Multi-stakeholder Collaborations & Biomarker Development and Implementation

By

Tingyu Liu

B.A. Neuroscience
Pomona College, 2010

SUBMITTED TO THE ENGINEERING SYSTEMS DIVISION IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

MASTER OF SCIENCE IN TECHNOLOGY AND POLICY
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2014

Signature of Author: **Signature redacted**

Engineering Systems Division
May 9, 2014

Certified by: **Signature redacted**

Kenneth A. Oye
Associate Professor of Political Science and Engineering Systems Division
Thesis Supervisor

Certified by: **Signature redacted**

Gigi Hirsch
Executive Director of the MIT Center for Biomedical Innovation
Thesis Supervisor

Certified by: **Signature redacted**

Magdalini Papadaki
Research Scientist of the MIT Center for Biomedical Innovation
Thesis Supervisor

Accepted by: **Signature redacted**

Dava Newman
Professor of Aeronautics and Astronautics and Engineering Systems Division
Director, Technology and Policy Program

Multi-stakeholder collaborations & biomarker development and implementation

By

Tingyu Liu

Submitted to the Engineering Systems Division on May 9, 2014 in Partial Fulfillment of the Requirements for the Degree of Master of Science in Technology and Policy

Abstract

The rise of genomic technologies and increasing interest in personalized medicine have triggered a renewed focus on the role of biomarkers in drug development and clinical care. While many stakeholders, including industry, regulatory agencies, patients, academia, and payers, are involved in biomarker development and use, these stakeholders reside in functional siloes that lead to fragmented outputs and efforts. Thus, although biomarker discovery is extensive, biomarker adoption in the R&D process and clinical care is slow, complex, and opaque. The gap between discovery and practice needs to be bridged by clear and consistent evidentiary standards, regulations and policies. Biomarker adoption requires a systems effort that includes active dialogue and understanding among all stakeholders. The level of discussions and resources (data, samples, expertise, financial, etc.) needed also makes this field particularly fertile for collaboration.

This thesis suggests that multi-stakeholder collaborations (MSCs) can address the challenges of biomarker adoption. It describes what has been achieved in existing prominent MSCs, extracts general lessons on their contributions, and provides recommendations on the use of MSCs in drug development. Specifically, MSCs have demonstrated impact across 4 key areas: 1) data generation and establishment of standards; 2) development of scientific processes for biomarker evaluation and use; 3) the structuring of innovative operational and organizational scaffolds for collaboration; and 4) delivery of vetted regulatory recommendations and robust biomarkers for use by drug developers and clinicians. Understanding the roles, evolution, and challenges of MSCs will be important in enhancing the capacity to collaborate to address the complex challenges for biomedical innovation and healthcare outcomes.

Thesis Supervisor: Dr. Kenneth A. Oye
Title: Associate Professor of Political Science and Engineering Systems Division

Thesis Supervisor: Gigi Hirsch
Title: Executive Director of the MIT Center for Biomedical Innovation

Thesis Supervisor: Magdalini Papadaki
Title: Research Scientist of the of the MIT Center for Biomedical Innovation

Table of Contents

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my advisors, Dr. Ken Oye, Dr. Gigi Hirsch, and Dr. Magda Papadaki, for their support and guidance throughout my research and my Master's studies. Special thanks goes to Dr. Papadaki, who patiently edited and re-edited my writing and always provided insightful comments. I would like to thank the members of the Center for Biomedical Innovation; my research would not have been possible without their uplifting words and banter.

I am also grateful for the subject matter experts who have kindly provided inputs and advice during the research and writing process.

I would like to thank my fellow Technology and Policy students for the community throughout my two years at MIT. I would especially like to thank the students who spent countless hours with me at libraries or cafes during the last few weeks before the thesis deadline. Finally, I would like to thank my loved ones, who have supported me fully throughout this entire process.

Executive Summary

Recent advances in genomic technologies have given rise to interest in new therapies that can develop safer, better drugs for patients in an affordable and efficient manner. In particular, the biomedical industry is devoting increased attention to the promises and potential of personalized medicine. Personalized medicine is the effort to prevent, diagnose, and treat disease with targeted therapies designed to be effective for individual patients based on genetic or other information.

The application of well-defined biomarkers is the cornerstone of personalized medicine. A biomarker is an indicator of biological processes or pharmacologic responses to drug therapy.[1] Identifying the presence, absence, or amount of a biomarker in a patient can predict disease susceptibility and guide tailored therapies by assessing individual benefit or risk towards a particular drug, or determining the optimal drug dose base on metabolism. Although genomic technologies have brought about a rapid increase of novel biomarkers, developing accepted terminologies and scientific standards, as well as thoroughly evaluating the biomarkers, is a slow, demanding and complex process that comes with a variety of challenges. The novelty that accompanies many genomics biomarkers presents a crucial regulatory challenge. Regulations around biomarkers and other cutting edge R&D innovations, are unclear and controversial, which both hinders their adoption and may lead to biomarker misuse. Reimbursement and pricing are also risky, since payers are still unclear and inconsistent on evidence requirements for coverage decisions.

Most importantly, biomarker adoption in pharmaceutical development and clinical practice is a systems effort that necessitates active dialogue and understanding among all the stakeholders, including industry, regulatory agencies, patients, academia, and payers, as well as consolidated outputs and efforts across the spectrum. The level of resources (data, samples, expertise, financial, etc.) needed also makes this field particularly fertile for collaboration. In this thesis, we argue that multi-stakeholder collaborations (MSCs) are useful tools to address the key challenges of biomarker adoption. Starting from the scientific challenges addressed by key MSCs, we expand to describe their approaches and achievements on novel processes and policies toward an increased systems capacity for therapeutic innovation through biomarker adoption.

We discuss the evolution of the first regulatory pathway, termed the Biomarker Qualification Process (BQP), to promote biomarker adoption in R&D. The development of the BQP was initiated by one-on-one discussions between industry and the U.S. FDA (Food and Drug Administration) on how to best integrate pharmacogenomics data into drug submissions. These discussions led to the formation of key MSCs that mediated data-sharing and analysis agreements and resulted in the qualification of several biomarkers for preclinical safety testing and clinical trial selection in Alzheimer's disease by both the FDA and the EMA (European Medicines Authority).

Thus far, MSCs involved in the BQP have mainly focused on early-stage R&D biomarkers. Other MSCs are focusing on biomarkers involved in later-stage drug R&D and patient care, including the development of in vitro diagnostics (IVDs). Significant achievements of the Biomarkers Consortium and the Early Detection Research Network (EDRN) include, respectively, a novel adaptive biomarker-driven screening trial for therapeutics and a rigorous process for evaluating biomarkers for early detection.

To enhance the understanding of collaborations in biomarker adoption, we developed a framework to distill key lessons and analyze the contributions of these key MSCs across four areas that involve important translational enablers for biomarker adoption, particularly data and standards, operational elements around novel scientific processes, key organizational aspects of collaborative

---

[1] Note that biomarkers can be used as a *tool* to help drug developers during the preclinical and clinical stages of R&D. A biomarker can also be used as a *product*, in which case it is a medical device termed an In Vitro Diagnostic (IVD). An IVD can be independent of a drug or it can be drug-specific. IVDs that must be paired with a specific drug because detecting and measuring this biomarker is essential for the clinical use of the drug is more precisely termed a companion diagnostic (CDx).

innovation, as well as the prominent tangible outputs around novel biomarkers and relevant regulatory guidelines. In specific:

1) **Data and data standards**. Data sharing allows MSCs to tackle projects that would not otherwise be feasible. Moreover, broad-based data harmonization can set important standards to increase consistency in cross-stakeholder interpretations, as well as subsequent stages of data deployment and decision-making such as regulatory submissions, the design of novel clinical trials, or drug labelling and use.

2) **Scientific processes**. The scientific processes developed by MSCs are essential to break down large, complex challenges into manageable workstreams. MSCs have been pivotal in the establishment of new pathways for biomarker development and use across R&D stages and disease areas. For example, the MSCs involved in the BQP are collaborating to move qualified preclinical biomarkers into the clinical stage. In addition, the I-SPY 2 trial for early-stage breast cancer has led to new adaptive trials for lung cancer and Alzheimer's disease. Such follow-on work will increase the predictive capacity and therapeutic impact of biomarkers.

3) **Collaborative frameworks**. Biomarker development is a complex field that requires an increasing degree of knowledge communication and integration across stakeholder groups and R&D stages. Learnings from the MSCs analyzed in this thesis depict important developments in the operational and structural scaffolds of collaborative innovation, allowing the design and execution of the scientific processes discussed above and their ultimate deliverables. MSCs have continued to evolve in scope and breadth through strategic partnerships and interactions; their follow-on work will create progressive new value generation from biomarker and related translational outputs. Most notably, MSCs have increased the level of understanding and alignment on important strategic directions and priorities, within and across the different stakeholder groups.

4) **Concrete deliverables**. A number of robust biomarkers and regulatory guidance documents that outline innovative policies for biomarker adoption have already emerged from the work of MSCs. A few of the prominent examples of progress include nine qualified nephrotoxicity biomarkers, two qualified biomarkers for Alzheimer's disease, and five FDA-approved diagnostics for early-stage cancer detection. Guidance documents include a regulatory pathway for qualifying biomarkers (the BQP) and methods to perform adaptive clinical trials.

MSCs have already contributed to innovative approaches to tackle the challenges of biomarker adoption in drug R&D and clinical care. Our distilled learnings about their contributions have led to a number of key recommendations that will be important as MSCs continue to accelerate biomarker innovation. The recommendations can be broadly summarized as:

1) **Avoid fragmentation and duplication of pre-existing efforts.** MSCs have broken down complex biomarker research challenges and brought together different stakeholders. As MSCs continue to evolve, it is imperative to pursue a higher level of collaboration by bringing together different MSCs to share key knowledge and processes. Such an effort will insure that MSCs are aware of and can leverage each other's contributions across the four areas described in the above framework.

2) **Keep scientific processes flexible and adaptive.** MSCs have already established key scientific processes for biomarker development and use. It is essential to keep the processes adaptive, with the ability to expand their scope as more data and information becomes available. An important enabler in this process will be regulatory oversight on the ground of flexible guidance documents that can readily incorporate new learnings in the field.

3) **Progressively include stakeholders.** MSCs in the biomarker field have mainly included regulators, industry, and academia thus far. However, our study showed that other stakeholder groups, including diagnostics companies, payers, and patients, need to be more

systematically included in MSCs because successful biomarker development and use requires their contributions and insights as well. Depending on the goals of an MSC, regulators and industry may join at different times. For example, an MSC that focuses on scientific research as opposed to commercial outputs may choose to engage industry later in the biomarker development process.

4) **Develop metrics to track progress and measure impact.** There is currently a lack of meaningful metrics and performance measurement approaches for collaborative efforts. A key barrier that is hindering metrics development in the biomarker field is the confidentiality provisions as biomarkers move back into the commercial space. MSCs should discuss ways to remove such barriers since meaningful evaluation is essential to set goals, identify gaps, and optimize collaborations.

The biomedical industry has gradually shifted towards a more collaborative approach to tackle challenges that exceed individual stakeholder capabilities and integrate the current fragmented paradigm in drug R&D. As science and technology become more diffuse and specialized, MSCs will be integral for successful therapeutic innovations. This thesis has charted the progress and contributions of key collaborations in the biomarker field to elucidate key learnings and future recommendations. Such analysis will help MSCs evolve into established operational models for innovation in biomedicine and healthcare.

Section 1: Introducing biomarkers as a key tool in drug discovery/development & patient care

Recent advances in biotechnology, prompted by the sequencing of the human genome, have resulted in a new array of therapeutics. Indeed, the rise of the so-called Omics Revolution (genomics, proteomics, pharmacogenomics) and the –ics technologies (informatics, bioinformatics, biostatistics) have expanded the ways in which disease is detected and treated. There are now over 30,000 primary genes and over 100,000 proteins and many of their protein and metabolic products serve as promising new molecular targets (Day et al. 2009).

To turn these possible targets into real treatments requires a complex, systemic and highly regulated process. In the United States, the main authority that regulates drugs is the Food and Drug Administration (FDA), and especially the Center for Drug Evaluation and Research (CDER). In the European Union, the FDA counterpart is the European Medicines Agency (EMA). Pharmaceutical R&D has traditionally followed four main stages (Figure 1):
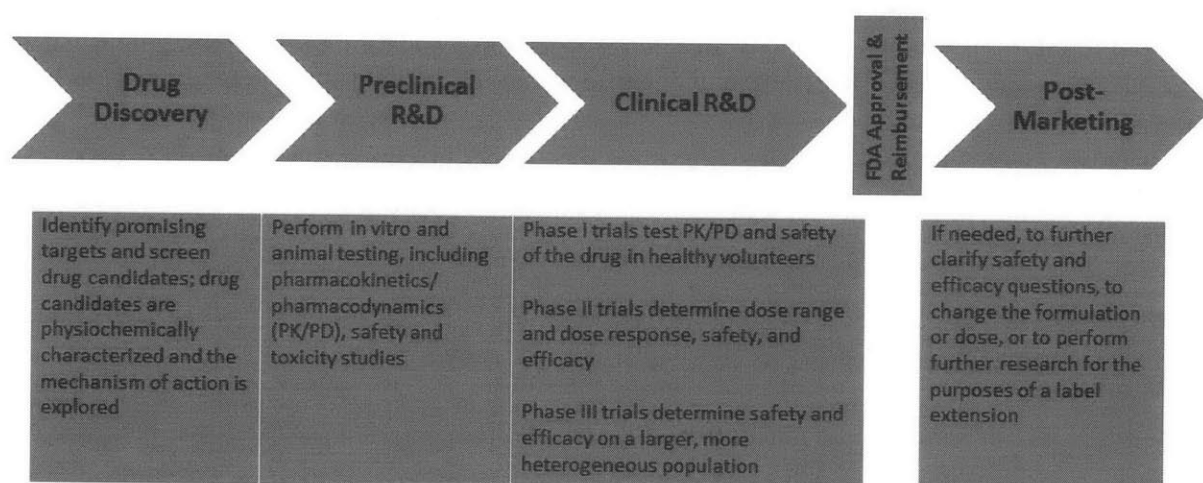

Drug Discovery

Identify promising targets and screen drug candidates; drug candidates are physiochemically characterized and the mechanism of action is explored

Preclinical R&D

Perform in vitro and animal testing, including pharmacokinetics/ phamacodynamics (PK/PD), safety and toxicity studies

Clinical R&D

Phase I trials test PK/PD and safety of the drug in healthy volunteers

Phase II trials determine dose range and dose response, safety, and efficacy

Phase III trials determine safety and efficacy on a larger, more heterogeneous population

FDA Approval & Reimbursement

Post-Marketing

If needed, to further clarify safety and efficacy questions, to change the formulation or dose, or to perform further research for the purposes of a label extension

Figure 1. Stages of pharmaceutical R&D.

On average, this process takes more than ten years, costs more than a billion dollars, and only one in ten thousand compounds actually succeed to become a marketed drug (Paul et al. 2010). Many experts have studied R&D productivity and success to propose solutions on improving the process (Pammolli et al. 2011).

The capacity of the healthcare system as a whole to provide newer, better drugs to patients in an affordable, reliable, sustainable, and efficient manner has been hindered by a series of bottlenecks in drug R&D and patient care (Table 1). The challenges are already apparent in the discovery stage; the proliferation of new molecular targets due to the genomics revolution is a key example. These targets are frequently so novel that there is insufficient data and/or a lack of harmonized data standards to understand and further assess the target in question. Another bottleneck is the high attrition rate of late-stage clinical trials (Carroll 2012); there is a lack of robust safety or efficacy profiling because of an imperfect understanding of the disease and the science, including the drug target and drug design. This has been particularly evident in the field of oncology; drug developers have struggled to obtain significant clinical results when the heterogeneity of the patients' genotypes and clinical presentation increase in later-stage clinical trials (Gerard et al. 2012). On the other hand, a drug may fail to reach a patient because of reimbursement requirements by payers which have become increasingly uncertain and stringent, especially in the light of the Affordable Care Act (Trusheim & Berndt 2012) and recent global fiscal pressures. A new drug may have very limited value for patients if insurance companies

refuse to cover the drug because they believe that it is not sufficiently superior to existing therapies or its cost-benefit analysis is not sufficiently robust.

Table 1. The healthcare value chain for drugs and devices has been hindered by a series of bottlenecks.

Healthcare value chain	Bottlenecks
Drug/device discovery	• Lack of clear criteria to prioritize targets • Lack of harmonized data standards
Drug/device development	• Lack of good animal models for diseases • Unclear safety and toxicity profile of drugs • Inability to demonstrate safety and efficacy • Inability to scale-up drugs/devices for large-scale manufacturing • Different decision criteria among stakeholder groups • Heterogeneity of patients leads to clinical trial failures
Patient care	• Safety and efficacy can only be optimized for patient sub-populations • Limited patient access due to high prices and/or lack of reimbursement • Unknown cost-effectiveness evidence compared to standard of care or for novel treatments
General	• Long timelines • High development costs • Low probability of success • Lack of communication & fragmented actions between the different stakeholder groups

The formal definition of biomarkers commonly used today was established by the Biomarkers Definitions Working Group's workshop in 2001. By this definition, a biomarker is "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic response to a therapeutic intervention" (Atkinson et al. 2001). Moreover, a "characteristic" is any "biological molecule found in blood, other body fluids, or tissues that is a sign of normal or abnormal process, or of a condition or disease and may be used to see how well the body responds to a treatment for a disease or condition." The potential value of the application of biomarkers in clinical care has been known for thousands of years. For example, in 4th to 5th century BCE, ancient clinicians detected glucose in urine by observing that urine from certain individuals attracted ants, thus providing a clinical description of diabetes. However, in the past twenty years, the term "biomarker" has come into focus and popularity, especially after the sequencing of the human genome (Vaidya & Bonventre 2010). Biomarkers are a key technology to address the bottlenecks which are preventing patients from obtaining innovative therapies and especially personalized medicines.

Biomarkers serve as the foundation of personalized medicine, an emerging healthcare approach that recognizes the importance of individual differences in treating disease. It is beginning to be understood that there are multiple subgroups and stages for each disease, and traditional drug development is not fully considering the development of targeted drugs tailored to a specific population subgroup. Personalized medicine aims to prevent, diagnose, and treat disease based on a thorough understanding of the genetic basis of the disease and its clinical presentation in individual patients. Identifying the presence, absence, or amount of a biomarker in a patient can predict disease

susceptibility and guide tailored therapies by assessing individual benefit or risk towards a particular drug, or determining the optimal drug dose base on metabolism.

A classic example in the field of personalized medicine is trastuzumab (Herceptin, developed by Genentech). Herceptin is a molecularly targeted cancer therapy that is designed to "shut off" the HER2 gene. The HER2 gene is responsible for the overexpression of the HER2 protein in approximately 20-25% of breast cancer cases and results in a very aggressive form of the disease. Thus, a breast cancer patient is first tested for HER2 using an *in vitro* diagnostic (IVD). The patient can only use Herceptin after obtaining a positive test result because the drug is only beneficial for patients with this genotype (FDA 2013b).

In this thesis, we have defined biomarkers to assume three main roles (Table 2).[2] Biomarkers used in R&D are typically basic assays and tests to help drug developers understand their drug and its effects in both the preclinical and clinical stages. However, sometimes a biomarker used in R&D can be further developed as a companion diagnostic (CDx), a subset of IVDs. This is a biomarker that is matched to the drug because detecting and measuring this biomarker is *essential* for the clinical use of the drug, as the Herceptin example illustrates. Note that a CDx does not *have* to be developed during drug R&D. IVDs which are not CDx are used independently of a drug (see examples below, Table 2).

Table 2. Examples of the different uses of biomarkers.

Biomarker uses	Examples
During the research & development phases for a drug	1) Identify targets for therapy, e.g. circulating mRNA as signaling molecules 2) Demonstrate drug safety, e.g. by measuring upregulation of KIM-1 biomarker to predict toxic kidney injury in animal models
Companion Diagnostic (an IVD that is *essential* for the safe, efficacious use of a drug)	1) Determine breast-cancer patients' therapy based on HER2 profile for Herceptin therapy 2) Determine melanoma patients' therapy based on presence of V600E or V600K mutation in BRAF gene for Tafinlar and Mekinist
Prognostic/Diagnostic device (non-drug-specific IVDs)	1) Oncotype DX – 21-gene panel test that predicts likelihood of cancer recurrence in women who are estrogen receptor (ER) positive 2) CertNDx – diagnostic test for bladder and upper urinary tract cancers by detecting FGFR3 mutation in urine 3) 23andMe – direct-to-consumer (DTC) test that assesses inherited traits, genealogy and possible congenital risk factors

The value that biomarkers bring to the R&D and, correspondingly, to patient care is evident (Table 3). Biomarkers have the potential to overcome the bottleneck examples in Table 1 above and thus significantly reduce the costs and uncertainty associated with drug development. Biomarkers are useful from the earliest stages of drug development, to identify drug targets and establish how the potential drug interacts. Biomarkers help elucidate how the drug behaves in animal models and human patients, and eventually define the target population. As mentioned previously, if a drug was elucidated

---

[2] Note that while these roles are commonly understood in the biomarker field, biomarkers are not typically thought of and categorized in such a manner because experts often only discuss one aspect of a biomarker's use, as opposed to all three.

with the use of a biomarker, then the biomarker can become a CDx. Moreover, biomarkers are useful to detect disease and understand a patient's profile as an IVD.

