonstrated using an independent validation set of adequate size, the resulting predictors cannot be accepted as giving definitive conclusions for use in clinical practices since they may lead to inappropriate treatment.[26] For example, data used to develop a composite biomarker to predict drug response should be distinct from the data generated and used to evaluate the actual drug response within patient subsets having the specific pharmacogenomic profile.[28,30]

Genotyping arrays

Microarrays can be used to identify heritable variation or somatic changes of DNA in individuals and across populations (ie genotype).[31,32] Arrays can detect many forms of heritable genetic variation, including single nucleotide polymorphisms (SNPs), single or multiple base insertions and deletions, gene conversions and repeats. Polymorphism analysis using microarrays can potentially be used to detect coding changes associated with a biologically relevant phenotype (eg drug metabolism enzyme polymorphisms), for linkage analysis if performed on a genome-wide scale or in a dense chromosomal region, or for assessment of loss of heterozygosity (LOH) in cancer.[23]

Most genotyping arrays fall into one of two general categories — SNP arrays or resequencing arrays.[33] Both types primarily use highly redundant hybridisation-based discrimination. Genotyping arrays may query multiple polymorphisms or mutations in a single gene (for example, cystic fibrosis transmembrane conductance regulator, cytochrome P450 2D6 LOH) or multiple genetic loci throughout the genome. Many human SNPs that are disease-causing or are involved in response to drug therapies are well known and can be assessed in a highly parallel fashion using high-density oligonucleotide microarrays. If used in infectious disease, this type of array would allow for the detection of genotypic variants of microbial pathogens.[34–36]

A rapidly developing clinical application of microarray genotyping is pharmacogenetic testing.[37,38] A pharmacogenetic test can be defined as an assay intended to study inter-individual variations in DNA sequence related to drug absorption and disposition (pharmacokinetics) or to drug action (pharmacodynamics).[39] This includes polymorphisms within the genes encoding metabolising enzymes, transporters, receptors and other proteins. The ultimate promise of pharmacogenetics is the possibility that knowledge of a patient's DNA sequence might be used in drug therapy to maximise treatment efficacy, avoid adverse drug reactions, optimise doses and target drug treatment only to patients who are likely to respond.

Depending on the intended use of the potential IVD, many of the areas of concern identified for gene expression array technologies, such as bias, over-fitting and generalisability, may also apply to genotyping arrays.

Comparative genomic hybridisation using arrays (array-CGH)

Changes in DNA copy number in a genome are associated with gains and losses of chromosomes and chromosomal segments due to deletions, insertions and duplications. These can be germline mutations or somatic events. DNA alterations can lead to copy number polymorphisms in normal individuals,[40–42] but can also be the cause of various disease states. Detection of these alterations facilitates the identification of crucial genes and pathways involved in biological processes and diseases, and may aid in diagnosis and therapy for genetic and somatic diseases.[43,44] Array-CGH is a method for identifying variations in genomic DNA copy number between normal and pathological samples. In a typical array-CGH experiment, total genomic DNA isolated from reference and test samples is differentially labelled with two fluorescent molecules with different excitation and emission properties and hybridised to a microarray containing probes representing different regions of genomic DNA. Information on relative DNA copy number differences is generated based on the intensity of the fluorescence signal generated by each labelling molecule.

Depending on the amount of genome coverage, array-CGH can be either marker based or contig based. Marker-based arrays target specific regions selected based on previous knowledge of widely known chromosomal alterations. Contig-based arrays target the whole genome derived from probe sequences representing short intervals (as low as a kilobase scale) of the entire genome for the purpose of a new biomarker discovery for a given condition.[45] These two array types may include platforms using clones of bacterial artificial chromosomes, fosmid clones, cDNA and short oligonucleotides. These different platforms provide different levels of performance, such that some are more suitable for particular applications than for others. Factors that determine performance requirements include the magnitudes of the copy number changes, the state and composition of the specimen, how much material is available for analysis and how the results of the analysis will be used. There are advantages and disadvantages, as well as inherent technical difficulties, associated with each of these platforms, which impact on the accuracy of data generated.