Table 3. Value that biomarkers bring to R&D and patient care.

		Healthcare value chain	Biomarker Impact
Diagnostic/Prognostic Devices	Companion Diagnostics	Biomarkers in R&D	**Drug discovery**	• Elucidate drug targets through mechanism of disease studies • ID & screen biomarkers for use as a target for intervention • Establish structure-activity-relationship (SAR)
			Preclinical	• Build PK/PD models • Elucidate mechanism of action of drugs • Establish safety & efficacy endpoints for animal models
			Clinical	• Select clinical trial subjects • Establish bioequivalence • Establish dose response & optimization • Elucidate mechanism of action of drugs in humans • Establish safety & efficacy endpoints for patients • Define initial target population
			Patient Care	• Stratify patients based on risk of disease • Screen, diagnose, and monitor patients • Monitor drug response & side effects
			Population Care	• Track public health status • Inform recommendations and policies for preventing, mitigating, and treating diseases

Biomarkers deliver value to each stakeholder across drug R&D stages and clinical practice to help experts make more informed decisions that will better apprise a patient's clinical profile and/or deliver safer, more effective drugs to patients (Table 4). Pharmaceutical companies use biomarkers to help demonstrate the safety and efficacy of new drugs. Some, such as Roche and Novartis, are building dedicated in-house diagnostics units while others such as AstraZeneca are forming partnerships with external developers (Opar 2011). Diagnostic companies can partner with pharmaceutical companies or launch diagnostic devices by themselves. Thus, it is of no surprise that the global diagnostics market is $42 billion (with $16 billion in the US in 2011) and over 6.8 billion diagnostic tests are performed each year (Advameddx 2013). Regulatory agencies have also prioritized biomarker research and development to better understand how to evaluate biomarker use and to promote biomedical productivity (Woodcock & Woosley 2008). Payers and providers want to understand if patients are being treated in a way that takes into account their individual differences, to maximize efficacy of cures and save both money and lives.

Table 4. Biomarkers deliver value to each stakeholder in the healthcare value chain.

Stakeholder	Biomarker value delivery
Industry	• Improve predictions on safety and efficacy, and patient effectiveness • Obtain reimbursement approval • Optimize standard of care • Inform drug labeling
Regulators	• Endorse new technologies • Understand how to evaluate new technologies
Providers	• Improve understanding of patient and disease profiles • Understand individual genetic profiles • Prescribe more targeted drugs for patients
Payers	• Support outcomes-relevant reimbursement decisions • Help determine which drugs are the most cost-effective for which patients
Patients	• Foster access to new drugs and/or targeted therapies • Reduce number of adverse responses to drugs

Still, biomarkers uptake comes with many challenges. This thesis discusses the challenges as well as the opportunities of biomarker integration into drug development and patient care. We then argue that multi-stakeholder collaborations (MSCs) can be seen as a useful instrument to address the complex and costly challenges that are currently constraining productive use of biomarkers. Successful biomarker adoption requires a systems-level change that involves and impacts all stakeholders (Barratt et al. 2012; Stephenson et al. 2013; Chataway et al. 2012; Goldman 2012; Wagner et al. 2007; Campion et al. 2013). The term "MSCs" (as distinct from "consortia" or "collaborations" in general) is defined as an environment that convenes two or more different stakeholder groups, such as the ones listed in Table 4. Abstracting from our research in the biomarker field, we make a case for the role of MSCs beyond scientific outcomes in the design and implementation of effective policies which will allow all stakeholders to incorporate emerging technologies and science innovation in drug discovery and healthcare in more sustainable and strategic ways.

Section 2: Biomarkers in drug development and patient care: background and critical challenges

The challenges to bring biomarkers into R&D and patient care span from nomenclature to policy; the details are complex and it is difficult for invested stakeholders, even field experts, to understand the whole picture and make informed decisions. In this section, the challenges are broken down into three aspects: the confusion around terminology/nomenclature, the complexity of evidentiary standards, and controversies in regulations and policies. It is important to note that each of these three aspects feeds into the other, and they are not mutually exclusive. This section ends with a summary of these challenges in the context of a biomarker's role in pharmaceutical R&D, as a CDx, or as a diagnostic/prognostic device.

Section 2.1: Terminology/Nomenclature

The definition of a biomarker, as discussed in Section 1[3], although detailed in content, is overly simplistic in concept, failing to illustrate the many different types and categorizations of biomarkers. The present study contains a Glossary that lists the most common terminologies used by various stakeholders to depict the current confusion in the biomarker landscape (Glossary A). The confusion is mainly due to the many different groups working in this field, each with their own agenda and points of focus.

It is important to note that for a biomarker to be used for any intended purpose, it needs to be measured with a corresponding device. The device is a way to assay or illustrate the state of the biomarker (Woodcock 2010).[4] Devices are regulated by the Center for Devices in Radiology and Health (CDRH), as opposed to CDER. One of the fundamental complexities in the field is that regulatory evaluation considers the biomarker separately from its measurement device, although in practice, using a biomarker necessitates an instrument or tool to measure it. The comprehensive definition of a biomarker with its corresponding measurement device highlights the complexity of the field and the reason why definitions and categorizations fail to capture all biomarkers. The term "biomarker" is vast, including any imaging technique (CT, MRI, PET, etc) and any clinical laboratory test, from immunochemistry, blood pressure, pain scales, and microbial cultures to ECGs. This situation is further complicated by the recent abundance of novel biomarker discoveries resulting from the Human Genome Project (Gutman & Kessler 2006).

The confusing nature of biomarker terminology extends to complicate the evaluation framework for assessing biomarkers. There is a common tendency to overuse or misuse the terms *validation/qualification*, as well as the terms *analytical validity, clinical validity*, and *clinical utility*. Glossary B provides examples of how different experts have defined the terms *validation* and *qualification*. Experts from different stakeholder groups tend to use the terms loosely and may have their preferred definition of the word since each group has its own areas of focus and decision-making,

---

[3] As discussed in Section 1, the formal definition of biomarkers as established by the Biomarkers Definitions Working Group is "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic response to a therapeutic intervention" (Atkinson et al. 2001). Moreover, a "characteristic" is any "biological molecule found in blood, other body fluids, or tissues that is a sign of normal or abnormal process, or of a condition or disease and may be used to see how well the body responds to a treatment for a disease or condition."

[4] A device is formally defined as "an instrument, apparatus, implement, machine, contrivance, implant, in vitro reagent, or other similar or related article, including any component, part, or accessory, which is...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...and which does not achieve its primary intended purposes through chemical action within or on the body of man or other animals and which is not dependent upon being metabolized for the achievement of its primary intended purposes" Section 201(h) of the Federal Food, Drug, and Cosmetic Act, 21 U.S.C 321(h).

languages, and norms, as well as different valuation of the data (EGAPP 2014). Other terms that have evolved into use seem to be derivations of these terms, including *analytical validation* or *analytical performance*, *clinical qualification*, *utilization* or *clinical utilization*.

Despite the overall confusion, seminal papers by industry authors and the FDA are in general agreement and have been widely cited and these terms will be used for the purposes of this thesis.[5] On this ground, after the initial biomarker discovery, there are three general stages of biomarker development:

1) *analytical validity*—validated method/assay performance;
2) *clinical validity*—validated link between the biomarker and biological/clinical endpoints;
3) *clinical utility*—the relevance and usefulness of the biomarker in patient care.

Section 2.2: Evidentiary Standards

Biomarker evaluation and acceptance is very slow compared to the speed of biomarker discovery (Bleavins 2010). There has been no shortage of novel biomarkers (Rinaldi 2011), but it is difficult to determine which ones have the greatest potential to be further developed for use in drug development and patient care. This difficulty arises from the fact that the key stakeholder groups work in functional silos of data and knowledge generation. Research labs, where most biomarker discovery and early development occurs, only have limited communication with and awareness of other labs. Moreover, industry and diagnostic companies are not aware of relevant research pursued at academic labs. Conferences, publications, and university technology transfer offices remain the key means of communication between stakeholders, although this is a slow and unreliable route. There is no systematic way of linking the stakeholders and promoting cross-stakeholder awareness of each other's work and needs. Thus, although academics may lead on important biomarker discoveries, they often do not understand the rigor needed to move a biomarker through the different stages of maturity.

To ensure that the biomarker has practical use beyond its initial discovery, the gap from discovery to practice needs to be bridged by consistent and well-accepted evidentiary standards that can be systematically applied to every novel biomarker within reason. Evidentiary standards are a set of parameters that underlie a biomarker's lifecycle as it steps through progressive stages of maturity, from analytical validity to clinical validity and clinical utility. A defined set of parameters does not yet exist, although field experts have agreed that certain key characteristics are important to keep in mind.

First, evidentiary standards depend on the intended use of a biomarker; a biomarker used in early stage R&D for an exploratory or observational study and/or for internal decision-making may not require the same evidentiary standards as a biomarker that is used for regulatory decision-making and/or for treatment decisions in clinical practice such as a CDx (Lesko 2007; Lesko & Atkinson 2001; Bleavins 2010). As the Biomarker Definitions Working Group stated, "robust linkage of a biomarker with a clinical endpoint is not essential in early development when the goal is confirmation of pharmacologic activity or optimization of dose regimen" (Atkinson et al. 2001). It is important to note that evidentiary standards are science- and context- specific, so they must be reevaluated if a biomarker is used in a different way, or in another medical context (Altar et al. 2008). Thus, finally, it has also been generally agreed by industry and academics alike that biomarkers should be reevaluated on a continual basis as more data is generated. This dynamic and resource-intensive nature of biomarker evolution is difficult to sustain. Multi-stakeholder collaborations are useful for such maintenance responsibilities because a single stakeholder cannot tackle such a large and multidisciplinary challenge.

---

[5] For a historical review of papers that detail the evidentiary standards of biomarkers, (IOM 2010a) gives special attention to nomenclature, systems of classification, and statistical methods developed for their evaluation.

Experts have been struggling with how to define the set of parameters for each evidentiary standard to bring biomarkers from the discovery laboratory to practical use in the healthcare value chain. There are no set definitions and standards, and no agreed-upon regulations that explicitly evaluate a biomarker as it progresses through subsequent stages of maturity (Khleif et al. 2010). Determining which biomarker possesses clinical validity or utility is agreed on a case-by-case basis. This is one of the key issues that results in the slow integration of biomarkers into mainstream patient management (Wilson et al. 2007). Below, the issues surrounding each stage of biomarker development are highlighted in greater detail.

Section 2.2.1: Analytical Validity

Analytical validation refers to the assessment of the biomarker's measurement device and its performance (Bleavins 2010). It has been said that analytical validity is "often an underappreciated component of biomarker evaluation" (Pennello 2013). Even the most basic assays, such as immunohistochemistry and fluorescence in-situ hybridization tests, can produce varying results depending on who is performing the assay, where the reagents and samples come from, and how the samples were prepared and stored (Gutman & Kessler 2006). This issue becomes magnified with novel, genomic biomarkers because technology is rapidly evolving and interpretations are not standardized or widely comprehensible (EGAPP 2014).

Another issue is inconsistent results across laboratories; laboratories may have different protocols for the same test or different quality systems, especially if they are not certified from the same auditing organizations (Wilson et al. 2007; Halim 2009; Khleif et al. 2010). Sample variability is yet another issue, as many errors and inconsistencies can occur in the pre-analytical phase and affect sample integrity and analysis suitability. For example, patients may have interfering medications or different nutritional diets, technicians may encounter organizational issues such as mislabeling a sample with the wrong ID, or one sample could be processed and stored differently from another (Drucker & Krapfenbauer 2013).

Many of these issues are due to a lack of industry agreement on validation standards and criteria in the field (IOM 2010a; Khleif et al. 2010) and the fact that much of the research concerning analytical validity is unpublished or proprietary, making it even more difficult to synthesize the data (EGAPP 2014).

Section 2.2.2: Clinical Validity

Table 3 (Section 1) defines the three main roles of a biomarker as a tool used in R&D, a CDx, or a drug-independent diagnostic device. Each biomarker use has its own set of evidentiary standards, and analytical method validation must be demonstrated for all of these roles. However, biomarkers used in the early stages of R&D, such as during the drug discovery phase, may not need to demonstrate clinical validity. To the contrary, biomarkers used in later stages of R&D, or as IVDs, should demonstrate clinical validity as well as clinical utility if possible, because the biomarker will be used in patient management.

Clinical validity assesses the link of analytical results to clinical and biological endpoints. An example of a clinically valid biomarker is the count of CD4+ immune system cells for HIV and AIDs. Clinical validity comprises a critical point of disconnect among the different stakeholder groups. A research lab typically performs the preliminary analytical validity testing, but they often do not have a grasp of what is needed in terms of sample size, test cross-validation, biospecimen quality, and the translation of the biomarker to a tangible product. Thus, their data will be of limited use downstream for industry and regulators. Indeed, one expert has stated, "Published academic studies...typically lack the scale, analytical stringency, and level of evidence needed to elicit industry confidence to invest in expensive clinical validation trials" (Poste 2012). Other stakeholder groups may not have a full understanding of the scene either. In many cases, especially with genomic technologies, the FDA does

not have a complete idea of what they are looking for, and drug developers are unclear on how to proceed (Khleif et al. 2010).

Meanwhile, clinical validity standards are being raised. Ideally, clinical validity standards would be determined by comprehensive retrospective studies and prospective clinical trials (Teutsch et al. 2009). However, such trials require significant resources and industry is hesitant to provide funding because it is unclear how regulators would evaluate the trials. Moreover, obtaining a comprehensive picture of the biomarker may be difficult, impossible, or unethical if it requires recruiting biomarker-negative participants for a prospective trial. For example, a biopsy has never been performed on a PSA (Prostate Specific Antigen)-negative individual for the purposes of completely validating the biomarker (Gutman & Kessler 2006). Also, errors from the analytical validity data can spill over and cause serious consequences for clinical validity studies if they remain unrealized or undetected (Baggerly & Coombes 2011). For example, in the case of a CDx, the drug development program can take ten years, making the use of different biomarker assay platforms or the switching of reagent vendors very likely (Halim 2011).

Section 2.2.3: Clinical Utility

Clinical utility assesses the quality and quantity of the generated data from the analytical and clinical validity studies, as well as its consistency and generalizability, to determine whether the biomarker is robust enough to be used in patient care for a specific context of use (COU) (Teutsch et al. 2009). However, generating sufficient, high-quality evidence is difficult due to resource constraints and uncertain market incentives (Parkinson et al. 2014). For example, at the National Cancer Institute (NCI), much of the financial resources are set for biomarker discovery, and less than 2% of the funding is focused on clinical utility (Schully et al. 2011). The need for both increased market incentives and sufficient evidence is a catch-22 problem: without sufficient evidence, there is no market incentive, and without an incentive, no evidence will be generated. Experts have stated that evidence for clinical utility should be identified by multiple stakeholders to enable efficient generation of the needed data (Zineh & Lesko 2009). Patient and public advocacy groups can also push for clinical acceptance of such biomarkers. However, without a formal patient inclusion process in place, they represent an "idiosyncratic process that depends on the emergence of leadership in a given disease area" (Mittleman et al. 2013).

There is still no general consensus on what clinical utility should involve. Some organizations, such as the National Institute for Health and Clinical Excellence (NICE) in the UK, and the CDC's Evaluation of Genomic Applications in Practice and Prevention (eGAPP) group have tried to define clinical utility (Poste et al. 2012). Some organizations believe that healthcare economics should be part of the evaluation process, which makes the concept of clinical utility more rigorous. Clinical utility can also be surprisingly elusive. Some of the most widely and extensively used biomarkers, like mammogram images and the PSA, are good case studies of this ambiguity. Although these biomarkers may be analytically and clinically valid, there is still no clear evidence that they lead to better health outcomes (Trusheim et al. 2013).

Section 2.3: Regulations and policies

When biomarkers become products that are used in patient care, they are termed IVDs and are regulated by the FDA as medical devices. This feeds into the challenges of biomarker adoption on the regulatory front, because the regulations around medical devices that measure biomarkers are very unclear and controversial, which may lead the biomarker to be misused in patient care.

There are two paths that a biomarker device can take. The first path goes through FDA's approval process for medical devices.[6] FDA-approved IVDs will demonstrate clinical validity (Kelloff & Sigman 2012), but will demonstrate clinical utility if it is an inherent part of clinical validity, as is the case with CDx, since its performance predicts patient outcomes (Parkinson et al. 2014). On the other hand, a device may not have to undergo FDA approval if it is considered a Laboratory Developed Test (LDT).[7] Both IVDs that are reviewed by the FDA or offered as an LDT are intended for clinical use and so need to be performed in laboratories which are inspected, certified, and accredited under the Clinical Laboratory Improvement Amendments (CLIA), administered through CMS (Parkinson et al. 2014).[8] However, CLIA focuses on technical performance and doesn't address clinical validity or utility (Trusheim et al. 2013). Thus, many experts regard LDTs as a "loophole" for industry to bypass the long and expensive FDA-approval pathway (Khleif et al. 2010).

LDTs can make the claim of clinical validity and utility and, with the right marketing forces, be used in many hospitals and doctor offices. Because many healthcare stakeholders work in functional siloes, doctors and hospitals may not understand the difference between an LDT and an FDA-approved test. LDTs can even mimic FDA-approved tests and erode the latter's market share (Hayes et al. 2013). Furthermore, a new type of LDTs, Direct-to-Consumer (DTC) tests such as 23andMe, pose additional problems, because most patients will not understand the scientific uncertainty and validity surrounding these tests well enough to make informed decisions, resulting in the possibilities of biomarker misuse (Kelloff & Sigman 2012).

However, LDTs are useful in a variety of cases. For example, the LDT pathway allows the evaluation of a cutting-edge technology before embarking on the costly and time-intensive FDA approval process (Hockett & Close 2010). In addition, sometimes patient groups may advocate for faster access for tests instead of waiting for well-designed clinical trials to demonstrate validity and utility, depending on the urgency of their disease (Teutsch et al. 2009).

The current proliferation of diagnostics has been frequently accompanied by high price demands, especially with biomarkers that use cutting-edge genomic technologies. In response, payer organizations are starting to demand demonstrates of clinical utility before reimbursing such devices. However, payers typically do not provide guidance as to what kind of data and how much data are

---

[6]The Medical Device Amendments to the Food, Drug, and Cosmetic Act in 1976 gave the FDA authority to regulate medical devices. The FDA established 3 classes for medical devices:
1) Class I have the least amount of risk and typically don't require premarket review or approval before they are used in patient care.
2) Class II have moderate risk and need to obtain 510(k) approval. 510(k) approval doesn't necessarily require clinical trials (Evans 2010), but does require the device manufacturer to show proof that the device is equivalent to a device already on the market.
3) Class III have the highest risk and have to undergo the PMA process which includes the submission of supporting clinical trial data (Weiss 2012). Note that in this case, an investigational device exemption (IDE) application may need to be filed. An IDE permits use of the device in a clinical investigation to evaluate the safety and/or efficacy of the investigational medical device.

[7] Originally, LDTs were well-characterized, low-risk diagnostics that were developed in one lab to be used only in that lab. Thus, FDA has exercised discretion in enforcing these tests. With the molecular and genetic science revolution, these tests became more complex, utilizing epigenetic profiling and single nucleotide polymorphism (SNP) analysis, and are increasingly used to assess high-risk diseases to inform critical treatment decisions even if the results do not actually support the decision (Hockett & Close 2010).These tests are also now offered by large commercial laboratories that provide services for many institutions (CDRH 2010).

[8] An exception to this rule are IVDs that are intended for over-the-counter (OTC) use by the patient and IVDs that are used at the point of care (POC), by the attending physician outside of the laboratory (Gibbs 2010).

needed to support their decision to reimburse medical devices, so industry is not sure how much testing and what kind of testing is required (Quinn 2010). Simultaneously, most LDTs lack sufficient evidence of their impact on healthcare outcomes, so providers and payers are unsure how to use the tests and how much they should pay (Hayes et al. 2013). FDA approval is one way to demonstrate value, but even FDA-approved devices may not be reimbursed. FDA has tried to help resolve this issue by introducing a pilot process where a diagnostic can pursue regulatory and reimbursement approvals at the same time, but it has not been widely useful or used (Hollmer 2013).

Section 2.3.1: The challenge of Companion Diagnostics (CDx)

An important subsection of biomarkers are the CDx, which have substantial challenges on their own. As discussed in Section 1, a CDx is a biomarker product that is *essential* for the safe and efficacious use of a drug. A CDx can be co-developed with a drug (the ideal case), or to "rescue" a drug that demonstrated spotty performance during trials, or it can be developed after the drug is on the market (in which case the drug label will be updated accordingly). CDx are high-risk products because the failure of a CDx is equal to the risk of drug misuse.

In most cases, the pace of drug development is faster than the ability to generate and assess biomarkers relevant to the drug (Woodcock 2010). Thus, it has been a major challenge to integrate CDx development into drug development. Developing a CDx also requires collaboration within and across stakeholder groups. For example, within the FDA, CDRH and CDER must play a collaborative regulatory role. Regulatory collaborations also need to occur on a global scale. Currently, CDx regulations are very different at the EMA; in the EU, a broad range of CDx are allowed and there are no EMA-approval requirements

There are currently 61 CDx on the market. In 2013, there were three drugs that were approved with companion diagnostics in their labeling. Experts have said that there will be three to five CDx approvals per year over the next five years (Mullard 2014a). Existing CDx have focused primarily on oncology, but the next wave will be autoimmune diseases and inflammation, as there are biomarkers readily available (Opar 2011).

FDA has provided concept papers and draft guidance documents on how best to integrate CDx into drug development in a clear and cohesive manner. However, industry has pushed back against these documents, calling them unrealistic guidelines that will hinder biomedical innovation (Ray 2013). Thus far, regulation has been done on a case-by-case basis within the context of the specific drug (Dennis et al. 2013; Woodcock et al. 2011; FDA 2014). The FDA has stated that they are trying to be as flexible as possible because they are still learning, even though flexibility doesn't lend itself to defining a predictable pathway (Mansfield 2014). Industry has complained that the regulations have been inconsistent—in some cases the drug and CDx must be approved together, in other cases they are not, and finally, in certain cases, the CDx can be an LDT as opposed to a FDA-approved IVD (Carver 2010), without clear criteria for distinction.

Box 1. The importance of CDx and the controversy between industry and the FDA.

before a company can market a CDx (Pignatti et al. 2014). Within industry, closer collaborations need to occur between diagnostic companies and pharmaceutical companies. These companies need to develop innovative methods for pooling their data and agreeing upon analytical methods. Experts have also proposed ways in which the companies can resolve IP issues in the co-development process (Keeling & Roth 2008). Across stakeholder groups, the FDA and industry need to further collaborate as well and resolve the debate around CDx (Box 1).

The CDx landscape is increasing in complexity and requires regulation. For example, there are at least ten HER2 tests that are FDA-approved CDx, which complicate assay selection (Parkinson et al. 2014). In addition, despite industry backlash, the FDA has repeatedly expressed its intent to regulate all CDx, even if they are LDTs (Trusheim et al. 2013). This is a decision with widespread ramifications because there is increasing evidence as we move into the realm of more personalized treatments that some kind of CDx will be a part of all (or most) drug approvals (Drucker & Krapfenbauer 2013).

2.4: Summary: The Importance of MSCs in Biomarker Adoption

The path from an exploratory biomarker to a biomarker qualified for a specific application context can be long and unpredictable (Vaidya & Bonventre 2010). The main gaps are summarized in Table 5. Note that these gaps are not mutually exclusive. For example, a biomarker used in drug development can later become a CDx, and thus R&D evidentiary standards issues also becomes inherent in CDx.

Table 5. Gaps and challenges preventing biomarker adoption.

Biomarker roles	Gaps/Challenges preventing biomarker adoption
Research & Development Tools	• Resource intensive (data, analysis, samples, etc) platforms • Industry/diagnostic companies are unsure what biomarkers are available for use (no comprehensive, open database) • Lack of stakeholder agreement on required evidentiary standards • Researchers & industry are not provisioning for future patient care (clinical utility), such as by informing clinical device development
Companion Diagnostics (Drug CDx)	• Co-development is difficult to achieve in practice • No high level of cross-stakeholder collaboration; e.g. pharmaceutical and diagnostic companies are not used to working together • Lack of global harmonization; FDA and EMA have distinctly different regulations • No formal regulatory pathway & difficulty aligning stakeholders towards an accepted pathway • Drug-IP conflicts limit sharing of knowledge & data
Prognostic/ Diagnostic Devices	• Lack of incentives to fund large multicentric trials • Very high accuracy thresholds (control vs. diseased patients) needed for early detection • Not fully capitalizing on the knowledge base of R&D biomarkers • Uncertain regulatory criteria from biomarker discovery to clinical utility • Weak and uncertain market incentives unless CDx is involved • Unclear pricing and reimbursement criteria and accepted thresholds
General	• Rapid pace of new data and knowledge generation • Difficulty demonstrating analytical validity at small companies or academic researchers • Unclear how to assess which biomarker is the most robust • Biological variability and heterogeneity of patient samples • Need for multiple BM signatures for disease detection and monitoring • Poor standards on patient selection & sample collecting and handling • Stakeholders work in functional silos with limited communication

It is clear from this section that the main theme underlying current challenges in biomarker adoption is the misalignment, misinformation, and misunderstanding between the different stakeholder groups. Stakeholders work in functional siloes and cross-stakeholder awareness is limited, leading to bottlenecks that prevent the movement of biomarker discoveries through increasing stages of maturity (Keeling & Roth 2008). The lack of global regulatory harmonization policies on CDx regulations is one pertinent example (Pignatti et al. 2014). Increased dialogue and understanding between stakeholders will help shed light on relevant issues and lead to agreeable resolutions.

The level of resources needed (data, samples, expertise, financial, etc.) to discover, develop, and evaluate biomarkers makes this field particularly fertile for collaboration and intervention among key stakeholder groups. MSCs are able to tackle large undertakings such as generating and harnessing high-quality evidence to demonstrate clinical utility or to maintain biomarker databases as more data is produced (Parkinson et al. 2014; Zineh & Lesko 2009). The work and impact of leading cross-stakeholder collaborations in addressing the main gaps listed in Table 5 are discussed in the next sections.

Section 3: The impact of MSCs in promoting biomarker adoption in drug R&D

The first regulatory pathway to promote biomarker adoption in R&D was officially introduced in the form of a guidance document released by the FDA and EMA (FDA 2014). The document described the qualification process for drug development tools (DDTs), which includes biomarkers, clinical outcome assessments (COAs), and animal models. These tools are intended to be used in multiple drug development programs to streamline the R&D process and increase the odds for success. For the purposes of this thesis, the focus will be on the biomarker aspect of the DDT and the specific framework for the biomarker qualification process (BQP) of the guidance.[9]

Before the BQP was established, any biomarker used in drug development would be evaluated by regulatory agencies on a case-by-case basis which was both sponsor- and drug-specific. This means that the sponsor will only perform the tasks necessary to develop the biomarker for a specific drug—even if the biomarker can be applied more widely. Also, in these cases, the information remains proprietary (Dennis et al. 2013; Woodcock et al. 2011; FDA 2014). The BQP aims to establish a global, rather than product-specific, fitness for use (Woodcock 2010).

The BQP represents the first concerted effort to address some of the gaps described in Section 2 (Table 5) by providing a regulatory pathway for biomarker adoption to aid in drug development (Woodcock 2010). The BQP was conceived and developed in a multi-stakeholder environment, with close collaborations between academics, industry, and regulators both in the US and EU. The BQP is one of the main outputs of FDA's 2004 Critical Path Initiative (CPI), which was launched to foster the development of new evaluation tools and drug trial standards by focusing on pre-competitive collaborations and sharing of data (Yu 2014; Mahajan & Gupta 2010; Kelloff & Sigman 2012). Indeed, the guidance itself specifically encourages collaborations since "substantial effort is involved in achieving qualification" (FDA 2014).

Qualifying a biomarker is seen as a lengthy, resource-intensive process that an individual drug company usually cannot accomplish on its own. As BQP participants have stated, a qualified biomarker is "likely most appropriate for roadblocks in particular therapeutic areas or aspects of drug development where there is pressing public health need and where cooperation between drug companies and other entities is imperative for advancement" (Dennis et al. 2013). Safety biomarkers represent a good example. For instance, drug-induced liver injury is the leading cause of drug failures in clinical development and market withdrawals (Chen et al. 2011). Safety biomarkers are critically needed along the healthcare value chain and can be used in many drug development programs. Qualifying safety biomarkers is more likely to be feasible in a collaborative environment because they require a large amount of evidence that would be difficult, if not impossible, to find within one drug company (Woodcock 2010; Amur et al. 2008). The BQP allows stakeholders to work together in new ways—not just between industry and regulatory agencies, but also between preclinical scientists and clinicians—to share expertise and identify and address key research needs (Dennis et al. 2013). Indeed, regulatory

---

[9] The framework can be broadly described as:
   1) Initiation: sponsor submits a letter of intent, which is deemed acceptable by CDER.
   2) Consultation and advice: an initial briefing package is submitted, iterative meetings and correspondence occurs to review the submission, highlight gaps, and review any additional data that is submitted to address those gaps.
   3) Review of the full qualification package: members of the review team offer their assessment and a combined executive summary is published for public comment before CDER issues the final qualification recommendation.

authorities have noted that over 80% of BQP submissions have been from consortia (Woodcock et al. 2011).

According to the FDA guidance, a biomarker that is qualified as a DDT signifies that "within the stated context of use (COU), the DDT can be relied on to have a specific interpretation and application in drug development and regulatory review. Once a DDT has been qualified for a specific COU in drug development, it can be used to produce analytically valid measurements. The DDT...can be used by drug developers...without the relevant CDER review group reconsidering and reconfirming the suitability of the DDT" (FDA 2014). There are a few key points to be noted from this definition.

First, the COU describes the circumstances in which the biomarker can be used, based on the available supporting data (FDA 2014). The COU was mentioned a few times in the Section 2 discussion around evidentiary standards, because it is imperative to understand that the supporting data is only valid within a specific COU. Thus, the data must be reevaluated if a biomarker is used in a different way, or in another medical context (Altar et al. 2008).

Second, the biomarker "can be used to produce analytically valid measurements." Again, it is important to note that the analytical validity of a biomarker depends on its COU; certain contexts might only require an analytical range that is a subset of the whole available range (Campion et al. 2013). Qualification prespecifies that a biomarker has demonstrated adequate reliability, sensitivity, and specificity for a specific context of use (Stephenson et al. 2013). Although *analytical validity* is a term that generally encompasses both the biomarker and its measurement device, this DDT guidance specifies that the biomarker is "conceptually independent" of its device. This distinction is due to the fact that the guidance is written by CDER, who has no authority over devices. Thus, evaluation of the device must be considered separately from biomarker qualification and the BQP uses devices that have already been evaluated (Woodcock et al. 2011).

Finally, the biomarker can be used without reconsidering and reconfirming its suitability. The biomarker may be used across multiple clinical disorders, drugs, or drug classes. This would save the resources of both the regulatory agencies and the drug companies. Many experts hope that the BQP will result in a database of robust biomarkers that have been vetted by regulatory authorities, which will help streamline drug R&D and mitigate the high attrition rates seen in clinical trials (Dennis et al. 2013). In addition, if another sponsor wants to expand a biomarker's qualified COU, it can build upon the available knowledge, as it is not proprietary.

The development of the BQP highlights the importance of collaboration and bold policy changes in healthcare innovation. The BQP has the potential to lead to a harmonized and trusted pathway for biomarker use. The emergence of the process and its success would not have been possible without the work of various MSCs involving the regulatory agencies, especially the MSCs that were initiated by the CPI. These organizations contributed to the various policies in the form of regulatory guidance documents to address some of the challenges of biomarker adoption discussed in Section 2 (Table 5).

Section 3.1: Prequel: Voluntary eXploratory Data Submissions (VXDS) at the regulatory agencies as an example of collaborations across stakeholders at the global scale

The genesis of the BQP is the VXDS program, a collaborative program initiated by the FDA to explore the use of pharmacogenomic data in R&D (Vaidya & Bonventre 2010). The program revealed to the stakeholders that a regulatory path from exploratory biomarkers to biomarkers qualified for a specific context of use was sorely needed.

The human genome was sequenced in 2000, leading to the discovery of many pharmacogenetic biomarkers (Goodsaid & Papaluca 2010). Such biomarkers were increasingly being used in preclinical drug development (Goodsaid & Frueh 2006), but the FDA was not sure how to handle this influx of novel

pharmacogenomic data. In an effort to promote regulatory science in the light of genomic progress, the FDA initiated meetings with industry to discuss appropriate applications of such data. At the first meeting in May 2002, the FDA introduced a novel, voluntary submission process for pharmacogenomics data, called Voluntary Genomic Data Submission (VGDS). VGDS allows sponsors to submit claims for qualification through informal meetings with regulatory agencies (Orr et al. 2007). It was envisioned that these informal meetings would occur under the concept of "Safe Harbor" in which the submissions were evaluated under an "exempt status" so they would not compromise regular drug submissions. In this manner, the stakeholders developed a platform for an open, scientifically-driven information exchange between the FDA and industry so that both parties could learn how to analyze and integrate biomarker data and enable its use (Frueh et al. 2006).

The FDA used what they learned from the first workshop to draft "Guidance for Industry: Pharmacogenomic Data Submissions" and held a second workshop immediately after to discuss this draft with industry in November 2003 (Frueh et al. 2006). The draft guidance formally introduced the VGDS process along with a new Interdisciplinary Pharmacogenomic Review Group (IPRG) charged with reviewing the submissions (Goodsaid & Frueh 2007). IPRG was FDA-wide and involved CDER, CDRH, and CBER collaborations to help resolve the issues surrounding analytical evaluation of new genomic-based biomarkers and their evidentiary standards (Gutman & Kessler 2006; Goodsaid & Frueh 2007). The IPRG had the power to create a review team (termed the Biomarker Qualification Review Team, BQRT) with the right scientific expertise to assess biomarker datasets.

Also in 2003, FDA and EMA authorities held their first *joint* discussions with sponsors on VGDS packages under the framework of a bilateral confidentiality agreement (Goodsaid & Papaluca 2010). Due to the success of this meeting, the VGDS was expanded, with an option for sponsors to have joint FDA-EMA VGDS briefing meetings. In 2006, the EMA and FDA issued the "Guiding Principles for Processing Joint FDA/EMA VGDSs" (Orr et al. 2007). EMA also released its own guideline for industry on how to prepare for dialogue and submissions with the Pharmacogenetics Working Party (EMA 2006).

When the FDA released its final guidance document, it held a third workshop to discuss the issue of varying data quality and standards in submissions (Frueh et al. 2006). The group also decided to change the phrase VGDS to VXDS, or Voluntary eXploratory Data Submissions, because they wanted to formally recognize that submissions extended beyond genomic data and also included AD and depression biomarkers, as well as other "-omics" type exploratory technologies (Goodsaid & Frueh 2007; Orr et al. 2007).
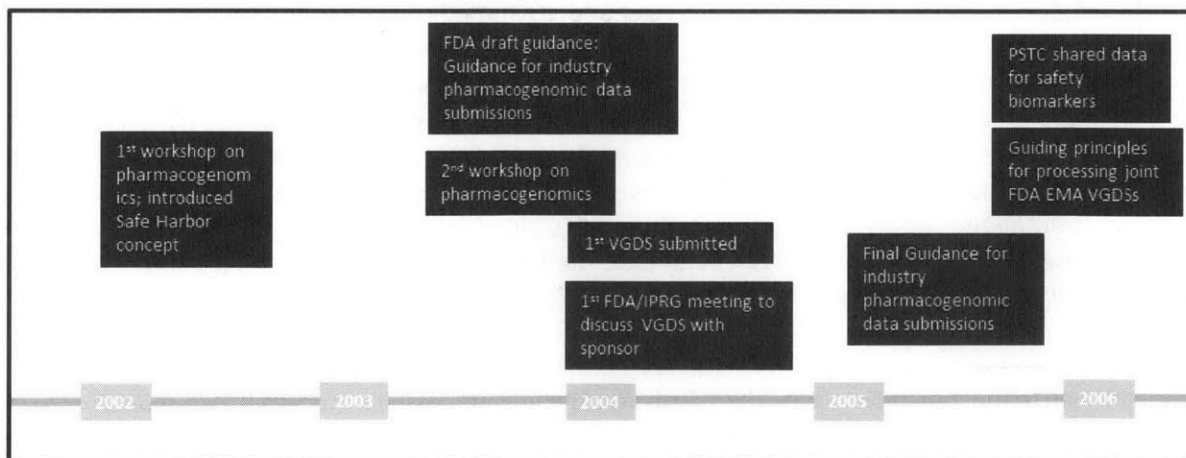

Figure 2. The timeline of voluntary data submissions at the FDA, adapted from Orr et al., 2007.

The impact of the VXDS process (Figure 2) was significant. Data sharing helped educate both the FDA and industry and increased the overall knowledge base (Goodsaid & Frueh 2006). The FDA learned how to analyze emerging data and innovative approaches. Industry also received feedback on trial designs, methodologies, and data interpretation. For example, during one of the first submissions, the FDA and the sponsor realized that they used different normalization and statistical approaches to analyze the data, which resulted in different interpretations (Goodsaid et al. 2010). Several sponsors have used these types of informal interactions with the FDA as a stepping stone to present the same or related data later in a regulatory context to streamline drug submissions, update a drug label, or construct novel adaptive phase III clinical trial designs (Orr et al. 2007).

The collaborative nature of this program led to new policies on genomic data in the form of regulatory guidance documents that both FDA and industry found agreeable, leading to discussions on what new regulations might be needed for the future. For example, AstraZeneca's VXDS submission led the team to the idea that conditional approval may be needed in the co-development of safety biomarkers to allow extensive monitoring of both the drug and biomarker after they had been launched (Goodsaid et al. 2010).

There have been more than fifty VGDS/VXDS submissions since the program's inception (Goodsaid & Mattes 2013). Approximately two-thirds of the submissions have focused on clinical applications of exploratory biomarkers (Goodsaid et al. 2010). However, the VXDS program seems to have outlived its original purpose, as there were no updates beyond 2010 (FDA 2010), although the EMA counterpart continued to hold meetings in 2014 (EMA 2014).

Participants have discussed ways of expanding the scope and goals of the VXDS. For example, regulatory authorities have discussed the possibility that they might serve as a broker to strategically bring together multiple parties to provide different perspectives on biomarkers in various contexts, which could lead to fruitful collaborations (Goodsaid et al. 2010). Indeed, a form of this vision did come to pass: multiple pharmaceutical companies collaborated with the guidance of the FDA to advance biomarker adoption via the Biomarker Qualification Program.

Section 3.2: The Biomarker Qualification Program

The evolution of the VXDS program into the Biomarker Qualification Program is an exemplary story that demonstrates how FDA's launch of CPI and its resulting precompetitive initiatives[10] pushed collaborations to the next level to promote innovations in regulatory science. Regulators recognized that while VXDS remains a valuable pathway, "there is a need…to provide a framework for the development and regulatory acceptance of scientific tools for use in drug development programs [which has] led to the establishment of a pilot project for biomarker qualification" (Goodsaid & Mattes 2013).

The initial process for safety biomarker qualification was done through industry collaboration in a regulatory pathway structure exercise. The collaborative boundaries were pushed further while cross-validating the safety biomarkers with other industry stakeholders who were members of the Critical Path Institute's Predictive Safety Testing Consortium (PSTC). The collaborative regulatory pathway structure exercise expanded upon a VXDS submission: the FDA and Novartis initiated a 2-year cooperative R&D agreement (CRADA) to design a process map for qualifying preclinical safety

---

[10] Note that CPI led to the formation of the Critical Path Institute (C-Path), which was created with support and funding from the FDA, Science Foundation Arizona, and the Tucson community. C-Path is a non-profit organization that serves as a neutral third party to manage funding and protect IP for its MSCs. Currently, there are six C-Path MSCs: Coalition Against Major Diseases (CAMD), Coalition for Accelerating Standards & Therapies (CFAST), Critical Path to TB Drug Regimens (CPTR), Multiple Sclerosis Outcome Assessments Consortium (MSOAC), Polycystic Kidney Disease Outcome Consortium (PKD), Patient-Reported Outcome Consortium (PRO Consortium), and Predictive Safety Testing Consortium (PSTC).

biomarkers and testing this pilot process by submitting the biomarker data to the FDA for qualification as Probably Valid Biomarkers (see Glossary A). CRADAs are important tools to promote innovations in regulatory science because they allow the FDA to leverage industry's resources (funds, data, facilities) while industry is able to develop a close working relationship with the FDA and obtain valuable consultations and advice (Clinton & Wechsler 2006). To cross-validate the biomarkers, Novartis' data was combined with data from other pharmaceutical companies who were PSTC members so that the biomarkers could be qualified to achieve Known Valid Biomarkers status (Goodsaid & Frueh 2007; Marrer & Dieterle 2007).

PSTC's final package of nine kidney safety biomarkers for preclinical use to describe nephrotoxicity in rats was submitted to the FDA and EMA for qualification in June 2007; seven were qualified a year later (Goodsaid et al. 2008). In light of this success, the Biomarker Qualification Program was officially introduced in 2009 (Goodsaid & Mattes 2013). Since then, a number of qualifications have occurred (Figure 3, bottom).