Alternative procedures are being developed to tackle technical issues such as highly complex genomic material, the presence of repetitive sequences, method of sample processing (eg fresh tissues versus formalin-fixed tissues) and data management and extrapolation of data generated from array-CGH.[46–48] Once these technical hurdles are overcome, array-CGH may potentially be used for the analysis of clinical samples to identify genomic alterations, as well as for measuring the loss of allelic heterozygosity and improving quantitative accuracy, resolution and the dynamic range of the detection of genomic copy number variations compared with existing cytogenetic methods.[44,45,47] High analytical performance is a prerequisite for array-CGH to have clinical utility. For example, noise level of the array measurements and difference in behaviour of different array elements may need to be appropriately addressed. Reliable detection of single copy changes may be difficult in heterogeneous cell populations (eg intermixed tumour and normal cells); however, hybridisations of defined specimens with known aberrations may help in establishing the performance characteristics of array-CGH.[23]

Regulatory approaches for evaluation of microarray-based IVDs

Since the US Congress enactment of the Medical Device Amendments in 1976,[1] the FDA has evolved risk-based regulations and policies designed to promote and protect the public health by regulating medical devices. The FDA regulations also set as an equally important goal the encouragement of the discovery and development of new medical products for the benefit and promotion of public health. IVDs are considered to be medical devices for the purposes of regulatory oversight, and are defined as reagents, instruments and systems intended for use in the diagnosis of disease or in the determination of the state of health in order to cure, mitigate, treat or prevent disease.[2] Therefore, by statute, IVD tests that are to be commercialised for the diagnosis and management of patients are subject to FDA regulation.

Within the FDA, the Office of *In Vitro* Diagnostics (OIVD) within the Center for Devices and Radiological Health (CDRH) regulates devices pre- and post-marketing to ensure that they demonstrate a reasonable assurance of safety and effectiveness for the intended use according to the directions for use.[49] Each regulated device is assigned one of three risk-based classes related to the level of FDA oversight prior to marketing. Class I devices are generally considered low risk and many are exempt from pre-market notification to the FDA. Class II devices (and non-exempt Class I devices) are considered to carry more risk and are subjected to pre-market notification review by the FDA to determine whether they are similar (in terms of safety and effectiveness) to another legally marketed device intended for the same type of use. Class III devices are considered the highest risk devices and these devices require pre-market application approval, involving a more in-depth review and documentation of the safety and effectiveness of the device.[50]

The FDA/OIVD has put in place a so-called pre-Investigational Advice Exemption (IDE) mechanism[51] to allow early communication between the FDA and sponsors of new technology and to provide protocol review and regulatory guidance with no cost to the sponsor. This process can prevent manufacturers from wasting resources on studies that would not support FDA approval for their intended use and allows the FDA the opportunity to become familiar with the test before reviewing the formal pre-market submission. This may be especially important for devices in emerging fields such as microarray-based diagnostics, where both review policy and regulatory science are continually evolving.

As with other areas of genetic testing, there are unique scientific and ethical issues that will need to be addressed for the field of genomics/microarrays to advance and allow utilisation of microarrays in clinical diagnostics. We will attempt to address some of the issues that OIVD has identified in relation to the development of microarrays for use in clinical decision making, including intended use and analytical and clinical performance. More detailed information referring to some specific types of tests can also be found in Special Controls guidance documents issued by OIVD.[52–54] It should be noted that since these are relatively new types of IVD assays, requirements and performance characteristics are evolving and are not completely defined for every type of platform and assay.

Intended use

The intended use of a microarray-based IVD, and the risk the device poses to the patient according to its intended use, are the main determinants for the regulatory classification. The intended use of the device for which pre-market approval or clearance is sought should specify the analyte which the device is intended to measure, the clinical purpose of measuring the analyte and the populations for which the device is indicated, where appropriate. The exact types of analytical and clinical data that have to be submitted depend greatly on the claims made in the intended use.

The risk to a patient from any IVD device is primarily related to the use of the information derived from the test results, and is not linked to the complexity of the technology used in the device. The risk to the patient of false-positive or false-negative results is often tied to the clinical decisions made based on the result. For example, if a device is being used to determine whether a patient will benefit from a drug that has significant toxicity, it could be a high-risk device: the patient may experience unnecessary side-effects with no clinical benefit following an incorrect test result. The opposite case is that a patient who could potentially benefit from the therapy could be sub-optimally treated based on incorrect diagnostic test results predicting adverse events. This emphasises the need for appropriate analytical and clinical validation of the assay.