The evolution from VXDS to the BQP illustrates how industry has answered regulators' call and helped them enhance regulatory science. This evolution led to new regulatory policies that have helped clarify some of the evidentiary standards around biomarker use. The knowledge that regulatory agencies gained from these collaborations and qualifications has allowed them to publish several guidance documents which serve to share their current thinking around biomarker policy with industry (Figure 3, top).
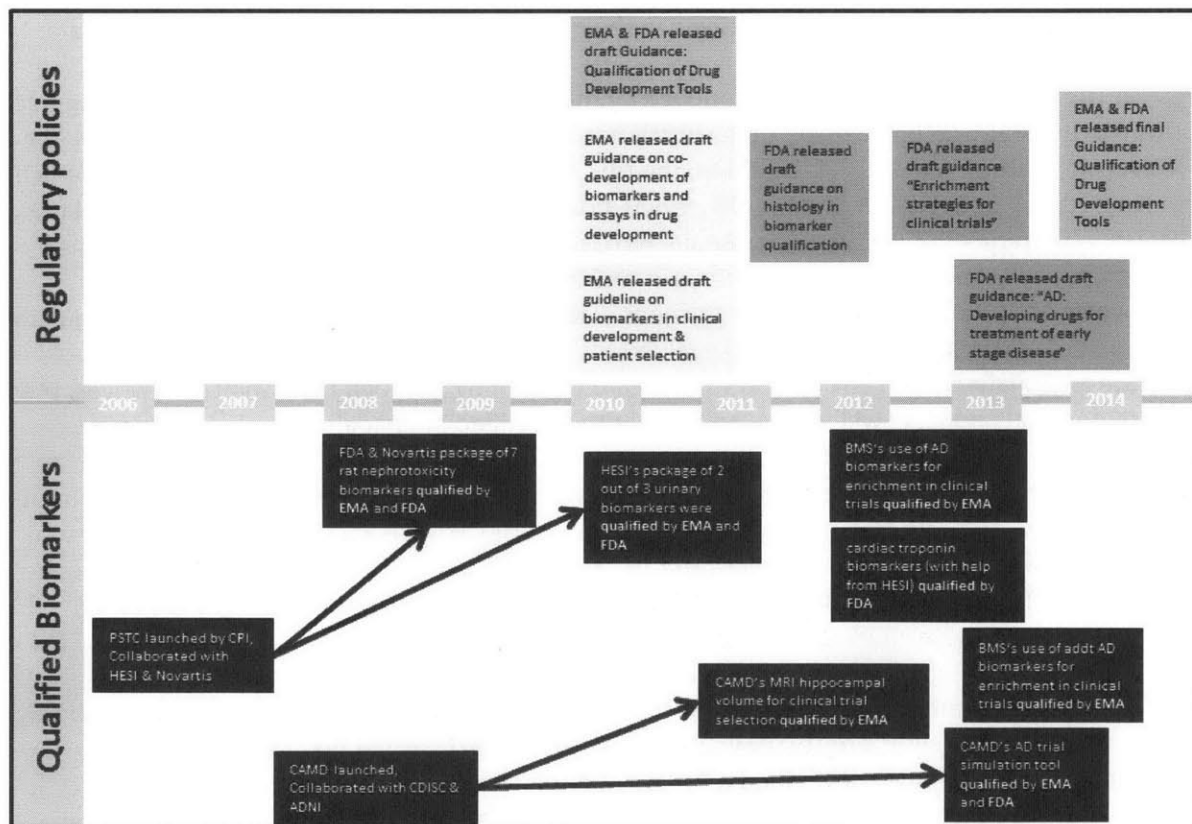

Figure 3. The timeline of the BQP at the FDA and EMA, including relevant tangible outputs of regulatory policies and qualified biomarkers. The MSCs that have submitted qualified biomarkers include: PSTC, CAMD, HESI, and ADNI. Additional information on these biomarkers can be found in Appendix A.

Section 3.3: Analyzing key BQP outputs

To enhance the understanding of collaborations in biomarker adoption, we developed a framework to analyze the contributions of the MSCs across four areas that involve important translational enablers for biomarker adoption. Starting from the scientific challenges addressed by key MSCs, we expand to describe their approaches and achievements on novel processes and policies toward an increased systems capacity for therapeutic innovation through biomarker adoption. The four areas can be described as: 1) creating data and data standards, 2) developing new scientific processes, 3) operationalizing and organizing innovative collaborations, 4) delivering tangible outputs (Table 6).

(1) The biggest data contribution is simply the fact that industry shared their proprietary data to be analyzed in a collaborative setting. This necessitated the creation of a standardized data format which is vital for broad-based data harmonization and knowledge sharing. For example, CAMD collaborated with Clinical Data Interchange Standards Consortium (CDISC) to publish the first data standard for AD, which led to the first database of combined clinical trials to be openly shared by companies with qualified researchers worldwide. The FDA has recognized the importance of such efforts in streamlining the R&D process and mandated that CDISC's clinical standards will be required for submissions by 2017 (Stephenson 2013). In addition, CAMD used this data in their AD biomarker qualification package (EMA). However, more work still needs to be done. For example, during the qualification of the first biomarker package, PSTC mentioned that a common database and data submission format for nonclinical and clinical biomarker data was not available; the group had been using Excel, but they stressed that it was a suboptimal platform for such processes (Dieterle et al. 2010). Thus, it is unclear if the different MSCs involved in the qualification process are using the same data formats and platforms.

(2) The main scientific process introduced by these MSCs is, of course, the BQP. According to the FDA, they have already received many drug submissions that make use of biomarkers qualified by the BQP (Goodsaid & Mendrick 2010). In addition, the MSCs contribute to the scientific process of the BQP by providing "detailed documentation of the review process, to provide guidance for future regulatory submissions" (Dieterle et al. 2010). For example, PSTC publishes all biomarker research, even if the qualification submission is not attained (Dennis et al. 2013). PSTC also explicitly introduced the "Progressive Qualification Framework" which keeps dialogue and evidence gathering open so regulatory agencies can anticipate additional data and evaluations in the future (Sistare et al. 2010). This would include data and analysis from any clinical studies of the currently qualified biomarkers. However, there is no guarantee that sponsors will update and curate the qualified biomarker's database as new evidence is generated (Walker et al. 2013).

The BQP has also stimulated similar scientific processes outside of CDER. Recently, CDER used what it has learned from the BQP to help CDRH frame a similar DDT guidance entitled Medical Device Development Tools (MDDT). CDRH can reach out to CDER to obtain advice about implementing such tools. Like the DDT, the MDDT guidance also places emphases on MSC contributions (FDA 2013a). Indeed, the Medical Device Innovation Consortium (MDIC) was launched in partnership with the FDA with the goal of advancing medical device regulatory science (Anon 2013).

(3) The collaborative structure that the MSCs have agreed upon allowed them to share and work with precompetitive data with a mutually valuable end result. PSTC, as the first MSC to submit a biomarker package to the BQP, developed a "unique Consortium Agreement" that defined the operational and organizational issues around the collaborative structure, including anti-trust issues, confidentiality guarantees, and so forth (Goodsaid et al. 2008; Mattes et al. 2010).

Moreover, innovative collaborations would foster biomarker adoption by structuring a streamlined qualification process between FDA and EMA. PSTC's project on renal safety biomarkers has demonstrated the early impact of streamlining a qualification process at the global level. Although both

FDA and EMA have similar qualification processes, they only share and discuss biomarkers that are jointly submitted (Walker et al. 2013; NCI 2012). Yet industry have stated that joint submissions are challenging in practice (Manolis et al. 2011). One of the key challenges towards global harmonization is the difference in qualification opinions. For example, unlike the FDA, the EMA is open to qualifying a biomarker for certain COU even in the absence of consensus in the field on standardization and cut-off values, with the understanding that the guidance can be updated as new data are obtained. CAMD is currently working to align the qualification process to ensure that data submitted to one regulatory body would also meet the requirements of another agency, although the details of this plan is still unclear (Stephenson et al. 2013).

(4) The BQP stimulated the publication of many guidance documents that outline innovative policies for biomarker adoption (Figure 3, top). However, for BQP to widen its impact, the guidance documents need to be expanded upon. For example, there is no guidance for how to apply retrospective studies to regulatory decision making (Khleif et al. 2010). Cardiac troponins was the only qualified biomarker whose submission was entirely based on retrospective data (Hausner et al. 2013). The FDA has also recognized that they may need to think about how they would handle biosignature qualification (more than 1 biomarker) submissions (FDA 2014). Some experts have proposed an updated document that provides more guidance around expected evidentiary standards (Campion et al. 2013). Explicit guidance should secure inclusion of stakeholder views. For example, some experts have voiced their concern that the BQP team at the FDA has moved towards a more narrow context of use and a conservative interpretation of the data, and the process is becoming more time-consuming and resource intensive (Dennis et al. 2013). As another example, in the 2012 NCI-FDA-NIST workshop report, a regulator involved in the process stated that often, the FDA and the submitter get stuck in an "infinite loop" during the consultation and advice stage of the BQP and he suggested abolishing this inefficient aspect of the process (NCI 2012).

Table 6. MSC contributions for biomarker adoption in R&D.

Data & Data Standards	• Shared data, standards, expertise, common practices • CDISC standards for submissions required in 2017 by the FDA
Scientific Processes	• Developed the BQP • Introduced the "Progressive Qualification Framework" (PSTC) ○ Evolving the qualified biomarkers for clinical use • CDER helped CDRH frame a qualification process for medical devices • Revised diagnostic criteria for AD and MCI (ADNI)
Innovative Collaborations	• Set up initial framework for preclinical trial collaborations(PSTC) • Unique Consortium Agreement as a foundation for future collaborations (PSTC) • Global sharing of data, such as the World Wide ADNI Model
Outputs: Qualified Biomarkers	• 9 qualified nephrotoxicity biomarkers (PSTC, HESI) • 1 (EMA only) qualified hippocampal volume biomarker (CAMD) • 1 AD simulation tool qualification (CAMD)
Outputs: Regulations & Policies	• FDA & EMA DDT Guidance on the regulatory pathway for biomarkers • FDA Guidance on clinical trial enrichment • FDA Guidance for AD treatment • FDA Guidance on the use of histology in qualification • EMA Guidance on pharmacogenomic biomarkers methodology • EMA Guidance on co-development of pharmacogenomic biomarkers

Section 3.4: Conclusion and next steps

The evolution of the VXDS program and the BQP are key historical examples that illustrate the impact of MSCs in promoting novel pathways for biomarker adoption in drug R&D. Experts in the field have discussed where the BQP should go next. One generally unanimous next step envisioned by the MSCs is progressing the qualified preclinical biomarkers into clinical qualification. Industry has already voiced its support for such a program, because currently, sponsors need to discuss their plans with regulatory groups on a case-by-case basis if they wish to use these qualified biomarkers in clinical studies (Campion et al. 2013). PSTC have already begun clinical studies and have initiated collaborations with more clinically-focused MSCs such as the Biomarkers Consortium and the SAFE-T consortium (Dennis et al. 2013). Once qualified clinical biomarkers exist, regulatory agencies can more fully explore how the BQP addresses issues of clinical validity and utility, thus widening the scope of the BQP in the healthcare value chain.

The key challenge with the current scientific process is that the pathway is slow and complex, which does not incentivize industry to contribute their resources to further the BQP. Industry also finds it difficult to integrate biomarker qualification with drug development when the former process is too slow and vague to be of significant use (Campion et al. 2013). Defining the level of assay validation required for each COU and investing in assay development and validation is tricky, and there are no specific guidance documents for this issue (Dennis et al. 2013).

Although MSCs have contributed significantly to the establishment of the BQP, the utility of the BQP alone is limited due to its narrow scope. In the next section, the work of other MSCs will be illustrated. These MSCs possess broader goals and focus on biomarker adoption for patient care, including biomarkers for later-stage clinical trials and diagnostics.

Section 4: MSCs' role in promoting biomarker adoption in later-stage R&D and patient care

From our analysis so far, it has become evident that many processes or regulatory policies are needed to address evidentiary standards for the different types of biomarkers and their COUs. Although the BQP represents one important guideline, it is not the only one being developed. MSCs involved in the BQP have thus far focused mainly on biomarkers for early-stage drug R&D. However, there are other MSCs that are focusing on biomarkers involved in later-stage R&D and clinical care, to allow patients to fully capitalize on the promises of personalized medicine. Biomarkers in the patient care setting cannot be decoupled from their measurement devices, as was seen in the case in the BQP. The BQP evaluated a biomarker in terms of its ability to be a *tool* for drug discovery, whereas a biomarker used in clinical practice for diagnosis, prognosis, or disease management becomes a *product* per se.

Traditional IVDs encompass hematology, glucose-monitoring, and histology. Newer IVDs include tests that measure DNA, RNA, and protein expression. There are different degrees of stringency in the use of IVDs. In the case of a CDx such as the Herceptest for the drug Herceptin, the IVD is *required* for the safe and efficacious use of the drug. An IVD can also be *recommended* in the use of a drug. For example, individuals with low/moderate thiopurine S-methyltransferase (TPMT) activity are predisposed to mercaptopurine toxicity and will show evidence of severe bone marrow toxicity; for such patients, a TPMT assay is recommended. A drug's label can also contain information about changes in efficacy, dosage, or toxicity in a subset of patients, but does not require or recommend the use of an IVD. And finally, an IVD can be thought of as being relatively independent of any drug; for example, 23andMe discusses genetic associations with various diseases, but does not point to any specific therapies.

MSCs in the IVD space encounter new challenging issues because 1) they need to deal with the uncertain outcome of the debate on IVD regulatory pathways; 2) they may need to conduct large clinical trials, which take significant resources but provide very few incentives; and 3) the line between precompetitive and competitive collaboration is more blurry at this stage. This thesis focuses on two key global MSCs that have distinct goals and have approached the particular challenges of biomarker adoption in the patient care and management settings.

Section 4.1: The Biomarkers Consortium

The Biomarkers Consortium was launched in October 2006 (Box 2) with the goal of discovering, developing, and qualifying biomarkers to support new drug development, preventive medicine, and medical diagnostics. From the outset, they have explicitly stated that they are helping to create a new era of personalized medicine, extending their goals beyond developing biomarkers for R&D by addressing the critical later stages of diagnostics development. "The Biomarkers Consortium can help advance evidence generation of clinical utility for molecular diagnostic tests with expanded access to sources of funding and data that are needed to support well-designed studies that produce evidence of clinical utility" (Parkinson et al. 2014). For example, the Cancer Steering Committee Clinical Utility Working Group is helping to provide practical definitions of clinical utility and guidelines on the information needed by payers to support reimbursement for molecular diagnostic assays used in cancer therapy (Parkinson et al. 2014).

The Biomarkers Consortium has completed five projects to date and currently has six that are active. Their first completed project was on adiponectin, demonstrating its value as a robust predictor of glycemic response to peroxisome proliferator-activated receptor agonist drugs used in the treatment of diabetes (Wagner et al. 2009). This first project contributed significantly to the optimization of performance for collaborative environments. Key learnings include the importance of having regular face-to-face meetings and teleconferences, and collaborative resources, such as a website to share information and project updates. The team also decided that the best method to handle legal matters was to designate one legal liaison per company (Wagner et al. 2010).

> The Biomarkers Consortium was an initiative of the Foundation of the National Institute of Health (FNIH), a nonprofit that is associated with but independent of NIH. FNIH is authorized by Congress to broker relationships between NIH and industry, academia, and philanthropies, so that it can seek funding from NIH institutes as well as private partners (IOM 2008). This sort of arrangement is advantageous to FNIH-fostered partnerships like the Biomarkers Consortium because it incentivizes industry participation; the basic investment has already been funded by NIH, and industry is more willing to leverage a preexisting investment (IOM 2010b). The Biomarkers Consortium is divided into a number of disease-specific steering committees (cancer, inflammation and immunity, metabolic disorders, and neuroscience) that are also composed of multiple stakeholders from academia, government, industry, and non-profit/advocacy organizations. The committees solicit project proposals from all sources. Although the Biomarkers Consortium have defined their projects to be pre-competitive, issues revolving around IP and data sharing are resolved on a case-by-case basis (IOM 2008).

Box 2. Background on the Biomarkers Consortium.

Section 4.1.1: I-SPY 2 Trials

A key example of the Biomarkers Consortium's work is the I-SPY 2 trial (Investigation of Serial Studies to Predict Your Therapeutic Response with Imaging And Molecular Analysis), a novel phase II adaptive trial model to accelerate the approval process of effective drugs for breast cancer using biomarkers. I-SPY 2 leveraged two features which were established during I-SPY 1 (Box 3): the infrastructure to ensure accurate and consistent data collection and sharing, and the use of pCR as a primary endpoint.

I-SPY 2 was launched in 2010 with a budget of $26 million and a timeline of five years (IOM 2010b). As a precompetitive adaptive screening trial, it is a pioneering example for both drug and biomarker development. Up to twelve investigational drugs, paired with biomarkers that are predictive of the drug response, will be tested (Box 4). The trial's adaptive design focuses on progressive, real-time learning of which biomarker profiles predict response to each drug (Barker et al. 2009). The trial uses Bayesian approaches to determine the predictive probability success in a phase III confirmatory trial (Barker et al. 2009). Bayesian approaches are considered quite novel and have faced limited adoption in drug development, so I-SPY 2 might increase its use (Allison 2010). Initially, randomization to new drug regimens is balanced. As results become available, predictive indexes are built for each drug and biomarker combination. At this point, randomization is adapted towards assigning new patients to the drugs that are predicted to provide the highest benefit. After approximately sixty patients, drugs that do not sufficiently benefit the patients with any biomarker profiles will be dropped from the study. This is a key development because it is the first time that a trial can compare drug effectiveness in real time and replace ineffective drugs with better candidates.

Biomarkers used in I-SPY 2 belong to three classes (Barker et al. 2009):

1) *Standard biomarkers* are already in clinical use or FDA-approved and will be used to determine patient eligibility and randomization.

2) *Qualifying biomarkers* have not yet been FDA-approved and will be evaluated under Clinical Laboratory Improvement Amendment (CLIA) quality standards. These biomarkers are termed "hypothesis-testing" because they show promise for predicting which patients will respond to which agents.

3) *Exploratory biomarkers* are also termed "hypothesis-generating"; they may possess predictive or prognostic value for breast cancer treatment, but not much information is known yet.

Note that these biomarkers are drug-specific, suggesting that after undergoing the necessary evaluation, they will become either a CDx or a recommended IVD for use during the diagnostic process (L. J. Esserman et al. 2012). Experts have stated that while the sojourn time for any drug in the trial is too short to assess the validity of the biomarker, positive results could lead to FDA approval to use the biomarker in a follow-up registration trial (IOM 2010b). Promising data can support a PMA or a regulatory request to use a biomarker to stratify patients in a Phase III validation study (IOM 2010b).

It took significant time and effort to plan, agree on, and coordinate the unique aspects of I-SPY 2. Indeed, developing the protocol alone took two-and-a-half years, with many multi-stakeholder meetings involving the FDA. The stakeholders that developed the I-SPY 2 protocol set up the innovative "master Investigational New Drug (IND)" concept (Patlak 2010). Usually, multiple drugs and biomarkers require multiple trials, each with its own IND application. However, for I-SPY 2, FNIH was chosen as a trusted intermediary to hold the master IND. Thus, when I-SPY 2 began, the IND protocol had five potential therapies. As the trial progresses, other therapies will be submitted to the regulatory authorities for approval, but the master IND will only be modified with an additional appendix. This highlights the importance of a neutral third party in successful collaborations.

The stakeholders have also worked out an IP agreement that allows them to collaborate in a trial that extends

> I-SPY 1 began in 2002. I-SPY 1 was not a Biomarkers Consortium project but a collaboration of the American College of Radiology Imaging Network (ACRIN), Cancer and Leukemia Group (CALGB), and the National Cancer Institute (NCI)'s Specialized Program of Research Excellence (SPORE). Its primary objective was to evaluate whether response to therapy, as measured by MRI volume and pathologic complete response (pCR), would predict recurrence-free survival (RFS). Its secondary objective was to build a resource and establish standards for collected clinical, molecular, genetic, and imaging biomarker data for a multicenter network that would support high-quality, real-time biomarker evaluation for future trials of such nature (L. Esserman et al. 2012). I-SPY 1 demonstrated that the predictive power of pCR is robust enough to be used as an early indicator of RFS, thus establishing pCR to be a surrogate endpoint (Esserman & Woodcock 2011). I-SPY 1 also set up standard methods for collecting core biopsy material that are then used in measuring and evaluating gene expression profiles, and for MRI-based tumor evaluation (L. J. Esserman et al. 2012; Allison 2010).

Box 3. Background on I-SPY 1.

the boundaries of the pre-competitive line (Goldman et al. 2013). The participants have agreed that the drug IP will remain with the company owning the IP, the biomarker IP will also remain within the company but be licensed for use in the project, and any new IP will be managed by FNIH, which will return royalties to the inventing organizations (IOM 2010b). However, it is unclear what new IP might be created, and what kind of royalties would be implemented. This will set a strong precedent for collaboration in the later stages of development, as collaboration is increasingly being pursued for therapeutic innovation.

The I-SPY 2 trial played an instrumental role in the use of biomarkers to support drug development. As a result of the learnings gained in I-SPY 2, FDA published a draft guidance on using pCR in neoadjuvant treatment for high-risk early-stage breast cancer; pCR would be used as a surrogate endpoint to support accelerated approval (Esserman et al. 2013; Prowell & Pazdur 2012). The guidance recognized for the first time that neoadjuvant and adjuvant therapy resulted in the same long-term outcomes, but the neoadjuvant approach allows earlier measurement of therapy response and provides information about the prognosis of each patient. This draft guidance addresses the lack of an approved regulatory pathway for CDx.

Section 4.1.2: Analyzing the outputs of the Biomarkers Consortium

As in Section 3, the Biomarkers Consortium's contributions to biomarker adoption can be broken down into the four key areas of 1) creating data and data standards, 2) developing new scientific processes, 3) operationalizing and organizing innovative collaborations, and 4) delivering tangible outputs (Table 7).

> The initial five drugs that entered the I-SPY 2 trial are: veliparib (ABT-888; Abbott), conatumumab (AMG 655 & AMG 386; Amgen), figitumumab, and neratinib (Pfizer, Puma). In order to enter I-SPY 2, drugs must have completed Phase I testing, and have evidence of potential efficacy from preclinical or clinical studies (Barker et al. 2009). I-SPY 2 just graduated its first two candidates: AbbVie's veliparib & Puma Biotechnology's neratinib (Fox 2014). Currently, the Biomarkers Consortium is in discussion with AbbVie and Puma about launching I-SPY 3, an adaptive collaborative follow-on, although the compounds might also go through traditional company-sponsored trials. It would be interesting to see how IP issues would be dealt with in the former case.