Analytical performance

Adequate analytical performance of a diagnostic test is crucial to allow confidence in the test results. Microarray-based devices need to be validated analytically to determine that they can measure the intended analyte reproducibly and accurately when the assay is performed by the intended user (eg a clinical laboratory with trained and experienced operators). The analytical performance of the test should be clearly known in order to evaluate the device risk. More technologically or conceptually complex tests may require higher levels of analytical evaluation. The users should be aware of any analytical limitations of the test, so that test results may be interpreted in light of these limitations. High analytical variation may decrease the reliability of test results, compromising the clinical utility of the test.

Evaluation of microarray analytical performance is complicated by the large number of separate analytes evaluated simultaneously, as well as by the massive amount

of post-analytical data processing required to obtain the result; however, pre-analytical steps should not be overlooked as they are an extensive source of variation in most microarray experiments. Methods for appropriate sample quality assessment at different steps during pre-analytical processing (eg RNA quality assessment, labelling, amplification) can be very useful in demonstrating the analytical performance of these tests. Analytical validation requires the use of appropriate controls such as hybridisation controls and sample controls. There are several ongoing joint efforts between the FDA, the National Institute of Standards and Technology (NIST) and the microarray community to make this process easier:[10,55–57] this is described in more detail below.

For quality control, both positive and negative controls should be included and controls should reflect sample composition and DNA concentration. Controls should show that all steps and critical reactions have proceeded properly and without contamination or cross-hybridisation. For each microarray feature or element, the manufacturer of a specific microarray device should assure the identity of the probe or feature and its reproducible placement on the array. Analytical verification should demonstrate that the test adequately detects what it is supposed to detect. This can be relatively straightforward for genotyping arrays, but potentially very complex in the case of arrays measuring gene expression levels. The same is true for demonstrating the diagnostic utility of the test.

For example, a microarray-based test that measures a set of SNPs to aid in the determination of drug therapy will need to demonstrate that the test can detect the polymorphisms it is designed to detect and, ideally, no others. This may be straightforward for genes that are easily amplified, without excessive numbers of repeats or interfering sequences; however, the inclusion of other genetic modifications (eg duplications, deletions, inversions) or additional SNPs in the assay will increase the potential for analytical complications.

The reproducibility of microarray-based devices should be thoroughly evaluated. This becomes especially important as devices become increasingly complex, and reproducibility should be extensively evaluated for each analyte.[8,52] Studies should include assessment of reproducibility between several laboratories (including at least two external sites), operators, instruments, array lots, reagent batches, scans, days etc, as appropriate. Values may be assessed and reported in a number of ways — that is, variance component analyses[58] may be relevant in identifying major sources of variation in the values observed.

In addition, the accuracy of the device should be established. A way to accomplish this may be to compare the results of the device with the results for the same samples using a reference method (ie 'gold standard' test). For genotyping tests, double-stranded bi-directional DNA sequencing is the one well-recognised method for establishing performance; however, depending on the clinical claims being made — for example, diagnosing a specific form of cancer, evaluating risk for future development of disease and/or determining drug selection or dosing — additional clinical information, such as biopsy results, imaging data and/or results of blood tests, may be required to confirm the true diagnosis. All analytical parameters (eg fluorescence intensity, subset of alleles) should be challenged during testing to substantiate the device performance at key decision limits.

It is important that studies be carried out with real clinical samples (either prospectively collected or appropriately stored) whenever possible, to evaluate any biological and matrix effects that may occur and to achieve greater assurance that there are no unforeseen endogenous interferences that may affect the analytical performance of the device. Analytical performance of a diagnostic test should be based on a dataset that is independent from and prior to the prospective or retrospective samples on which it is to be clinically validated.

Clinical performance

While analytical validation presents its own challenges, the question of the clinical validation[59] of microarray-based tests is likely to prove to be an even larger issue for many assays and manufacturers to confront.

A diagnostic device has clinical utility when it provides information about a disease or clinical condition that is useful and/or meaningful to the healthcare provider and the patient. Different types of microarray tests may require different types of clinical validation. For some polymorphisms and genetic alleles, there may be a sufficient literature base to establish the clinical utility of the new test without extensive clinical studies. The FDA recently approved the Roche AmpliChip for Cytochrome P450[60] and the TM Bioscience Cystic Fibrosis assay[61] using analytical studies and clinical literature. Many microarray-based devices that are in development, however, may need clinical studies to establish the safety and effectiveness of their use in a clinical setting. If certain biomarkers have been extensively studied biologically or clinically, and have an underlying biological rationale, their validation may be easier because of support from existing scientific data. Biomarkers or classifiers without an established biological/mechanistic relationship to the disease, however, can also have a utility as a potential diagnostic tool. Clinical validation of biomarkers with no known biological mechanism related to the specific disease or drug action can be established using solid statistical methods and robust clinical trial data.[30]

A major hurdle for establishing clinical performance of the test is obtaining and securing adequate numbers of quality specimens from clinical trial subjects that can be used in support of the clinical utility of the device for its intended use. Regardless of whether a clinical trial is performed using prospective samples or banked samples collected in a prospective manner, the trial should be hypothesis driven and test a pre-specified and well-defined diagnostic biomarker/classifier claim (ie all cut-off points, platforms etc should be established in advance of the trial). If using samples that were

stored from a previously conducted clinical trial, clinical study specimens should be banked without selection bias, and under optimal storage conditions, to enable valid interpretation of results.

Similarly to analytical validation, the complexity of the test can influence the types of studies that are needed. Any patterns or combinations of analytes should be defined prior to the start of the clinical studies to allow for verification of the pattern as a unit. In addition, the test population should be clearly defined and considered to support the intended use in the intended population.

Quality control standardisation efforts for microarrays

As described above, several types of factors affect the outcome of a microarray study; a major concern is to assure reproducibility and accuracy. Intra- and inter-laboratory data consistency is the foundation of reliable knowledge extraction and meaningful comparisons. The microarray community and regulatory agencies are working together to establish a set of consensus quality assurance and quality control criteria for assessing and ensuring data quality to identify critical factors affecting data quality and to optimise and standardise microarray procedures so that biological interpretation and decision making are not based on unreliable data.[13] Developing standards and quality measures should enable the successful and reliable use of microarray technology in regulatory decision making and in clinical practice.

Microarray controls generally can be divided into internal and external. Internal (intrinsic) sample controls, which control for the sample integrity, can be a gene or a group of genes intrinsic to the sample, chosen because they are known not to vary in the particular condition. A major obstacle to using this type of microarray control is finding genes that would not vary in studied conditions. External controls are generally developed as spike-ins that are not part of the sample and can be used to characterise the performance of certain steps in the microarray experiment. Here, we list several current large quality control and standardisation efforts in the microarray community.

External RNA Controls Consortium (ERCC)

The ERCC is working to develop tools for experiment control and performance evaluation for gene expression analysis. These tools will include spike-in controls, protocols and informatics tools intended to be useful for various microarray platforms and QRT-PCR.[55] These platform-independent control materials are needed for performance evaluation of reproducibility, sensitivity and robustness in gene expression analysis. They will be readily accessible and used as a tool for verification of technical performance of a microarray assay, but not for sample integrity (they are not intrinsic controls). There are plans for the controls to be made commercially available as individual plasmid clones for the synthesis of polyadenylated transcripts.[56]

Microarray Quality Control (MAQC)

The purpose of the MAQC project is to provide quality control tools to the microarray community and to develop guidelines for microarray data analysis by providing the public with large reference datasets along with readily accessible reference RNA samples.[57] The MAQC project involves six FDA centres, major providers of microarray platforms and RNA samples, the Environmental Protection Agency (EPA), NIST, academic laboratories and other stakeholders. The MAQC project aims to establish quality control metrics and thresholds for objectively assessing the performance achievable by various microarray platforms and evaluating the advantages and disadvantages of various data analysis methods.[13,57]

Two human RNA samples have been selected and the differential gene expression levels between the two samples have been calibrated with both microarrays and QRT-PCR. The resulting microarray datasets will be used to assess to the precision and cross-platform/laboratory comparability of microarrays. The large QRT-PCR datasets will also enable evaluation of the nature and magnitude of any systematic biases that may exist between microarrays and QRT-PCR. The availability of the calibrated RNA samples combined with the resulting microarray and QRT-PCR datasets, which will be made readily accessible to the microarray community, will allow individual laboratories more easily to identify and correct procedural failures. The MAQC project aims to help improve microarray technology and foster its proper applications in discovery, development and review of FDA-regulated products.