Box 4. I-SPY 2 drugs and progress.

(1) Like the MSCs mentioned in the previous section, the members of the Biomarkers Consortium also shared precompetitive biomarker data that were aggregated and analyzed to produce robust biomarkers. The I-SPY trials established data standards and data collection; industry stakeholders demonstrated that their drugs' phase I trials were robust and shared their drug candidates to be used in a phase II trial with common data formats.

(2) The scientific process of linking a biomarker to a drug through a trial like I-SPY 2 provides a clearer pathway for the biomarker to become a diagnostic device, thus addressing the co-development pathway gap discussed in Section 2 (Table 5). To streamline the process for subsequent trials, it is imperative that the scientific process for determining the robustness of the drug and the requirements for a biomarker to achieve exploratory, qualifying, or standard categorization are clearly established.

Secondly, the use of a Bayesian approach in drug trials is currently uncommon, but it is integral for an adaptive trial. Thus, I-SPY 2 may lead to wider use and understanding of this scientific approach. Finally, the I-SPY 2 trial provides a policy model for an adaptive, biomarker-driven screening trial with multiple compounds. I-SPY 2 experts have discussed how this model can be used for other cancers that possess emerging predictive markers and poor outcomes (L. J. Esserman et al. 2012). Similarly designed trials such as BATTLE-2 and the Lung Cancer Master Protocol have already begun in the wake of I-SPY 2 (Fox 2014).

(3) The collaborative structure of the Biomarkers Consortium extends a step beyond MSCs such as PSTC because stakeholders are working together on projects that tackle later-stage R&D biomarkers. I-SPY 2 is especially groundbreaking as stakeholders are able to achieve collaboration even though the trial has strong clinical and commercial implications. Such a concept may lead to the formation of a collaborative I-SPY 3 trial. However, there has not been any public discussion of partnerships between pharmaceutical companies and diagnostics companies following the results of I-SPY 2, although this between-stakeholder collaboration is integral to the trial's ultimate success.

(4) Most of the Biomarkers Consortium's completed projects so far have only resulted in presentations and publications. Thus, it is not yet clear whether their work had concrete impact. The main policies that have stemmed from the Biomarkers Consortium are FDA's draft guidance on pCR in neoadjuvant treatment of high-risk early-stage breast cancer and the "master IND" procedure, which is being adopted by other adaptive cancer trials. In addition, although the concept of I-SPY 2 can be seen as an innovative, new regulatory pathway for CDx approval, it is still too early to appraise its success since no drug or biomarker has been approved yet.

Table 7. The Biomarkers Consortium's contributions for biomarker adoption.

Data & Data Standards	• Shared industry data for projects • Contributed common data and standards for breast cancer trials (I-SPY 2)
Scientific Processes	• Developed an adaptive, biomarker-driven screening trial for therapeutics (I-SPY 2) • Developed an adaptive clinical trial that assesses biomarker validity (I-SPY 2) • Establishment of Bayesian approaches in drug trials
Innovative Collaborations	• Framework for late-stage R&D biomarkers collaborations • First time that pharmaceutical companies contributed their drug candidates in a multi-drug trial • Developed unique IP agreement for I-SPY 2 that may lead to novel collaborative processes (I-SPY 3)
Outputs: Biomarkers	• Completed projects (i.e. adiponectin) may be used by industry in efficacy studies • Biomarkers emerging from I-SPY 2 trials may result in diagnostics after undergoing supporting trials & FDA approval • pCR can be used as a surrogate endpoint
Outputs: Regulations & Policies	• FDA Guidance on pCR in neoadjuvant treatment of high-risk early-stage breast cancer • Developed concept of a "master IND" which is now being used in other trials worldwide • First pathway for CDx development (I-SPY 2)

Section 4.2: The Early Detection Research Network (EDRN)

The EDRN focuses on accelerating the translation of biomarker information to clinical application and to evaluate ways of testing cancer in its earliest stages. The EDRN was created by NCI in 1999, with the initial goal of creating an informatics platform to connect different research networks, facilitate collaboration, and promote efficiency and rigor in biomarker research. Originally, the different research networks involved in biomarker development at NCI—including discovery/basic scientists, translational scientists, clinical scientists, and data analysts—worked in relative isolation in their respective siloes. The EDRN was developed to address the resulting disconnect between the abundance of biomarker discoveries to final product output.

The EDRN is composed mainly of research laboratories where biomarker studies are conducted and governed by a Steering Committee of Principal Investigators (PIs) and NCI staff. A Network Consulting Team of external experts from different stakeholder groups reviews progress and recommends new research opportunities. Moreover, the EDRN's Prioritization Subcommittee establishes procedures for allocating resources and prioritizing research, including the establishment of metrics for evaluating progress.

The EDRN collaborates with organizations such as the Jet Propulsion Laboratory (JPL), to develop informatics and standards, and discusses diagnostics requirements at biennial meetings with the FDA and CMS. However, although the EDRN is part of the NCI, it is unclear if the EDRN share practices with other NCI collaborations such as the Cancer Biomarker Collaborative (CBC), the Clinical Assay Development Program (CADP) and the Molecular Characterization Clinical Assay Development Laboratory (MC-CADL) (Williams et al. 2012).

The EDRN is composed mainly of scientists who, unlike traditional academic researchers, have a broad understanding of the rigor and sample size necessary to move a biomarker through development. The EDRN has built a rigorous and widely-known method of biomarker development and evaluation, which serves as an informative example for MSCs in the field. The head of the EDRN has commented that the "EDRN is one of the first organizations to recognize the importance of standardized, prospective collection of biologic samples in the context of clinically relevant circumstances" (Srivastava 2013).

Table 8 illustrates the EDRN's 5-phase biomarker development guideline (Wagner & Srivastava 2012; Pepe et al. 2001). The EDRN notes that "there is no similar collaborative group that is capable of …research from basic discovery to clinical validation" (Wagner & Srivastava 2012). The EDRN's large reference laboratory and the EDRN Knowledge Environment (EKE) allow it to carry out rigorous analytical and clinical validation studies. The EDRN has over 300,000 specimens and ten standard reference specimen sets from standardized, well-characterized, matched cases and controls. EKE is the informatics infrastructure that connects information about biomarkers, studies, specimens, and resulting scientific data (Crichton et al. 2010). The EDRN used JPL's data system as a model to implement their biomarker data analysis and research centers, demonstrating successful technology infusion and agency transfer.

Table 8. The EDRN's 5-phase guideline of biomarker development.

Phase	
1) Discovery	Identify potentially useful biomarkers
2) Clinical assay and validation	Determine the capacity of biomarkers to distinguish between patients with cancer and those without cancer (most biomarker don't progress beyond this phase)
3) Retrospective longitudinal	Determine how well biomarkers detect preclinical disease by testing the markers against specimens collected longitudinally from research cohorts
4) Prospective screening	Identify the extent and characteristics of disease detected by the test and determine the false referral rate
5) Cancer control	Evaluate both the role of biomarkers for detection of cancer and the overall impact of screening on the population through large-scale population studies

With these processes and datasets in place, the EDRN is a pivotal case in the adoption of a "vertical approach" to biomarker research. EDRN is divided into four main functions (Pepe et al. 2001):

1) The Biomarker Developmental Laboratories (BDL) develop and characterize new biomarkers or refine existing biomarkers;
2) The Biomarker Reference Laboratories (BRL) analytically validate biomarkers using blinded specimens, including technical development, refinement, and quality control;
3) Clinical Validation Centers (CVC) clinically validate biomarkers by conducting a multi-site biomarker validation trial; and
4) The Data Management and Coordinating Center (DMCC) performs statistical analysis and supports the informatics infrastructure.

These laboratories and centers work together to move a biomarker along the 5-phase process described in Table 8. Further, the DMCC team designed a process for phase 2/3 biomarker validation trials termed

PRoBE (Prospective-specimen-collection-Retrospective-Blinded-Evaluation) in which specimens are collected before diagnosis to determine whether the biomarker's performance characteristics are sufficient for clinical translation (Prensner et al. 2012). This concept is not unique to the EDRN, although it may possess a more specific, well-validated protocol for such purposes.

It is important to note that out of the 5-phase process, the EDRN only focuses on phases 1-3 (Table 8). Phases 4 and 5 require collaboration with the clinical trial community (Prensner et al. 2012). This demonstrates that progressive involvement of stakeholders is needed to move a biomarker to full clinical validation. It is not clear how smooth the transition is from phase 3 to 4. In addition, the EDRN must engage with industry to license the biomarker and take it through FDA approval. In this case, "industry" typically refers to diagnostic companies. The EDRN holds biannual industry forums that bring together academic investigators and industry representatives interested in commercializing cancer biomarkers (Wagner & Srivastava 2012). The EDRN views the practice of only engaging industry during the latter half of the biomarker development process as its core strength. As the EDRN has stated, "Without the pressure of venture capital, 'best science' rather than 'best business' practice permits development of novel concepts that otherwise, because of commercial needs, might never reach fruition" (Wagner & Srivastava 2012).

Section 4.2.1: Analyzing the EDRN outputs

Again, as we have seen, the EDRN's key areas of contribution to biomarker adoption are listed in Table 9 as 1) creating data and data standards, 2) developing new scientific processes, 3) operationalizing and organizing innovative collaborations, and 4) delivering tangible outputs.

(1) In terms of data and standards, the EDRN stresses the maintenance of robust and well-curated EKE and reference sets, which is a key foundation for successful biomarker development. In addition, introducing the concept of rigorous standards and uniform data sets to research scientists allows them to understand the magnitude of efforts necessary for biomarker adoption. The limitation of the EDRN's data and standards is that their focus on developing early detection biomarkers may be too narrow. Such a structured methodology can be applied beyond this limited category of biomarkers. Thus, it may be more broadly beneficial if the EDRN could reach out to other biomarker collaborations or companies and provide guidance on how to improve their systems.

(2) As with data and standards structures, the EDRN's process of biomarker discovery and development also focuses on robust and rigorous processes. The EDRN's 5-phase approach and PRoBE design can likewise be more broadly applied to biomarkers beyond early detection, to enable wider learnings in the biomarker field. In addition, the EDRN's documents do not explicitly discuss their scientific process in relation to the evidentiary standards framework of "analytical validity, clinical validity, and clinical utility." Hence, the EDRN's scientific process may add to the confusion highlighted in Section 2, where stakeholders use multiple frameworks and terminologies to describe similar concepts. The EDRN has stated that its methods are "widely used." However, it is difficult for any organization to validate such a statement.

(3) The EDRN often acknowledges the importance of collaborations. Indeed, their collaboration with JPL allowed them to apply learnings to build their own informatics system. The EDRN was also one of the seminal organizations that educated siloed scientific researchers on the importance of a broad understanding of the whole value chain in biomarker discovery and development, including how to work together in this space. In addition, although the EDRN typically does not engage industry until a biomarker reaches phase 4 (Table 8), it understands that engagement is necessary because it does not have the capability or resources to commercialize a biomarker. EDRN also prides itself on openness; it has stated that new members constantly join the organization. The Associate Membership option is an easy way to allow a new member access to the EDRN's reference sets for evaluating its own biomarkers. More than two hundred collaborative projects have been initiated, and some have resulted in public-

private partnerships (Srivastava 2013). Currently, the EDRN has more than fifteen collaborations with biotechnology and diagnostics companies (Srivastava 2013).

(4) The EDRN has identified over 1000 biomarkers, prioritized 300, and completed 10 validation trials. These biomarkers begin as CLIA-certified LDTs, but once enough samples have been obtained, the end goal is FDA approval. Thus far, the EDRN has five FDA-approved biomarkers for various clinical endpoints (Anderson & Kodukula 2013). As an example, one of the approved biomarkers was licensed to Vermillion as the OVA1 test. OVA1 was endorsed by the American Congress of Obstetricians and Gynecologists and the product revenue in 2013 was $2.6 M (PRNewswire 2013; Vermillion 2013). However, one of the challenges that the EDRN faces is the ability to track these commercial metrics; industry may not wish to release data on how well an EDRN-licensed diagnostic is performing in the marketplace. But such information is vital because it allows stakeholders to recognize the EDRN's success and further its mission.

Although the EDRN's efforts have not led to explicit regulations or policies, its rigorous pathways and fruitful collaborations can be seen as developing important policy tools. Collaborative research is a cultural mindset, and it is one that is increasingly being adopted by researchers, in part due to organizations such as the EDRN.

Table 9. EDRN's contributions for biomarker adoption.

Data & Data Standards	• Provided an informatics infrastructure to connect all biomarker information through the EDRN Knowledge Environment (EKE) • Contributed EDRN Pre/Validation Reference Sets that contains specimens from well-characterized and matched cases and controls from specific diseases
Scientific Processes	• Developed a 5-phase guideline for biomarker discovery and development • Developed a process for phase 2/3 biomarker validation trials (PRoBE design) • Enabled learning, evaluation, and optimization via process communications and publications that describe both successes and failures
Innovative Collaborations	• Strategic approach in progressively allowing for more collaborations to move outputs forward and address challenges o Collaborated with JPL to learn informatics system infrastructure o Developed collaboration framework between different scientists & different research centers/functions o Developed an open collaborative scheme in which any investigator or company can use the EDRN's reference sets for evaluating their biomarkers o Collaborated with industry to license biomarkers and bring products through FDA approval and market launch
Outputs: Diagnostics	• 5 FDA-approved diagnostics • 5 diagnostics in CLIA-certified labs (pending future FDA-approval)

Section 5: Lessons and recommendations to enhance the role of multi-stakeholder collaborations (MSCs) in biomarker development and implementation

The last few decades have seen a gradual shift towards collaborations in an effort to generate, diffuse, and utilize knowledge and innovation in science and technology. A significant body of research has analyzed various distributed innovation systems, like the Triple Helix analytical framework that conceptualizes university-industry-government interactions and identifies key gaps and opportunities (Carlsson et al. 2002; Etzkowitz & Leydesdorff 1995; Ranga & Etzkowitz 2013).

On a similar ground, the present thesis has developed a proposed framework to map the unique contributions of MSCs in the biomarker field across four areas that involve important translational enablers, particularly data and data standards, novel scientific processes, key operational and organizational aspects of collaborative innovation, as well as the prominent tangible outputs like novel biomarkers and relevant regulatory guidelines. In this concluding section, we distill our analysis in each of these areas into a general set of lessons and recommendations to highlight the role of MSCs in advancing biomarker progress, and discuss current gaps, future challenges, and points of intervention.

Section 5.1: Creating data and data standards

Biomarkers comprise a particularly data and labor intensive field. The level of data sharing through MSCs allows them to tackle projects that would not otherwise be feasible. The data needed for qualification of nephrotoxicity biomarkers, for example, required contributions from multiple industry members in PSTC (Goodsaid & Frueh 2007; Marrer & Dieterle 2007). On this ground, broad-based data gathering and harmonization can support progressive standard setting in terms of formatting requirements and quality control to facilitate subsequent stages of decision making (regulatory submissions, designing novel clinical trials, etc.) This is what happened when ADNI began collecting MRI scans, PET scans, and CSF samples from thousands of patients in hundreds of centers worldwide, and devised common protocols to combine and compare the data (Miller 2009). Through this level of data consolidation, ADNI has contributed substantially to recent revisions of diagnostic criteria for AD and MCI (Aisen 2011). Moreover, other MSCs such as the Alzheimer's Association are building on ADNI's methods, demonstrating increasing importance of the MSC in the field (Carrillo et al. 2013).

Data consolidation from various sources also serves to enhance understanding between stakeholders and across stakeholder groups by decreasing uncertainty and increasing consistency in their evidence interpretations and decision-making criteria. For example, the EDRN Knowledge Environment (EKE) connects basic scientists with translational and clinical scientists, allowing these stakeholders to develop a broader understanding of the magnitude of efforts necessary to develop robust biomarkers (Crichton et al. 2010). In the VXDS program, industry shared data one-on-one with regulators, revealing simple, yet crucial disconnects that could be remedied more easily (Orr et al. 2007; Frueh et al. 2006). For example, during one of their meetings, the two stakeholder groups realized that they used different normalization and statistical approaches which resulted in different interpretations of the same data (Goodsaid et al. 2010). Several sponsors have used the VXDS program to gain a better understanding of FDA's reasoning so they can streamline subsequent drug submissions, update a drug label, or construct novel adaptive phase III clinical trial designs (Orr et al. 2007).

Section 5.2: Developing new scientific processes

The scientific processes developed by these MSCs are essential to break down large, complex bodies of work into manageable workstreams. A characteristic example has been the definition of the COU (context of use) by participants during the design of the DDT; explicitly defining the COU makes

qualification more feasible (FDA 2014). Similarly, the EDRN's vertical approach moves biomarkers along different well-defined workstreams with structured sets of criteria for assessment. The EDRN proactively coordinates its resources so that there is a rapid and efficient "hand-off" of biomarkers and biomarker-relevant information along the chain of laboratories and data centers (Wagner & Srivastava 2012).

The MSCs discussed in the previous sections have also helped establish new pathways for biomarker development and use across R&D stages and disease areas. The BQP represents a consolidated new path to gather diffused data and harness resources in a commercially uncertain space by qualifying biomarkers for use in early-stage R&D. Moving into the clinical setting, the I-SPY 2 trial is the first adaptive, biomarker-driven screening trial with multiple compounds for breast cancer. At this stage, it is too early to determine the project's success, since the trial has not led to new therapies and/or approved biomarkers as of this writing. The I-SPY 2 trial, and all subsequent efforts to be informed, provide an important example of how MSCs can develop novel paths of knowledge and outputs communication to increase our predictive capacity for key decisions around the deployment and therapeutic impact of biomarkers across R&D stages and disease areas. The I-SPY model is already being copied for other pressing diseases, such as lung cancer, in the BATTLE-2 trial and the Lung Cancer Master Protocol (Fox 2014). Further, the Innovative Medicines Initiative (IMI), Europe's largest public-private initiative, recently announced an unprecedented €53 million ($73 million) call for proposals for such a trial in Alzheimer's disease; at least twelve drug companies have already declared their participation (Mullard 2014b). I-SPY 2 also represents a possible pathway for IVD development, if successful (IOM 2010b). Similarly, EDRN's 5-phase process for early detection biomarkers represents a rigorous, well-characterized pathway to bring diagnostics to market (Wagner & Srivastava 2012; Pepe et al. 2001).

Section 5.3: Operationalizing and organizing innovative collaborations

As biomedical knowledge and science become more complex, specialized, and fragmented, it is important to enhance the capacity for their integration across R&D functional siloes and stakeholder groups. MSCs are important vehicles to pull together necessary resources for this convergence. They allow industry to provide the necessary funds, data, and facilities while the FDA contributes advice and guidelines for standards. Neutral parties such as nonprofits and academia offer an impartial space for meetings and discussions. The importance of a neutral, trusted third party, such as FNIH, to hold the master IND document for I-SPY 2 trials is also widely recognized (IOM 2010b). The very involvement in an MSC provides stakeholders with a better understanding of what is needed to work together effectively, including designing detailed IP agreements (IOM 2010b), designating legal liaisons (Wagner et al. 2010), and creating a neutral setting to discuss sensitive information and resolve conflicts of interest (Frueh et al. 2006).

MSCs comprise important operational and organizational scaffolds for collaborative innovation. A key example is the concept of Safe Harbor applied within the VXDS program, which was pivotal in building trust between industry and the FDA. The program allowed the successful development of the pharmacogenomics data submission guidance documents, in which each draft was followed by a collaborative FDA-industry meeting (Frueh et al. 2006). This collaboration led the pilot project for the BQP, which was made possible by collaboration between Novartis and the FDA through the cooperative R&D agreement (CRADA) as well as the Critical Path Institute (CPI). The CPI was created for the explicit purpose of fostering precompetitive collaborations and data sharing and comprises a key example of the subsequent shift towards precompetitive work (Clinton & Wechsler 2006; Woodcock 2010).

This shift was evident even internally, where the different divisions of the FDA—CDER, CDRH, and CBER—collaborated as part of the IPRG review group (Gutman & Kessler 2006; Goodsaid & Frueh 2007). This organizational collaboration led to the development of the BQP, but more broadly, the FDA divisions are more open to sharing knowledge and expertise. Although the BQP has developed mainly

under CDER, CDER is sharing knowledge that they have gained from implementing the BQP and working with the various MSCs involved. Thus, CDRH learned from CDER's DDT and designed a similar initiative termed the Medical Devices Development Tool (MDDT) (FDA 2013a).

As MSCs continue to evolve in breadth and scope, an important issue to be addressed is the concept of global regulatory harmonization. So far, the FDA and EMA have collaborated at certain points during both the VXDS and BQP process, with the future possibility to submit a joint biomarker qualification package (Rinaldi 2011; Goodsaid & Mattes 2013). However, the two biomarker qualification processes are not identical; for example, EMA requires a fee for submission and FDA does not, and the latter has no formal deadlines for response times (Walker et al. 2013). In addition, the FDA and the EMA have widely different regulations on IVDs; the FDA is generally considered more stringent (Senderowicz & Pfaff 2014). Global harmonization in the case of the BQP process will be important to attain for aligning on a topic of international consequence.

Finally, the increasing shift towards strategic partnerships across MSCs allows incremental value generation from independent outputs and achievements. The evolution of the VXDS program to the BQP, ADNI's growing Alzheimer's disease database, and PSTC's expansion beyond its preclinical focus by collaborating with the Biomarkers Consortium and SAFE-T Consortium are all key examples (Dennis et al. 2013).