Discussion

The development of DNA microarray technology for use in both exploratory studies and as a potential medical diagnostic tool has sparked excitement in the scientific and medical communities. Microarrays have the potential to simultaneously detect multiple DNA sequence variations, monitor relative expression levels of thousands of genes in selected tissues and identify infectious disease organisms or the host response to infection in biological specimens. When developed as medical IVDs, the results obtained by these assays may complement or be used in combination with other diagnostic methods and potentially be utilised for patient management and treatment decisions.

Microarrays promise to be a powerful technology when used correctly, but can be misleading without acceptable reproducibility and accuracy. The inherent strength of microarrays — their ability to query multiple analytes in a single assay, thereby providing results of increased complexity

compared with more traditional assays — also necessitates caution in interpreting the results for medical decision making. Performance data for use with specific applications, however, is often not sufficiently established scientifically to be suitable for regulatory decision making.

As noted in this paper, an important issue in the microarray field is platform-to-platform variability, since there may be inconsistencies between different platforms on a gene-by-gene basis. If a manufacturer develops a specific IVD involving a single microarray platform, however, and can demonstrate analytical and clinical performance of this diagnostic microarray, the issue of the inter-platform variability may not be crucial. To be able to use such platforms clinically, the variability of that specific microarray platform should be thoroughly assessed and proven to be robust and operating consistently and reproducibly across laboratories and over time. The appropriate evaluation should be performed to establish analytical and clinical performance characteristics of the IVD test, including whether the test accurately and reproducibly detects the gene markers it claims to detect for its intended use in clinically relevant populations.

The FDA recognises the potential of microarray-based tests for use in clinical practice. We have collaborated in a series of public meetings[4–6,10] to obtain input on relevant issues from the scientific community and interested stakeholders. Several guidance documents have been, or are being, developed that provide information on the FDA's current thinking on using new technologies in clinical decision making.[7–9,52–54] We are working with the whole microarray community towards streamlining this exciting technology and developing or assisting in the development of standards to help to validate the performance and expedite transfer of these devices from the research to the clinical use setting.

Disclaimer

This paper represents the current opinion of the Office of *In Vitro* Diagnostic Device Evaluation and Safety (OIVD), Office of Science and Engineering Laboratories (OSEL) and Office of Surveillance and Biometrics (OSB), in the Center for Devices and Regulatory Health (CDRH) within the US Food and Drug Administration (FDA).

Acknowledgments

We thank Dr Maria Chan and Roxanne Shively for helpful discussions. We also appreciate helpful comments from our colleagues in the Genomics/genetics group at the CDRH/FDA.

References

1. Food, Drug, and Cosmetic Act, 21 U.S.C. 321(h); definition of the term 'device' ('an instrument, apparatus, implement, machine, contrivance, implant, *in vitro* reagent, or other similar or related article, including any component, part or accessory, which is [2] intended for use in the diagnosis of disease or other conditions, or in the cure, mitigation, treatment or prevention of disease, in man or other animals').
2. 21 CFR §809.3(a); definition of *in vitro* diagnostic devices ('*In vitro* diagnostic products are those reagents, instruments, and systems intended for use in the diagnosis of disease or other conditions, including a determination of the state of health, in order to cure, mitigate, treat, or prevent disease or its sequelae. Such products are intended for use in the collection, preparation, and examination of specimens taken from the human body. These products are devices as defined in section 201(h) of the Federal Food, Drug, and Cosmetic Act (the act), and may also be biological products subject to section 351 of the Public Health Service Act.'), available at http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm, FDA (cited 1st April, 2004).
3. FDA (2004), FDA White Paper: 'Innovation or stagnation, challenge and opportunity on the critical path to new medical products', available at http://www.fda.gov/oc/initiatives/criticalpath/whitepaper.html. Updated March 29th, 2004; accessed 1st September, 2005.
4. 'Pharmacogenomics in drug development and regulatory decision making' — Workshop 3: 'Three years of promise, proposals, and progress on optimizing the benefit/risk of medicines', 11–13th April, 2005, Bethesda, MD.
5. FDA/Drug Information Association (2004), 'Co-development of drug, biological and device products', 29th July, 2004, Arlington, VA, available at http://www.diahome.org/Content/Events/04040.pdf. Accessed 1st September, 2005.
6. 'Drug metabolizing enzymes and pharmacogenomics' — a joint FDA/Johns Hopkins University/PhRMA educational workshop, 13–14th September, 2004, Rockville, MD.
7. FDA (2005), 'Guidance for industry: Pharmacogenomic data submissions', available at http://www.fda.gov/cber/gdlns/pharmdtasub.pdf. Posted 31st March, 2005; accessed 1st September, 2005.
8. FDA (2003), 'Multiplex tests for heritable DNA markers, mutations and expression patterns: Draft guidance for industry and FDA reviewers', March, available at http://www.fda.gov/cdrh/oivd/guidance/1210.html. Updated 21st April, 2003; accessed 1st September, 2005.
9. FDA (2005), FDA White Paper: 'Drug-device co-development concept paper (draft) April', available at http://www.fda.gov/cder/genomics/pharmacoconceptfn.pdf. Posted 8th April, 2005, accessed 1st September, 2005.
10. 'Microarrays in transcriptional profiling' — a joint FDA/Johns Hopkins University/PhRMA workshop, 20–21st July, 2005, Rockville, MD.
11. Dave, S.S., Wright, G., Tan, B. *et al.* (2004), 'Prediction of survival in follicular lymphoma based on molecular features of tumor-infiltrating immune cells', *N. Engl. J. Med.* Vol. 351, pp. 2159–2169.
12. Staudt, L.M. (2003), 'Molecular diagnosis of the hematologic cancers', *N. Engl. J. Med.* Vol. 348, pp. 1777–1785.
13. Shi, L., Tong, W., Goodsaid, F. *et al.* (2004), 'QA/QC: Challenges and pitfalls facing the microarray community and regulatory agencies', *Expert Rev. Mol. Diagn.* Vol. 4, pp. 761–777.
14. Bakay, M., Chen, Y.W., Borup, R. *et al.* (2002), 'Sources of variability and effect of experimental approach on expression profiling data interpretation', *BMC Bioinformatics [Internet]* Vol. 3, p. 4, available at http://www.biomedcentral.com/1471-2105/3/4. Published 31st January, 2002; accessed 1st September, 2005.
15. Bammler, T., Beyer, R.P., Bhattacharya, S. *et al.* (Members of the Toxicogenomics Research Consortium) (2005), 'Standardizing global gene expression analysis between laboratories and across platforms', *Nat. Methods* Vol. 2, pp. 351–356.
16. Irizarry, R.A., Warren, D., Spencer, F. *et al.* (2005), 'Multiple-laboratory comparison of microarray platforms', *Nat. Methods* Vol. 2, pp. 345–350.
17. Larkin, J.E., Frank, B.C., Gavras, H. *et al.* (2005), 'Independence and reproducibility across microarray platforms', *Nat. Methods* Vol. 2, pp. 337–344.
18. Thompson, K. (2005), 'Interrogation of microarray platform variation and performance with a mixed tissue RNA reference material', at 'Microarrays