Section 5.4: Delivering tangible outputs

Many robust biomarkers and regulatory guidance documents that outline innovative policies for biomarker adoption have already emerged from the work of MSCs. Qualified biomarkers include nine nephrotoxicity biomarkers, one cardiac troponin biomarker, and three Alzheimer's disease-related biomarkers. The EDRN has produced five FDA-approved diagnostics for early-stage cancer detection and five diagnostics for use in CLIA-certified labs.

Moreover, a number of MSCs have contributed to the development of guidance documents that are key in facilitating biomarker adoption. The VXDS program led to new policies on genomic data in the form of regulatory guidance documents that both the FDA and industry found agreeable (Goodsaid et al. 2010). Furthermore, the MSCs involved in the BQP helped launch the DDT guidance document and submitted biomarker packages that have been qualified for use in drug R&D. The Biomarkers Consortium's I-SPY 2 trial contributed to the pCR draft guidance document, the "master IND" procedure, and a regulatory pathway for CDx in the case of high-risk early breast cancer (L. J. Esserman et al. 2012). However, although developing policies in a multi-stakeholder space can lend itself to increased agreement and understanding, this is not always the case. For example, as of this writing, the IVD guidance document is still open and controversial, with no agreement between the FDA and industry (Carver 2010).

Section 5.5: Recommendation—Avoid fragmentation and duplication of pre-existing efforts

One of the key drivers of future success for both collaborative and individual efforts is to avoid fragmentation and duplication of work. An example of possible fragmentation is in the goals of the PSTC and the Biomarkers Consortium. PSTC is clearly working towards biomarker qualification, but the Biomarkers Consortium's work is less explicit. The Biomarkers Consortium has only stated that it plans to take one of its biomarker projects—FDG-PET and volumetric CT for outcome measures for lung cancer and lymphoma—through the qualification process at the FDA.[11] Thus, although the PSTC and the Biomarkers Consortium are collaborating to move qualified preclinical biomarkers into the clinical space

---

[11] The final objectives of active projects at the Biomarkers Consortium can be found at: http://www.biomarkersconsortium.org/projects_active.php

(Dennis et al. 2013), it is not clear if they are working towards the same goal of qualification. This can be a point of confusion for other MSCs or individual stakeholders in the biomarker field. To give another example, the EDRN has developed a rigorous data repository for biomarkers of early detection (Srivastava 2013). However, it is unclear if the EDRN's work is being imparted to other MSCs working towards similar rigorous standards. It is imperative for other MSCs who are building similar databases, even if expanding beyond early detection biomarkers, to leverage the EDRN's learnings in terms of data rigor and formatting.

It will be crucial for MSCs to be aware of other collaborative or individual efforts in the same field and build upon each other's expertise. Even within the collaborative space, it is common to see competition or conflicts of interests among MSCs. To avoid current fragmentation and duplication of efforts, it will be important to have open, direct dialogue between the key members of each MSC. Such conversations will serve to highlight not just the ultimate goals of their initiatives but also the necessary consolidation of activities to achieve ultimate impact from the adoption of their outputs.

Section 5.6: Recommendation—Maintain flexible scientific processes

In the current environment of rapid scientific progress, scientific processes should be flexible, with the ability to expand their scope as more data and information become available. While the current use of a biomarker might be limited, it is good practice to keep the broader uses of the biomarker in mind and have necessary infrastructure in place to capture relevant emerging data and knowledge. Such information can be used to move the biomarker across subsequent stages of maturity (analytical validity to clinical validity, for example) or across disease stages.

On this note, progressive knowledge expansion requires detailed documentation of the scientific process to allow future efforts to build upon previous work. The PSTC and EDRN, for example, publish their failures as well as successes in order to be comprehensive (Dennis et al. 2013). Efforts to improve the accuracy of biomarker tests and integrate them into elaborate algorithms will be pivotal in achieving consolidated information for drug treatment or preventive approaches.

Even further, the very process by which guidance documents and other such policies are created and reviewed should be flexible as well. Dialogue needs to be kept open to make sure that policies evolve with increased consensus of the affected stakeholders. As mentioned in Section 3, experts have voiced concerns that the BQP team at the FDA has moved towards a more narrow context of use and a conservative interpretation of the data, making the process more time-consuming and resource-intensive (Dennis et al. 2013). On an operational level, such concerns should be addressed in open meetings between FDA and industry. To best ensure that such meetings occur and that policies continue to evolve, MSCs should set firm rules for "check-in" meetings after certain milestones.

A key example of a progressive build-up from an existing guidance document is CAMD's novel data-driven trial simulation model as a suitable clinical endpoint for describing cognitive changes in patients with mild and moderate AD. The model underwent the same regulatory process as any other biomarker and obtained a positive qualification decision from the EMA and the FDA. However, although this model was not a "biomarker," FDA expanded the current definition of a DDT, rendering CAMD's submission a "fit-for-purpose" drug development tool (Rogers 2013).

The DDT guidance document can accommodate further progressive expansion into other areas of R&D as well as clinical practice. Experts have suggested that the DDT guidance can be adapted to evaluate biomarkers for safety monitoring, personalized healthcare strategies, and post-market safety surveillance (Dennis et al. 2013). As biomarker qualification expands in scope, further guidance will be needed for designing acceptable retrospective studies (Khleif et al. 2010) and outlining expected evidentiary standards (Campion et al. 2013), including stakeholder-wide consensus on the criteria for analytical validity, clinical validity, and clinical utility. These next steps can be challenging in terms of both necessary incentives and resources, highlighting the key role of MSCs to achieve agreement in

purpose and prioritization of goals across different stakeholder groups. CRADAs, the industry-FDA cooperative R&D platform, can be a good model to use, as the one-on-one factor provides for more focused commitment.

In this context, it will be important, especially for regulators, to clarify the breadth and scope of new policy paradigms such as the BQP. For example, although this was not stated by the regulators, perhaps pCR as a surrogate endpoint did not undergo the BQP process because it is already described in a draft guidance (FDA 2012) and is thus perceived as an accepted biomarker. If true, this would signify that draft guidance documents can represent another biomarker adoption pathway. It would be helpful for the learning process if the regulators explicitly stated that this was the case.

Section 5.7: Recommendation—Strategic and timely inclusion of more stakeholders

The inclusion of more stakeholders, such as patients, diagnostic companies, and payers, diversifies the decision criteria for an MSC. For example, patients and public advocacy groups can represent a powerful force to push for biomarker development, by both bringing awareness and funding. However, without a formal patient inclusion process in place, they still represent an "idiosyncratic process that depends on the emergence of leadership in a given disease area" (Mittleman et al. 2013). Furthermore, given the central role of patients in successful biomarker development and use, earlier and more encompassing engagement should accelerate the field.

Diagnostic companies, on the other hand, have deep technical expertise on how to build the analytically robust measurement devices that are necessary for the accurate performance of a biomarker. However, diagnostics companies do not possess the commercial power of drug developers, and what drug developers consider "precompetitive data" is typically considered "competitive" by diagnostic companies. Perhaps for this reason, it has been difficult thus far to bring diagnostics companies into biomarker collaborations, such as the I-SPY 2 trial, where there is little information on the role they could play. It is possible that the drug developers involved in the trials have in-house diagnostic divisions or have already acquired or partnered with diagnostic companies. However, such approaches are not sustainable or sufficient to resolve the critical issues that are plaguing diagnostics companies today in the commercial and regulatory space. Given the future impact of diagnostics in personalized medicine, increased dialogue and awareness for every stakeholder group could help resolve these challenges.

The engagement of payers is also of paramount importance because reimbursement decisions play a major role in the actual use of biomarkers in the clinical setting. Although payers are starting to demand clinical utility, evidence requirements are still vague (Quinn 2010). It is unclear how payers will evaluate the biomarkers used in I-SPY 2 or the other biomarker projects from the Biomarkers Consortium. In this respect, the Cancer Steering Committee Clinical Utility Working Group's efforts to provide practical definitions for COU and resolve evidence requirements for IVDs will be critical (Parkinson et al. 2014).[12] The pilot process for moving an IVD jointly through the FDA and Center for Medicare & Medicaid Services (CMS) represents an emerging example of payer engagement, but its effectiveness remains to be seen, given its limited use (Hollmer 2013).

Although regulators represent a key stakeholder group in many of the MSCs we discussed, it will be more beneficial to involve them earlier in discovery and development studies. Regulators can help plan for the type of data required, at the same time gaining a deeper understanding of new technologies and methodologies to better foster evidence-based regulatory decisions. Biomarker development is moving into an era where the pace of innovation needs a new and more flexible regulatory mindset on

---

[12] The Cancer Steering Committee Clinical Utility Working Group is a part of the Biomarkers Consortium.

how to evaluate the development process—from rigid decision points to progressive evaluations. Regulators may have an interest in keeping decisions flexible so that they remain more in control of the process.

Much like the earlier engagement of regulators, a later engagement of industry would be desirable under certain situations, such as in the development of early detection biomarkers when the commercial value may not be immediately apparent. At the EDRN, this approach is viewed as its core strength; they have stated that, "Without the pressure of venture capital, 'best science' rather than 'best business' practice permits development of novel concepts that otherwise, because of commercial needs, might never reach fruition" (Wagner & Srivastava 2012). The EDRN is funded mainly by NCI, an organization that is anchored around strong scientific priorities. Although industry would be more interested in the commercial outputs of scientific research, organizations like the EDRN can broaden the focus from the end product to include the generation of knowledge on the complexity of disease and physiology. Such new financial models can balance commercial incentives by supporting and incentivizing more science-focused work.

Section 5.8: Recommendation—Develop metrics to track progress and measure impact

Establishing metrics is important to set goals, manage collaborations, identify gaps and drive improvements in any system. Innovation systems and management research have extensively discussed how to measure the performance of a system and how the measurements must adapt to its lifecycle (Carlsson et al. 2002). Especially in the field of collaborations, measurements must also be transparent and subject to discussion and revision, so as to include all stakeholder viewpoints and avoid biased behavior. For example, a metric such as "number of qualified biomarkers" may result in a large number of biomarkers qualified based on a COU that is too narrow to be useful for sponsors; it will be necessary for the MSCs involved to insure that qualified biomarkers provide value to drug developers.

Despite the extensive work with biomarker-focused MSCs, there is currently a lack of metrics and performance measurement approaches. Some experts have criticized the Biomarkers Consortium for its lack of output, remarking: "after 6 years in operation, does the Biomarkers Consortium represent a good return on investment for its members? No qualified biomarker developed through the Biomarkers Consortium has yet delivered tangible results to patients or other stakeholders" (Mittleman et al. 2013). Developing metrics to track the progress and measure the impact of the Biomarkers Consortium's deliverables will help them address such criticisms. Metrics are also broadly relevant to all other ongoing collaborations and individual efforts to enable wider understanding in the biomarker field.

A similar challenge is faced by the BQP—a key issue being the lack of transparency in the program. For example, experts have stated that there is no public information about cycle time, the process of how the BQP team assesses biomarkers, or its efficiency (Goodsaid & Mattes 2013). However, certain members involved in the BQP are aware of the importance of such evaluations and have suggested metrics that include "faster decision making for advancing or terminating developmental compounds, increased numbers of compounds progressing through the pipeline, and reduced cost of nonclinical and clinical development programs" (Walker et al. 2013). Moreover, the ability to track a qualified biomarker's uptake is hindered by confidentiality provisions since biomarkers used in a drug development program are considered to have moved back into the competitive space (Walker et al. 2013). An imperative for the sustainability of the qualification process is sharing information about successes in biomarker uptake. Some members involved in the BQP feel that it is still too early to develop metrics, especially to compare the BQP across agencies for strengths and weaknesses (Campion et al. 2013). More data need to be gathered before informative analysis can occur. Such considerations

led the EMA to agree that the BQP should be reviewed only after ten qualifications, when further revision will be considered to adjust to the needs of all participants (Vaidya & Bonventre 2010).

Transparency is also an issue with metrics development for the EDRN. The EDRN's Data Management and Coordination Center tracks a set of metrics which includes the number of publications, patents, public-private partnerships, LDTs, and FDA-approved products (Crichton et al. 2010). However, industry's commercial rights prevent the EDRN from being completely transparent about other crucial metrics, such as how EDRN-licensed products are performing in the marketplace.

Section 5.9: Conclusion

MSCs have introduced a variety of approaches that have the potential to significantly enhance our capacity to effectively use biomarkers to optimize both drug R&D and clinical care. The analysis and recommendations provided in this section serve to highlight the main gaps and opportunities, based on selected key MSC examples. Tracking the development and progress of the various MSCs involved in biomarker adoption has revealed many lessons for analysis and discussion. As MSCs are moving to the next level, the different stakeholder groups should be careful to keep communication open to share data, standards, scientific processes, and collaborative ideas, and to avoid fragmentation of efforts and duplication of outcomes. It is imperative to keep in mind that all scientific and regulatory processes, new or old, should be kept flexible, as technologies and science continue to evolve. As the magnitude of challenges is too large for any single stakeholder, greater consensus and alignment will have to be attained. On this front, it is important to involve stakeholders in a progressive and strategic manner with MSCs, and to progressively engage new ones, including patients, diagnostic companies, and payers. Key metrics will need to be developed to monitor performance and progress of individual and collaborative efforts. The variety of stakeholder incentives and perceived conflicts of interests should be further identified and formally considered.

During the past two decades, the biomedical industry has gradually shifted towards a more collaborative approach to tackle challenges that exceed individual stakeholder capabilities and to integrate the current fragmented paradigm in drug R&D. With careful design and implementation, this crucial cultural shift will lead to greater openness and broader understanding to accelerate and drive sustainable progress in biomedical innovation through the establishment of the necessary operational, organizational, and management structures.

Glossary A: The main biomarker categorizations proposed by field experts

FDA

In order to try and establish a process for the submission and review of pharmacogenomics and toxicogenomic data, the FDA developed a Guidance for Pharmacogenomics Data Submissions (2005) (Goodsaid & Frueh 2006). The guidance defines biomarkers by its appropriateness in regulatory decision making:

Biomarkers	Definition
Probably valid biomarker	A biomarker that is measured in an analytical test system with well-established performance characteristics and for which there is a scientific framework or body of evidence that appears to elucidate the physiologic, toxicologic, pharmacologic, or clinical significance of the test results. A probable valid biomarker may not have reached the status of a known valid marker because, for example, of any one of the following reasons: • The data elucidating its significance may have been generated within a single company and may not be available for public scientific scrutiny. • The data elucidating its significance, although highly suggestive, may not be conclusive. • Independent verification of the results may not have occurred.
Known valid biomarker	A biomarker that is measured in an analytical test system with well-established performance characteristics and for which there is widespread agreement in the medical or scientific community about the physiologic, toxicologic, pharmacologic, or clinical significance of the results.

The FDA, in its Guidance: Qualification Process for Drug Development Tools (2014) has defined biomarkers by what disease and treatment characteristics they help define:

Biomarkers	Definition
Diagnostic	A disease characteristic that categorizes a person by the presence or absence of a specific physiological or pathophysiological state or disease.
Prognostic	A baseline characteristic that categorizes patients by degree of risk for disease occurrence or progression of a specific aspect of a disease. A prognostic biomarker informs about the natural history of the disorder in that particular patient in the absence of a therapeutic intervention. It can be used as an enrichment strategy to select patients likely to have clinical events of interest or to progress rapidly.
Predictive	A baseline characteristic that categorizes patients by their likelihood of response to a particular treatment relative to no treatment. A predictive biomarker can be used as an enrichment strategy to identify a subpopulation likely to respond to a treatment intervention in a particular way. It may predict a favorable response or an unfavorable response (i.e., adverse event).
Pharmacodynamics	A change in the biomarker shows that a biological response has occurred in a patient who has received a therapeutic intervention and for which the magnitude of the change is considered pertinent to the response. A pharmacodynamic biomarker may be treatment-specific or more broadly informative of disease response.

Industry

Experts from industry have defined biomarkers by graded categories based on the biomarker's intended application (Wagner et al. 2007):

Biomarkers	Definition
Exploration	Biomarkers are R&D tools accompanied by *in vitro* and/or preclinical evidence, but no consistent information linking the biomarker to clinical outcomes in humans. Used for hypothesis generation in drug development.
Demonstration	Biomarkers are associated with adequate preclinical sensitivity and specificity and linked with clinical outcomes, but have not been reproducibly demonstrated in clinical studies. Used for decision-making and supporting evidence with primary clinical evidence in drug development. Corresponds to "Probable Valid Biomarker" in FDA Guidance for Pharmacogenomics Data Submissions.
Characterization	Biomarkers that have linked preclinical sensitivity and specificity with clinical outcomes, and have been reproducibly demonstrated in more than 1 prospective clinical study in humans. Used for decision-making, dose finding, secondary/tertiary claims. Corresponds to "Known Valid Biomarker" in FDA Guidance for Pharmacogenomics Data Submissions.
Surrogacy	Biomarkers that have undergone a holistic evaluation of the available data and can substitute for a clinical endpoint. Used for registration in drug development. The designation of "surrogate end point" requires agreement with regulatory authorities.

Thomson Reuters

Many companies use the Integrity biomarker database from Thomson Reuters in their R&D process. This database defines biomarkers more specifically, in order of continuum of care:

Biomarkers	Definition
Risk stratification	The role of this biomarker is to determine a person's risk of suffering a particular clinical event within a specified period of time.
Risk factor	The role of this biomarker is to determine a person's risk of a disease on the basis of epidemiological evidence.
Disease profiling	The role of this biomarker is to obtain information about the disease, but there is insufficient data to assign a clinical role. The data is often obtained from high throughput analyses, for example transcriptional profiling, and might be extrapolated to the processes that cause the disease.
Screening	The role of this biomarker is to sort a population into 'healthy' and 'non-healthy.' Screening is an epidemiological process, though the same process may serve for diagnosis as well.
Diagnosis	The role of this biomarker is to identify or detect a disease
Differential diagnosis	The role of this biomarker is to distinguish between two or more diseases with similar signs and symptoms
Staging	The role of this biomarker is to describe how far a disease has progressed in a patient. The stage at diagnosis is often a prognostic indicator of overall survival and can be used as a guide for subsequent therapy.
Selection for therapy	The role of this biomarker is to select a sub-group of patients suitable for a particular therapy.
Toxicity profiling	The role of this biomarker is to obtain information about the underlying cause of

	an adverse or toxic event, but there is insufficient data to assign a predictive or monitoring role. A biomarker use with "toxicity profiling" represents the birth of that use, after the first mention of the association between the biomarker and the adverse event and is always experimental. When new studies that focus on its predictive or monitoring role are added to our database, the role will be changed and upgraded.
Prediction of drug resistance	The role of this biomarker is to detect possible resistance to therapy and thus to exclude that therapy from the possible therapies available to the patient.
Monitoring of disease progression	The role of this biomarker is to detect possible resistance to therapy and thus to exclude that therapy from the possible therapies available to the patient.
Monitoring of treatment efficacy	The role of this biomarker is to identify signs of a change (usually beneficial) as a result of treatment. A biomarker used for monitoring treatment efficacy is usually measured before the treatment starts (baseline) in and at stages throughout the treatment (follow up).
Monitoring treatment toxicity	The role of this biomarker is to identify signs of adverse effects as a result of treatment. Measured at baseline and at stages throughout treatment.

The Biomarkers Definitions Working Group

The Biomarkers Definitions Working Group (FDA, NIH, academia, industry, 2001) also provided a classification system for biomarkers:

Biomarkers	Definition
Type 0 (natural history)	Markers of natural history of a disease and that coordinate longitudinally with known clinical indices, such as symptoms over the full range of disease states.
Type I (therapeutic intervention)	Markers that capture the effects of an intervention in accordance with the mechanism of action of the drug, even though the mechanism might not be known to be associated with clinical outcome.
Type II (surrogacy)	Markers that are surrogate endpoints, and distinguishing this from clinical endpoints. They must be relevant to both the mechanism of action of the drug and to the pathophysiology of the disease.

Glossary B: Mapping differences in semantics on the terms *validation* and *qualification*

The terms *validation* and *qualification* are not well-defined. It has been stated: "The biomarker literature occasionally uses 'validation' and 'qualification' or 'evaluation' interchangeably. We have avoided this because the validation and qualification processes must be distinguished, and the term 'validation' does not adequately describe the qualification process" (Wagner et al. 2007). The following tables highlight key subject experts' views:

Validation	Expert Remarks
(Wagner 2008)	"fit-for-purpose process of assessing the assay and its measurement performance characteristics, determining the range of conditions under which the assay will give reproducible and accurate data"
(Bleavins 2010)	"biomarker validation [is] a continuum of technical bioanalytical method validation, which is a familiar process that does not generally pose a problem for an organization embarked on assay development" "the term validity is a broad concept that has been used to describe everything from the analytical methods to the characteristics of the biomarkers identified. Validity is also used across multiple industries, not only medical or health disciplines. Therefore, when referring to biomarkers, validation is sometimes termed qualification for clarity."
Biomarkers Definitions Working Group (Atkinson et al. 2001)	"The process of retrospectively linking a surrogate endpoint to a clinical endpoint has often been referred to as validation. In addition, the term validation is also often used to address performance characteristics (ie, sensitivity, specificity, and reproducibility) of a measurement or an assay technique... the term validation is unsuitable for the description of the process of linking biomarkers to clinical endpoints."
(Lee et al. 2006)	"Fit-for-purpose method validation provides for efficient drug development by conserving resources in the exploratory stages of biomarker characterization. For example, a biomarker under exploratory development in an early phase clinical trial would be less rigorously validated than an already well-qualified biomarker in the same trial. By definition, exploratory biomarker data would be used for less critical decisions"

Qualification	Expert Remarks
(Wagner et al. 2007)	Some have defined qualification as a scientific linkage to physiology: "The evidentiary process of linking a biomarker with biological processes and clinical end points. Biomarker qualification has been defined as a graded fit-for-purpose evidentiary process linking a biomarker with biology and clinical endpoints"
(IOM 2010a)	Qualification can also be defined as a scientific linkage to disease: "A description of the evidence relating to the biomarker in question—as measured using validated tests—to the intervention and disease outcome"
(Lee et al. 2006)	Some have more narrowly refined *qualification* as *clinical qualification*: "The evidentiary and statistical process linking biologic, pathologic and clinical endpoints to the drug effect, or linking a biomarker to biologic and/or clinical

	endpoints."
(FDA 2014)	For the FDA, qualification aligns with regulatory decision and review through the BQP: "A conclusion that within the stated context of use, the results of assessment with a biomarker can be relied upon to have a specific interpretation and application in drug development and regulatory review" [BQP]

Appendix A: Background information on the MSCs that have participated in the BQP

Health and Environmental Science Institute (HESI) was established in 1989 as a global branch of the ILSI (International Life Sciences Institute), with members from industry, academia, and regulatory agencies. The committee's aims are to pool resources to support scientific research, sponsor symposia or workshops, and conduct other technical activities. In 2002, it established a new program for biomarker activity, with a focus on nephrotoxicity biomarkers. HESI was working towards peer-reviewed publication, but when CPI met with them in March 2006 to discuss FDA's aim to establish a new process for biomarker qualification review, HESI decided to submit data to FDA instead (Goodsaid & Mattes 2013). Although HESI played an instrumental role in the qualification of nephrotoxicity biomarkers and cardiac troponin biomarkers, they have since decided that they will concentrate on biomarker discovery instead and peer-reviewed publication as an endpoint (Goodsaid & Mattes 2013).

Predictive Safety Testing Consortium (PSTC) was formed by CPI in 2006, with members from industry, EMA, FDA, and participating scientists, including clinicians and preclinical scientists. The Consortium's mission is to collect and summarize data and assess drug safety with preclinical and clinical biomarkers. PSTC is not oriented towards new biomarker discovery, but instead focuses on translational biomarkers that the participants have at least some experience and confidence with (Sistare et al. 2010). PSTC developed a "Progressive Qualification Framework" that defines the critical core set of data needed initially, but also keeps dialogue and evidence gathering open, to allow regulatory agencies to anticipate that additional data and further evaluations are expected and wanted (Sistare et al. 2010). The Nephrotoxicity Working Group submitted full biomarker qualification package for seven biomarkers of nephrotoxicity to the pilot process of biomarker qualification at FDA and EMA in July 2007 (Goodsaid et al. 2008). These biomarkers were qualified in 2008 for non-clinical application in drug safety evaluation using the rat model (Goodsaid & Mendrick 2010). These qualified biomarkers are now being tested for use in safety tests in human clinical trials by the PSTC's Kidney Working Group and Translational Team, collaborating with the Biomarker Consortium (Rinaldi 2011) as well as the Safer and Faster Evidence-based Translation Consortium (SAFE-T).

Coalition Against Major Diseases (CAMD) was launched in February 2008 with a mission to create new tools and methods that can be applied during the development process of new treatments for Alzheimer's and Parkinson's Disease. CAMD is focused on creating common data sharing standards, establishing a database of pharmaceutical clinical trial data, developing disease progression models to aid in clinical trial designs, and identifying biomarkers for use in patient selection (CAMD 2010). They are basing their plan of action on PSTC's successful collaborative framework (Romero et al. 2009). CAMD collaborated with Clinical Data Interchange Standards Consortium (CDISC) to publish the first data standard for AD by pooling together data on more than 6000 patients who had been included in 21 clinical trials and combined this data with ADNI's data (Goldman et al. 2013). This is the first database of combined clinical trials to be openly shared by companies to qualified researchers worldwide. Due in part to their large-scale efforts, the FDA has mandated that CDISC's clinical standards will be required for submissions by 2017 (Stephenson 2013). In addition, CAMD's work resulted in FDA issuing draft guidance for enrichment strategies for clinical trials and developing drugs for treatment of early stage AD (Stephenson 2013). CAMD has also submitted an AD simulation tool that was qualified by EMA and FDA and a hippocampal volume measurement biomarker for clinical trial selection (qualified by the EMA) (Sinha 2013). Currently, they have submitted a qualification package for the use of DAT imaging as a biomarker to enrich clinical trials in patients with early-onset PD (Stephenson et al. 2013).

The **Alzheimer's Disease Neuroimaging Initiative (ADNI)** began in 10/2004 as a joint business venture between IND (a boutique CRO) and Synarc (an imaging CRO). Then, a large neuroimaging AD conference brought together these companies with other like-minded industry stakeholders and they jointly formulated the idea of industry co-funding a large gold standard AD study with NIH, thus leading to the concept of ADNI (Khachaturian 2010). This background story demonstrates the power of bringing different stakeholders together to promote innovation. ADNI initially encompassed clinical sites in the US and Canada to identify neuroimaging measures and biomarkers with patients who have MCI and AD (Mueller et al. 2005). ADNI's methods standardization is being built up by other consortia like the Alzheimer's Association (Carrillo et al. 2013) and has since led to an umbrella organization called World Wide ADNI for the different countries involved in the initiative, including Europe, China, and South Korea (Burton 2011). Hundreds of scientists have made tens of thousands of downloads from ADNI's large database (Miller 2009). Eli Lilly used ADNI-style clinical trials for their 2 AD compounds (Miller 2009) and ADNI's work has contributed to recent revisions of diagnostic criteria for AD and MCI (Aisen 2011).

Appendix B: Expert biographies

To substantiate our research and lend evidence to our findings, we had informal discussions with key experts in the biomarker fields. Their biographies and some of the key issues discussed are listed below.

Shashi Amur, Ph.D

Shashi Amur is currently the Biomarker Qualification Science Coordinator in the office of Translational Sciences, CDER, FDA. She received her Ph.D. in biochemistry from Indian Institute of Science, India and completed post-doctoral fellowships at Temple University and at UCLA. She then gained experience in diagnostic and biotech sectors before joining FDA as a senior genomics reviewer in the Office of Clinical Pharmacology (OCP) and reviewed genomics- and biomarker-relevant sections of regulatory submissions. She has been an invited speaker at national and international conferences, and the author of 37 scientific publications. Her current research interests include pharmacogenomics, HLA-associated adverse events and biomarkers in Autoimmune Diseases and in Alzheimer's disease. She has served as Chair of the Pharmacogenomics Science Interest Group and Chair of OCP Science Day Committee at OCP, CDER, FDA and has organized seminars and workshops in CDER, FDA.

Jonathan Fleming, M.S.

Jonathan Fleming is the Managing Partner of Oxford Bioscience Partners, an international venture capital firm specializing in life science technology based investments, with offices in Boston, Massachusetts and Seoul, Korea. He is a General Partner of the Korea-Seoul Life Science Fund. Mr. Fleming has been in the investment business for over twenty seven years, starting and financing growth companies in the United States, Europe, Israel and Asia. Prior to joining OBP in 1996, he was a Founding General Partner of MVP Ventures in Boston, MA. He began his investment career with TVM Techno Venture Management in Munich, Germany. Mr. Fleming has also co-founded Medica Venture Partners, a venture capital investment firm specializing in early stage healthcare and biotechnology companies in Israel. Mr. Fleming is a director of several public companies including Xencor (NASDAQ: XNCR) and several private companies. He is also a director of Leerink Swann, a Boston based investment bank specializing in healthcare companies. In addition, Mr. Fleming is a Senior Lecturer at the MIT Sloan School of Business and a Member of the Board of NEHI, a healthcare oriented think tank and policy advocacy group. Mr. Fleming is on the Chancellor's Commercialization Advisory Board of the University of Texas System. Mr. Fleming holds a Master's degree in Public Administration from Princeton University and a Bachelor of Arts degree from the University of California, Berkeley.

Bill Helming

Bill Helming is a Managing Director of Arryve Consulting. Helming's career has concentrated on accelerating product development to help organizations gain competitive advantage and realize their full strategic intent. His consulting engagements have focused on achieving synergies in operational strategy, supply chain strategy and implementation, acquisition integration, R&D, technology transfer and manufacturing scale-up. Helming has also worked extensively in the public sector, implementing operational strategy to achieve important policy initiatives, such as strategic management for government-funded research and development portfolios devoted to infectious disease countermeasures for public health and biodefense. Prior to joining ARRYVE, Helming was partner and led the Life Sciences practice at PRTM Consultants, where he worked for 21 years before it became a division of PwC, where he was managing director. He earned his Master of Business Administration from the Robert H. Smith School of Business at the University of Maryland, Master of Science in Engineering from the Massachusetts Institute of Technology, and Bachelor of Science in Engineering from Webb Institute.

Huimin Kong, Ph.D

Huimin Kong serves as the President and Chief Executive Officer of Biohelix Corporation. BioHelix's mission is to improve the quality of healthcare through the development of simple molecular diagnostic tests for the near patient setting.

Andrea Schievella, Ph.D

Andrea Schievella is a scientist with 26 years of experience in cell and molecular biology, including significant experience in licensing and business development with biotechnology and pharmaceutical companies across the US and Europe. Her current role is a Business Manager at Cancer Research Technology, where she facilitates licensing and collaborative arrangements between the non-profit cancer charity Cancer Research UK and biotech and pharma companies working in cancer.

Sudhir Srivastava, Ph.D, MPH

Dr. Sudhir Srivastava has been Chief of the Cancer Biomarkers Research Group since 2000. His efforts focus on molecular biology of malignancies, early malignancies, risk assessment, and informatics, providing leadership in the areas of molecular screening and early detection. He is one of the principal authors of the Bethesda Guidelines for diagnosing Hereditary non-polyposis colorectal cancer.
He received several national/international honors and awards and is a member of a number of scientific committees. In 1995, he was elected to the American Joint Committee on Cancer, which is responsible for developing staging criteria for cancers for worldwide use, and serves on the Executive Committee—he was the first Asian American and non-MD to do so. Under his leadership, AJCC accepted the inclusion of tumor markers in the staging guidelines for colorectal cancer. He received several NIH and NCI honors and awards, has initiated and chaired state-of-the-science national and international-level workshops and conferences, and was principal architect of the first Gordon Research Conference on New Frontiers in Cancer Detection and Diagnosis. He initiated new areas of research (molecular signatures of infectious agents in cancer; micro-imaging in classifying preneoplastic lesions; nano-technology in Earlier Cancer Detection; Metabolomics and Glycomics Alliances with other NIH Institutes). He published more than 170 peer-reviewed papers and has edited four books. He is Editor-In-Chief for the journal *Disease Biomarkers*; led creation of the journal *Cancer Biomarkers*; was appointed to the editorial board of the journal *Cancer Prevention Research*; is a member of several scientific committees; founding member of HUPO; principal architect of NCI's Early Detection Research Network; and founding Editor and Board member for *Journal of Proteomics & Bioinformatics*. He has a PhD in biological science; MS in computer science; and MPH from the Johns Hopkins University.

Mark Trusheim, M.S.

Mark Trusheim is founder of Co-Bio Consulting, LLC as well as Executive in Residence and Visiting Scholar at the MIT Sloan School of Management and a Special Government Employee for the FDA's Office of the Commissioner. He is a former member of the Massachusetts Biotechnology Council's Board of Directors, which helps its over 500 members succeed in the state. In 2004 he further served as the Interim President of the MBC, leading its successful legislative agenda, its expansion of MassBioEd education programs and its continued membership growth. As an entrepreneur, Mark founded and was the first President and CEO of Cantata Laboratories. Cantata marketed clinical diagnostics and pharmaceutical biomarker services based on its biochemical profiling platform. Prior to Cantata, Mark spent over 10 years at Monsanto/Pharmacia, culminating his career there as Co-President and Chief Operating Officer of Cereon Genomics, LLC. He holds degrees in Chemistry from Stanford University and Management from MIT.

Andrey Zarur
Andrey Zarur is managing partner at Kodiak Venture Partners (KVP), a venture capital firm specializing in formation stage information and life technology investments. He has been active in early stage life sciences companies for over 15 years, and has actively participated in the creation of over a dozen companies in the healthcare and clean energy sectors. Prior to joining KVP, he was founder and chief executive officer of BioProcessors, a leading provider of high-throughput systems to the biopharmaceutical industry. In addition to BioProcessors, he has led four life science companies from inception to exit, and has been an active investor or board member on several others. Zarur is a co-founder and chairman of the board of GreenLight Biosciences and Lumicell. He also is chairman of the board of Allegro Diagnostics. Zarur is a director of several private companies, including Fluxion Biosciences, Astadia, and WTI, and is a director of several nonprofit organizations such as Infantia, a foundation dedicated to providing education to children in developing countries. Zarur holds a BS from the National University of Mexico and an SM and a PhD in chemical engineering from MIT.

Indicative Questions
1) What do you think are the biggest gaps/bottlenecks in the introduction of biomarkers in R&D and clinical practice?
2) What impact do you think the different stakeholders can have in terms of addressing these challenges?
3) How do you think the different stakeholders can be brought to work closer together?
4) In which ways do you think MSCs are useful? In which ways could they be improved? What are some metrics of success that you think would be important to evaluate these MSCs by?
5) Do you think the lack of harmonization in biomarker terminology and standards have hindered the adoption of innovative biomarkers in drug development and patient management?
6) Biomarker development is a very data-rigorous field. What would be the best way to address such challenges?
7) What are your thoughts on the BQP? Do you think that it can make a difference in providing a better regulatory path for biomarker use?
8) What are your thoughts on the reason behind the regulatory differences between the FDA and EMA?

Bibliography

Advameddx 2013, 'Infographic: Diagnostics Transforming Care'. *Advameddx*. Available at: http://advameddx.org/go.cfm?do=News.View&nid=14 [Accessed April 23, 2014].

Aisen, PS 2011, 'Q&A: The Alzheimer's Disease Neuroimaging Initiative.'. *BMC Medicine*, 9, p.101.

Allison, M 2010, 'Biomarker-led adaptive trial blazes a trail in breast cancer.'. *Nature Biotechnology*, 28(5), pp.383–4.

Altar, C, Amakye, D, Bounos, D, Bloom, J, Clack, G, Dean, R, Devanarayan, V, Fu, D, Furlong, S, Hinman, L, Girman, C, Lathia, C, Lesko, L, Madani, S, Mayne, J, Meyer, J, Raunig, D, Sager, P, Williams, S, Wong, P & Zerba, K 2008, 'A prototypical process for creating evidentary standards for biomarkers and diagnostics'. *Clinical Pharmacology & Therapeutics*, 83(2), pp.368–371.

Amur, S, Frueh, F, Lesko, L & Huang, S 2008, 'Integration and use of biomarkers in drug development, regulation and clinical practice: a US regulatory perspective'. *Biomarkers in Medicine*, 2(3), pp.305–311.

Anderson, D & Kodukula, K 2013, 'Biomarkers in pharmacology and drug discovery'. *Biochemical Pharmacology*, 87(1), pp.172–88.

Anon 2013, 'Medical Device Innovation Consortium'. Available at: http://mdic.org/ [Accessed April 27, 2014].

Atkinson, A, Colburn, W, DeGruttola, V, DeMets, D, Downing, G, Hoth, D, Oates, J, Peck, C, Schooley, R, Spilker, B, Woodcock, J & Zeger, S 2001, 'Biomarkers and surrogate endpoints: preferred definitions and conceptual framework.'. *Clinical Pharmacology & Therapeutics*, 69(3), pp.89–95.

Baggerly, KA & Coombes, KR 2011, 'What information should be required to support clinical "omics" publications?'. *Clinical Chemistry*, 57(5), pp.688–90.

Barker, A, Sigman, C, Kelloff, G, Hylton, N, Berry, D & Esserman, LJ 2009, 'I-SPY 2: an adaptive breast cancer trial design in the setting of neoadjuvant chemotherapy.'. *Clinical Pharmacology & Therapeutics*, 86(1), pp.97–100.

Barratt, RA, Bowens, SL, McCune, SK, Johannessen, JN & Buckman, SY 2012, 'The critical path initiative: leveraging collaborations to enhance regulatory science.'. *Clinical Pharmacology & Therapeutics*, 91(3), pp.380–3.

Bleavins, MR 2010, '*Biomarkers in Drug Development*', Wiley.

Burton, A 2011, 'Big science for a big problem: ADNI enters its second phase.'. *Lancet Neurology*, 10(3), pp.206–7.

CAMD 2010, '*FACT SHEET: Coalition Against Major Diseases*', Tuscon, AZ. Available at: http://c-path.org/pdf/CoalitionFactSheet.pdf.

Campion, S, Aubrecht, J, Boekelheide, K, David, W, Vaidya, VS, Anderson, L, Burt, D, Hwang, K, Pacheco, S & Saikumar, J 2013, 'The current status of biomarkers for predicting toxicity'. *Expert Opinion on Drug Metabolism & Toxicology*, 9(11), pp.1–28.

Carlsson, B, Jacobsson, S, Holmén, M & Rickne, A 2002, 'Innovation systems: analytical and methodological issues'. *Research Policy*, 31, pp.233–245.

Carrillo, M, Blennow, K, Soares, H, Lewczuk, P, Mattsson, N, Oberoi, P, Umek, R, Vandijck, M, Salamone, S, Bittner, T, Shaw, L, Stephenson, D, Bain, L & Zetterberg, H 2013, 'Global standardization measurement of cerebral spinal fluid for Alzheimer's disease: An update from the Alzheimer's Association Global Biomarkers Consortium'. *Alzheimer's & Dementia*, 9, pp.137–140.

Carroll, J 2012, 'The Top Phase III Disasters of 2012'. *FierceBiotech*. Available at: http://www.fiercebiotech.com/special-report/top-phase-iii-disasters-2012/2012-10-02 [Accessed March 9, 2014].

Carver, K 2010, 'Companion Diagnostics: Evolving FDA Regulation and Issues for Resolution'. In S. Danzis & E. Flannery, eds. *In Vitro Diagnostics: The Complete Regulatory Guide*. The Food and Drug Law Institute, pp. 149–184.

CDRH 2010, 'FDA/CDRH Public Meeting: Oversight of Laboratory Developed Tests (LDTs)'. In *Workshops & Conferences (Medical Devices)*. Center for Devices and Radiological Health. Available at: http://www.fda.gov/MedicalDevices/NewsEvents/WorkshopsConferences/ucm212830.htm [Accessed March 27, 2014].

Chataway, J, Fry, C, Marjanovic, S & Yaqub, O 2012, 'Public-private collaborations and partnerships in stratified medicine: making sense of new interactions'. *New Biotechnology*, 29(6), pp.732–40.

Chen, M, Vijay, V, Shi, Q, Liu, Z, Fang, H & Tong, W 2011, 'FDA-approved drug labeling for the study of drug-induced liver injury.'. *Drug Discovery Today*, 16(15-16), pp.697–703.

Clinton, BP & Wechsler, J 2006, 'What Ever Happened to Critical Path'. *Pharmaceutical Executive*, pp.1–7.

Crichton, DJ, Mattmann, CA, Thornquist, M, Anton, K & Hughes, JS 2010, 'Bioinformatics: biomarkers of early detection.'. *Cancer Biomarkers*, 9(1-6), pp.511–30.

Day, M, Rutkowski, J & Feuerstein, G 2009, 'Translational Medicine--A Paradigm Shift in Modern Drug Discovery and Development: The Role of Biomarkers'. In C. Guzman & G. Feuerstein, eds. *Pharmaceutical Biotechnology*. pp. 1–12.

Dennis, EH, Walker, EG, Baker, AF & Miller, RT 2013, 'Opportunities and Challenges of Safety Biomarker Qualification: Perspectives from the Predictive Safety Testing Consortium'. *Drug Development Research*, 74(2), pp.112–126.

Dieterle, F, Sistare, F, Goodsaid, F, Papaluca, M, Ozer, J, Webb, C, Baer, W, Senagore, A, Schipper, M, Vonderscher, J, Sultana, S, Gerhold, D, Phillips, J, Maurer, G, Carl, K, Laurie, D, Harpur, E, Sonee, M,

Ennulat, D, Holder, D, Andrews-Cleavenger, D, Gu, Y, Thompson, K, Goering, P, Vidal, J, Abadie, E, Maciulaitis, R, Jacobson-Kram, D, Defelice, A, Hausner, E, Blank, M, Thompson, A, Harlow, P, Throckmorton, D, Xiao, S, Xu, N, Taylor, W, Vamvakas, S, Flamion, B, Lima, B, Kasper, P, Pasanen, M, Prasad, K, Troth, S, Bounous, D, Robinson-Gravatt, D, Betton, G, Davis, M, Akunda, J, McDuffie, J, Suter, L, Obert, L, Guffroy, M, Pinches, M, Jayadev, S, Blomme, E, Beushausen, S, Barlow, V, Collins, N, Waring, J, Honor, D, Snook, S, Lee, J, Rossi, P, Walker, E & Mattes, W 2010, 'Renal biomarker qualification submission: a dialog between the FDA-EMEA and Predictive Safety Testing Consortium.'. *Nature Biotechnology*, 28(5), pp.455–62.