in transcriptional profiling' — 'a joint FDA/Johns Hopkins University/ PhRMA workshop, 20– 21st July, Rockville, MD.
19. Petersen, D., Chandramouli, G.V.R., Geoghegan, J. *et al.* (2005), 'Three microarray platforms: An analysis of their concordance in profiling gene expression', *BMC Genomics [Internet]* Vol. 6, p. 63, available at http://www.biomedcentral.com/1471-2164/6/63. Published 5th May, 2005; accessed 1st September, 2005.
20. Shi, L., Tong, W. and Fang, H. *et al.* (2005), 'Cross-platform comparability of microarray technology: Intra-platform consistency and appropriate data analysis procedures are essential', *BMC Bioinformatics [Internet]* Vol. 6(Suppl. 2), pp. S12, available at http://www.biomedcentral.com/1471-2105/6/S2/S12. Published 15th July, 2005, accessed 1st September, 2005.
21. The Tumor Analysis Best Practices Working Group (2004), 'Expression profiling-best practices for data generation and interpretation in clinical trials', *Nat. Rev. Genetics* Vol. 5, pp. 229–237.
22. Irizarry, R.A., Hobbs, B., Collin, F. *et al.* (2004), 'Exploration, normalization, and summaries of high density oligonucleotide array probe level data', *Biostatistics* Vol. 4, pp. 249–264.
23. Clinical and Laboratory Standards Institute (CLSI) (2005),'Diagnostic nucleic acid microarrays', proposed guideline, Document MM12-P.
24. Carter, S.L., Eklund, A.C., Mecham, B.H. *et al.* (2005), 'Redefinition of Affymetrix probe sets by sequence overlap with cDNA microarray probes reduces cross-platform inconsistencies in cancer-associated gene expression measurements', *BMC Bioinformatics [Internet]* Vol. 6, p. 107, available at http://www.biomedcentral.com/1471-2105/6/107. Published 25th April, 2005; accessed 1st September, 2005.
25. Ransohoff, D.F. (2005), 'Bias as a threat to the validity of cancer molecular-marker research', *Nat. Rev. Cancer* Vol. 5, pp. 142–149.
26. Ransohoff, D.F. (2004), 'Rules of evidence for cancer molecular-marker discovery and validation', *Nat. Rev. Cancer* Vol. 4, pp. 309–314.
27. Simon, R. (2005), 'Development and validation of therapeutically relevant multi-gene biomarker classifiers', *J. Natl. Cancer Inst.* Vol. 97, pp. 866–867.
28. Simon, R., Radmacher, M.D., Dobbin, K. and McShane, L.M. (2003), 'Pitfalls in the analysis of DNA microarray data: Class prediction methods', *J. Natl. Cancer Inst.* Vol. 95, pp. 14–18.
29. Simon, R. (2003), 'Diagnostic and prognostic prediction using gene expression profiles in high dimensional microarray data', *Br. J. Cancer* Vol. 89, pp. 1599–1604.
30. Trepicchio, W.L., Essayan, D., Hall, S.T. *et al.* (2006), 'Designing prospective clinical PG trials – Effective use of genomic biomarkers for use in clinical decision-making', *Pharmacogenomics J.*, In press.
31. Cutler, D.J., Zwick, M.E., Carrasquillo, M.M. *et al.* (2001), 'High-throughput variation detection and genotyping using microarrays', *Genome Res.* Vol. 11, pp. 1913–1925.
32. Syvanen, A.C. (2001), 'Accessing genetic variation: Genotyping single nucleotide polymorphisms', *Nature Rev. Genet.* Vol. 2, pp. 930–942.
33. Warrington, J.A., Shah, N.A., Chen, X. *et al.* (2002), 'New developments in high-throughput resequencing and variation detection using high density microarrays', *Hum. Mutat.* Vol. 19, pp. 402–409.
34. Wang, D., Coscoy, L., Zylberberg, M. *et al.* (2002), 'Microarray-based detection and genotyping of viral pathogens', *Proc. Natl. Acad. Sci. USA* Vol. 99, pp. 15687–15692.
35. Volokhov, D., Chizhikov, V., Chumakov, K. and Rasooly, A. (2003), 'Microarray-based identification of thermophilic *Campylobacter jejuni*, *C. coli*, *C. lari*, and *C. upsaliensis*', *J. Clin. Microbiol.* Vol. 41, pp. 4071–4080.
36. Zwick, M.E., Mcafee, F., Cutler, D.J. *et al.* (2005), 'Microarray-based resequencing of multiple *Bacillus anthracis* isolates', *Genome Biol. [Internet]* Vol. 6, pp. R10, available at http://genomebiology.com/2004/6/1/R10. Published 17th December, 2004; accessed 1st September, 2005.
37. Weinshilboum, R. and Wang, L. (2004), 'Pharmacogenomics: bench to bedside', *Nat. Rev. Drug Discov.* Vol. 3, pp. 739–748.
38. Lesko, L. and Woodcock, J. (2004), 'Translation of pharmacogenomics and pharmacogenetics: A regulatory perspective', *Nat. Rev. Drug Discov.* Vol. 3, pp. 763–770.
39. Harper, CC., Philip, R., Robinowitz, M. and Gutman, S.I (2005), 'FDA perspectives on pharmacogenetic testing', *Expert Rev. Mol. Diagn.* Vol. 5, pp. 643–648.
40. Sebat, J., Lakshmi, B., Troge, J. *et al.* (2004), 'Large-scale copy number polymorphism in the human genome', *Science* Vol. 305, pp. 525–528.
41. Iafrate, A.J., Feuk, L., Rivera, M.N. *et al.* (2004), 'Detection of large-scale variation in the human genome', *Nat. Genet.* Vol. 36, pp. 949–951.
42. Carter, N.P. (2004), 'As normal as normal can be?', *Nat. Genet.* Vol. 36, pp. 931–932.
43. Pinkel, D. and Albertson, D. (2005), 'Comparative genomic hybridization', *Annu. Rev. Genomics Hum. Genet.* Vol. 6, pp. 331–354.
44. Davies, J.J., Wilson, I.M. and Lam, W.L. (2005), 'Array CGH technologies and their applications to cancer genomes', *Chromosome Res.* Vol. 13, pp. 237–248.
45. Ishkanian, A.S., Malloff, C.A., Watson, S.K. *et al.* (2004), 'A tiling resolution DNA microarray with complete coverage of the human genome', *Nat. Genet.* Vol. 36, pp. 299–303.
46. Barrett, M.T., Scheffer, A., Ben-Dor, A. *et al.* (2004), 'Comparative genomic hybridization using oligonucleotide microarrays and total genomic DNA', *Proc. Natl. Acad. Sci. USA* Vol. 101, pp. 17765–17770.
47. Dhami, P., Coffey, A.J., Abbs, S. *et al.* (2005), 'Exon array CGH: detection of copy-number changes at the resolution of individual exons in the human genome', *Am. J. Hum. Genet.* Vol. 76, pp. 750–762.
48. Hsu, L., Self, S.G., Grove, D. *et al.* (2005), 'Denoising array-based comparative genomic hybridization data using wavelets', *Biostatistics* Vol. 6, pp. 211–226.
49. Mansfield, E., O'Leary, T.J. and Gutman, S.I. (2005), 'Food and Drug Administration regulation of *in vitro* diagnostic devices', *J. Mol. Diagn.* Vol. 7, pp. 2–7.
50. FDA (1998), New Section 513(f)(2) — 'Evaluation of automatic Class III designation, guidance for industry and CDRH staff', February 1998, available at http://www.fda.gov/cdrh/modact/classiii.html. Updated 28th September, 2004; accessed 1st September, 2005.
51. Altaie, S. (2003), 'The pre-IDE process', presentation at the IVD 510(k) Workshop, 22nd April, 2003, Rockville, MD, available at http://www.fda.gov/cdrh/oivd/presentations/042203-Altaie.html, FDA. Updated 20th May, 2003; accessed 1st September, 2005.
52. FDA (2005), Class II special controls guidance document: 'Drug metabolizing enzyme genotyping system — Guidance for industry and FDA staff', available at http://www.fda.gov/cdrh/oivd/guidance/1551.html. Posted 9th March, 2005; accessed 1st September, 2005.
53. FDA (2005), 'Instrumentation for clinical multiplex test systems: Class II special controls guidance document — Guidance for industry and FDA staff', available at http://www.fda.gov/cdrh/oivd/guidance/1546.pdf. Posted 9th March, 2005; accessed 1st September, 2005.
54. FDA (2005), 'CFTR gene mutation detection systems: Class II special controls guidance document — Guidance for industry and FDA staff', available at http://www.fda.gov/cdrh/oivd/guidance/1564.pdf. Posted 25th October, 2005.
55. NIST (2003), http://www.cstl.nist.gov/biotech/workshops/ERCC2003/.
56. FDA/NCTR (2005), 'Use of external RNA controls in gene expression assays', proposed guideline, Clinical and Laboratory Standards Institute (CLSI), Document MM16-P.
57. http://www.fda.gov/nctr/science/centers/toxicoinformatics/maqc/. Cited 29th August, 2005.
58. Searle, S.R., Casella, G. and McCulloch, C.E. (1992), Variance Components, Wiley, New York, NY, p. 501.
59. FDA (2004), 21 CFR §860.7. 'Determination of safety and effectiveness', available at http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm. Cited 1st April.
60. FDA (2005), Roche AmpliChip Cytochrome P450 Genotyping test and Affymetrix GeneChip Microarray Instrumentation System, available at http://www.fda.gov/cdrh/mda/docs/k042259.html. Cited 8th March.
61. FDA (2005), Tag-It™ Cystic Fibrosis Kit, available at http://www.fda.gov/cdrh/mda/docs/k043011.html. Cited 26th August.