Drucker, E & Krapfenbauer, K 2013, 'Pitfalls and limitations in translation from biomarker discovery to clinical utility in predictive and personalised medicine.'. *The EPMA Journal*, 4(1), p.7.

EGAPP 2014, 'The EGAPP initiative: lessons learned.'. *Genetics in Medicine*, 16(3), pp.217–24.

EMA 2006, '*Guideline on Pharmacogenetics Briefing Meetings*', London, UK. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC5000 03886.pdf.

EMA 2014, 'Pharmacogenomics Working Party'. *EMA*. Available at: http://www.ema.europa.eu/ema/index.jsp?curl=pages/contacts/CHMP/people_listing_000018.jsp &mid=WC0b01ac0580028d91 [Accessed April 7, 2014].

Esserman, L, Berry, D, DeMichele, A, Carey, L, Davis, S, Buxton, M, Hudis, C, Gray, JW, Perou, C, Yau, C, Livasy, C, Krontiras, H, Montgomery, L, Tripathy, D, Lehman, C, Liu, MC, Olopade, OI, Rugo, HS, Carpenter, JT, Dressler, L, Chhieng, D, Singh, B, Mies, C, Rabban, J, Chen, Y, Giri, D, van 't Veer, L & Hylton, N 2012, 'Pathologic complete response predicts recurrence-free survival more effectively by cancer subset: results from the I-SPY 1 TRIAL'. *Journal of Clinical Oncology*, 30(26), pp.3242–9.

Esserman, LJ, Barker, AD, Woodcock, J, Buxton, MA, Berry, DA, Patterson, R, Jolly, MK, Demichele, A, Hylton, N, Rubin, EH, Parkinson, D, Wholley, D, van 't Veer, L, Yee, D, Park, J, Tripathy, D, Perlmutter, J, Buetow, K, Hogarth, M, Gray, J & Dilts, DM 2012, 'A Model for Accerlerating Identification and Regulatory Approval of Effective Investigational Agents'. *Cureus*, 4(12), pp.1–17.

Esserman, LJ, Buxton, M & Berry, DA 2013, 'Accelerating approval of promising agents in the neoadjuvant setting: a path for facilitating targeted drug development for breast cancer'. *Breast Diseases: A Year Book Quarterly*, 24(2), pp.126–128.

Esserman, LJ & Woodcock, J 2011, 'Accelerating Identification and Regulatory Approval of Investigational Cancer Drugs'. *JAMA: The Journal of the American Medical Association*, 306(23), pp.2608–2609.

Etzkowitz, H & Leydesdorff, L 1995, 'The Triple Helix: University-Industry-Government Relations: A Laboratory for Knowledge-Based Economic Development'. *EASST Review*, 14, pp.14–19.

Evans, BJ 2010, 'Establishing clinical utility of pharmacogenetic tests in the post-FDAAA era.'. *Clinical Pharmacology & Therapeutics*, 88(6), pp.749–51.

FDA 2014, 'Guidance for Industry and FDA Staff: Qualification Process for Drug Development Tools', Silver Spring, MD. Available at: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM230597.pdf.

FDA 2012, 'Guidance for Industry: Pathologic Complete Response in Early-Stage Breast Cancer: Use as an Endpoint to Support Accelerated Approval', Available at: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM305501.pdf.

FDA 2013a, 'Medical Device Development Tools: Draft Guidance for Industry, Tool Developers, and Food and Drug Administration Staff', Available at: http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM374432.pdf.

FDA 2010, 'Number of voluntary exploratory data submissions (VXDS) and biomarker qualification reviews completed per month'. FDA-TRACK: Agency-wide Program Performance. Available at: http://www.accessdata.fda.gov/FDATrack/track?program=cder&id=CDER-OCP-VXDS-Reviews-Completed [Accessed April 7, 2014].

FDA 2013b, 'Paving the Way for Personalized Medicine', Available at: http://www.fda.gov/downloads/scienceresearch/specialtopics/personalizedmedicine/ucm372421.pdf.

Fox, JL 2014, 'Master Protocol for squamous cell lung cancer readies for launch.'. Nature Biotechnology, 32(2), pp.116–8.

Frueh, FW, Rudman, A, Simon, K, Gutman, S, Reed, C & Dorner, AJ 2006, 'Experience with voluntary and required genomic data submissions to the FDA: summary report from track 1 of the third FDA-DIA-PWG-PhRMA-BIO pharmacogenomics workshop.'. The Pharmacogenomics Journal, 6(5), pp.296–300.

Gerard, B, Betard, M, Smith, B & Denyer, M 2012, 'Development of targeted agents and companion diagnostics'. Drug Information Journal, 46(4), pp.405–419.

Gibbs, J 2010, 'Regulatory Pathways for Clearance or Approval of IVDs'. In S. Danzis & E. Flannery, eds. In Vitro Diagnostics: The Complete Regulatory Guide. Food and Drug Law Institute, pp. 43–68.

Goldman, M 2012, 'The innovative medicines initiative: a European response to the innovation challenge.'. Clinical Pharmacology & Therapeutics, 91(3), pp.418–25.

Goldman, M, Compton, C & Mittleman, BB 2013, 'Public-private partnerships as driving forces in the quest for innovative medicines'. Clinical and Translational Medicine, 2(2), pp.1–3.

Goodsaid, F, Amur, S, Aubrecht, J, Burczynski, M, Carl, K, Catalano, J, Charlab, R, Close, S, Cornu-Artis, C, Essioux, L, Fornace, A, Hinman, L, Hong, H, Hunt, I, Jacobson-Kram, D, Jawaid, A, Laurie, D, Lesko, L, Li, H, Lindpaintner, K, Mayne, J, Morrow, P, Papaluca-Amati, M, Robison, T, Roth, J, Schuppe-

Koistinen, I, Shi, L, Spleiss, O, Tong, W, Truter, S, Vonderscher, J, Westelinck, A, Zhang, L & Zineh, I 2010, 'Voluntary exploratory data submissions to the US FDA and the EMA: experience and impact.'. *Nature Reviews. Drug Discovery*, 9(6), pp.435–45.

Goodsaid, F & Frueh, F 2007, 'Biomarker Qualification Pilot Process at the US Food and Drug Administration'. *AAPS Journal*, 9(1), pp.105–108.

Goodsaid, F & Frueh, F 2006, 'Process map proposal for the validation of genomic biomarkers'. *Pharmacogenomics*, 7(5), pp.773–782.

Goodsaid, F, Frueh, F & Mattes, W 2008, 'Strategic paths for biomarker qualification'. *Toxicology*, 245, pp.219–223.

Goodsaid, F & Mattes, W 2013, '*The Path from Biomarker Discovery to Regulatory Qualification*', Elsevier Inc.

Goodsaid, F & Mendrick, D 2010, 'Translational medicine and the value of biomarker qualification.'. *Science Translational Medicine*, 2(47), pp.1–3.

Goodsaid, F & Papaluca, M 2010, 'Evolution of biomarker qualification at the health authorities.'. *Nature Biotechnology*, 28(5), pp.441–3.

Gutman, S & Kessler, L 2006, 'The US Food and Drug Administration perspective on cancer biomarker development.'. *Nature Reviews. Cancer*, 6(7), pp.565–71.

Halim, A 2011, 'Biomarkers in Drug Development: A Useful Tool but Discrepant Results May Have a Major Impact'. In *Drug Discovery and Development - Present and Future*. pp. 401–424. Available at: http://www.intechopen.com/books/drug-discovery-and-development-present-and-future/biomarkers-in-drug-development-a-useful-tool-but-discrepant-results-may-have-a-major-impact.

Halim, A 2009, 'Impact of discrepant results from clinical laboratories on patients and pharmaceutical trials: evidence from proficiency testing results'. *Biomarkers in Medicine*, 3(3), pp.231–238.

Hausner, EA, Hicks, KA, Leighton, JK, Szarfman, A, Thompson, AM & Harlow, P 2013, 'Qualification of cardiac troponins for nonclinical use: A regulatory perspective'. *Regulatory Toxicology and Pharmacology*, 67, pp.108–114.

Hayes, DF, Allen, J, Compton, C, Gustavsen, G, Leonard, DGB, McCormack, R, Newcomer, L, Pothier, K, Ransohoff, D, Schilsky, RL, Sigal, E, Taube, SE & Tunis, SR 2013, 'Breaking a vicious cycle.'. *Science Translational Medicine*, 5(196), pp.1–7.

Hockett, RD & Close, SL 2010, 'Regulation of laboratory-developed tests: the case for utilizing professional associations.'. *Clinical Pharmacology & Therapeutics*, 88(6), pp.743–5.

Hollmer, M 2013, 'FDA, CMS extend little-used pilot program that grants parallel review for some devices'. *FierceMedicalDevices*. Available at: http://www.fiercemedicaldevices.com/story/fda-cms-

extend-little-used-pilot-program-grants-parallel-review-some-device/2013-12-19 [Accessed March 27, 2014].

IOM 2010a, '*Evaluation of Biomarkers and Surrogate Endpoints in Chronic Disease*' C. M. Micheel & J. R. Ball, eds., National Academies Press.

IOM 2010b, '*Extending the Spectrum of Precompetitive Collaboration in Oncology Research: Workshop Summary*' IOM, ed., National Academies Press.

IOM 2008, '*Neuroscience Biomarkers and Biosignatures*', The National Academies Press.

Keeling, P & Roth, M 2008, '*IP Considerations in Early Biomarker Development in the Personalized Medicine Age*', Available at: http://www.diaceutics.com/sites/default/files/publications/IP considerations in early Biomarker development in PM.pdf.

Kelloff, GJ & Sigman, CC 2012, 'Cancer biomarkers: selecting the right drug for the right patient.'. *Nature Reviews. Drug Discovery*, 11(3), pp.201–14.

Khachaturian, ZS 2010, 'Perspective on the Alzheimer's disease neuroimaging initiative: progress report and future plans.'. *Alzheimer's & Dementia*, 6(3), pp.199–201.

Khleif, SN, Doroshow, JH & Hait, WN 2010, 'AACR-FDA-NCI Cancer Biomarkers Collaborative consensus report: advancing the use of biomarkers in cancer drug development.'. *Clinical Cancer Research*, 16(13), pp.3299–318.

Lee, J, Devanarayan, V, Barrett, YC, Weiner, R, Allinson, J, Fountain, S, Keller, S, Weinryb, I, Green, M, Duan, L, Rogers, J a, Millham, R, O'Brien, PJ, Sailstad, J, Khan, M, Ray, C & Wagner, J 2006, 'Fit-for-purpose method development and validation for successful biomarker measurement.'. *Pharmaceutical Research*, 23(2), pp.312–28.

Lesko, L 2007, 'Paving the Critical Path: How can Clinical Pharmacology Help Achieve the Vision?'. *Clinical Pharmacology & Therapeutics*, 81(2), pp.170–177.

Lesko, L & Atkinson, A 2001, 'Use of biomarkers and surrogate endpoints in drug development and regulatory decision making: criteria, validation, strategies'. *Annual Review of Pharmacology Toxicology*, 41, pp.347–66.

Mahajan, R & Gupta, K 2010, 'Food and drug administration's critical path initiative and innovations in drug development paradigm: Challenges, progress, and controversies.'. *Journal of Pharmacy & Bioallied Sciences*, 2(4), pp.307–13.

Manolis, E, Vamvakas, S & Isaac, M 2011, 'New pathway for qualification of novel methodologies in the European medicines agency'. *Proteomics. Clinical Application*, 5, pp.248–255.

Mansfield, EA 2014, 'FDA Perspective on Companion Diagnostics: An Evolving Paradigm'. *Clinical Cancer Research*, 20(6), pp.1453–1457.

Marrer, E & Dieterle, F 2007, 'Promises of Biomarkers in Drug Development – A Reality Check'. *Chemical Biology & Drug Design*, 69, pp.381–394.

Mattes, WB, Walker, EG, Abadie, E, Sistare, FD, Vonderscher, J, Woodcock, J & Woosley, RL 2010, 'Research at the interface of industry, academia and regulatory science'. *Nature Biotechnology*, 28(5), pp.432–433.

Miller, G 2009, 'Alzheimer's biomarker initiative hits its stride'. *Science*, 326, pp.386–389.

Mittleman, B, Neil, G & Cutcher-Gershenfeld, J 2013, 'Precompetitive consortia in biomedicine--how are we doing?'. *Nature Biotechnology*, 31(11), pp.979–85.

Mueller, SG, Weiner, MW, Thal, LJ, Petersen, RC, Jack, C, Jagust, W, Trojanowski, JQ, Toga, AW & Beckett, L 2005, 'The Alzheimer's disease neuroimaging initiative.'. *Neuroimaging Clinics of North America*, 15(4), pp.869–77, xi–xii.

Mullard, A 2014a, '2013 FDA drug approvals.'. *Nature reviews. Drug discovery*, 13(2), pp.85–9.

Mullard, A 2014b, 'Multicompany trials adapt to disciplines beyond cancer.'. *Nature medicine*, 20(1), p.3.

NCI 2012, '*NCI-FDA-NIST Workshop on Standards in Molecular Diagnostics for the Discovery and Validation of Clinically Useful Cancer Biomarkers*', Rockville, MD. Available at: http://prevention.cancer.gov/files/groups/NCI-FDA-NIST_Biomarkers_Workshop_Report.pdf.

Opar, A 2011, 'Bridging the drug-diagnostic divide.'. *Nature Reviews. Drug Discovery*, 10(5), pp.323–4.

Orr, M, Goodsaid, F, Amur, S, Rudman, A & Frueh, F 2007, 'The experience with voluntary genomic data submissions at the FDA and a vision for the future of the voluntary data submission program.'. *Clinical Pharmacology & Therapeutics*, 81(2), pp.294–7.

Pammolli, F, Magazzini, L & Riccaboni, M 2011, 'The productivity crisis in pharmaceutical R&D.'. *Nature Reviews. Drug Discovery*, 10(6), pp.428–38.

Parkinson, DR, McCormack, RT, Keating, SM, Gutman, SI, Hamilton, SR, Mansfield, EA, Piper, MA, DeVerka, P, Frueh, FW, Jessup, JM, McShane, LM, Tunis, SR, Sigman, CC & Kelloff, GJ 2014, 'Evidence of Clinical Utility: An Unmet Need in Molecular Diagnostics for Patients with Cancer'. *Clinical Cancer Research*, 20(6), pp.1428–1444.

Patlak, M 2010, 'Competitors try collaboration to speed drug development.'. *Journal of the National Cancer Institute*, 102(12), pp.841–3.

Paul, SM, Mytelka, DS, Dunwiddie, CT, Persinger, CC, Munos, BH, Lindborg, SR & Schacht, AL 2010, 'How to improve R&D productivity: the pharmaceutical industry's grand challenge.'. *Nature Reviews. Drug Discovery*, 9(3), pp.203–14.

Pennello, G 2013, 'Analytical and clinical evaluation of biomarkers assays: when are biomarkers ready for prime time?'. *Clinical Trials*, 10(5), pp.666–76.

Pepe, MS, Etzioni, R, Feng, Z, Potter, JD, Lou, M, Thornquist, M, Winget, M & Yasui, Y 2001, 'Phases of Biomarker Development for Early Detection of Cancer'. *Journal of the National Cancer Institute*, 93(14), pp.1054–1061.

Pignatti, F, Ehmann, F, Hemmings, R, Jonsson, B, Nuebling, M, Papaluca-Amati, M, Posch, M & Rasi, G 2014, 'Cancer Drug Development and the Evolving Regulatory Framework for Companion Diagnostics in the European Union'. *Clinical Cancer Research*, 20(6), pp.1458–1468.

Poste, G 2012, 'Biospecimens, biomarkers, and burgeoning data : the imperative for more rigorous research standards'. *Trends in Molecular Medicine*, 18(12), pp.718–723.

Poste, G, Carbone, DP, Parkinson, DR, Verweij, J, Hewitt, SM & Jessup, JM 2012, 'Leveling the playing field: bringing development of biomarkers and molecular diagnostics up to the standards for drug development.'. *Clinical Cancer Research*, 18(6), pp.1515–23.

Prensner, JR, Chinnaiyan, AM & Srivastava, S 2012, 'Systematic, evidence-based discovery of biomarkers at the NCI.'. *Clinical & Experimental Metastasis*, 29(7), pp.645–52.

PRNewswire 2013, 'Vermillion's OVA1 Receives New Statement by Society of Gynecologic Oncology'. *PRNewswire*. Available at: http://www.prnewswire.com/news-releases/vermillions-ova1-receives-new-statement-by-society-of-gynecologic-oncology-210984641.html [Accessed April 14, 2014].

Prowell, T & Pazdur, R 2012, 'Pathological complete response and accelerated drug approval in early breast cancer'. *New England Journal of Medicine*, 366(26), pp.2438–2441.

Quinn, B 2010, 'Payers and the assessment of clinical utility for companion diagnostics.'. *Clinical Pharmacology & Therapeutics*, 88(6), pp.751–4.

Ranga, M & Etzkowitz, H 2013, 'Triple Helix systems: an analytical framework for innovation policy and practice in the Knowledge Society'. *Industry and Higher Education*, 27(4), pp.237–262.

Ray, T 2013, 'After Long Silence on LDT Regulation, FDA Commish Revives Thorny Topic at ASCO'. *Phamacogenomics Reporter*. Available at: http://www.genomeweb.com/after-long-silence-ldt-regulation-fda-commish-revives-thorny-topic-asco.

Rinaldi, A 2011, 'Teaming up for biomarker future'. *EMBO Reports*, 12(6), pp.500–4.

Rogers, M 2013, 'AD Trial Simulation Tool Receives Regulators' Blessings'. *Alzforum*. Available at: http://www.alzforum.org/news/research-news/ad-trial-simulation-tool-receives-regulators-blessings?id=3542 [Accessed April 7, 2014].

Romero, K, Mars, MD, Frank, D, Anthony, M, Neville, J, Kirby, L, Smith, K & Woosley, RL 2009, 'The Coalition Against Major Diseases: Developing Tools for an Integrated Drug Development Process for Alzheimer's and Parkinson's Diseases'. *Clinical Pharmacology & Therapeutics*, 86(4), pp.365–368.

Schully, SD, Benedicto, CB, Gillanders, EM, Wang, SS & Khoury, MJ 2011, 'Translational research in cancer genetics: the road less traveled.'. *Public Health Genomics*, 14(1), pp.1–8.

Senderowicz, AM & Pfaff, O 2014, 'Similarities and Differences in the Oncology Drug Approval Process between FDA and European Union with Emphasis on In Vitro Companion Diagnostics.'. *Clinical Cancer Research*, 20(6), pp.1445–52.

Sinha, V 2013, '*Model-based Drug Development in Neurodegenerative Diseases: Regulatory Perspectives*', Available at: http://c-path.org/wp-content/uploads/2013/11/Session-I_CAMD-Annual-Meeting_2013_Sinha.pdf.

Sistare, F, Dieterle, F, Troth, S, Holder, D, Gerhold, D, Andrews-Cleavenger, D, Baer, W, Betton, G, Bounous, D, Carl, K, Collins, N, Goering, P, Goodsaid, F, Gu, Y, Guilpin, V, Harpur, E, Hassan, A, Jacobson-Kram, D, Kasper, P, Laurie, D, Lima, B, Maciulaitis, R, Mattes, W, Maurer, G, Obert, L, Ozer, J, Papaluca-Amati, M, Phillips, J, Pinches, M, Schipper, M, Thompson, K, Vamvakas, S, Vidal, J, Vonderscher, J, Walker, E, Webb, C & Yu, Y 2010, 'Towards consensus practices to qualify safety biomarkers for use in early drug development'. *Nature Biotechnology*, 28(5), pp.446–455.

Srivastava, S 2013, 'The early detection research network: 10-year outlook.'. *Clinical Chemistry*, 59(1), pp.60–7.

Stephenson, D 2013, '*Coalition Against Major Diseases: Viewing the 2013 Landscape*', Available at: http://c-path.org/wp-content/uploads/2013/11/Intro_11-12-2013-DStephenson.pdf.

Stephenson, D, Bain, L, Brumfield, M, Carrillo, M, Comery, T, Compton, C, Corrigan, B, Gordon, M, Jack, C, Katz, R, Logovinsky, V, Satlin, A, Marek, K, Nicholas, T, Polhamus, D, Angersbach, B, Raghavan, N, Romano, G, Romero, K, Shaw, L, Woodcock, J, Vradenburg, G & Isaac, M 2013, 'Coalition Against Major Diseases: Precompetitive Collaborations and Regulatory Paths to Accelerating Drug Development for Neurodegenerative Diseases'. *Therapeutic Innovation & Regulatory Science*, 47(6), pp.632–638.

Teutsch, SM, Bradley, L a, Palomaki, GE, Haddow, JE, Piper, M, Calonge, N, Dotson, WD, Douglas, MP & Berg, AO 2009, 'The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Initiative: methods of the EGAPP Working Group.'. *Genetics in Medicine*, 11(1), pp.3–14.

Trusheim, MR, Austin, MJF, Rausch, C & Berndt, ER 2013, 'Uncertain prognosis for high-quality diagnostics: clinical challenges, economic barriers and needed reforms.'. *Pharmacogenomics*, 14(3), pp.325–34.

Trusheim, MR & Berndt, ER 2012, 'Economic challenges and possible policy actions to advance stratified medicine'. *Personalized Medicine*, 9(4), pp.413–427.

Vaidya, VS & Bonventre, J V 2010, '*Biomarkers in Medicine, Drug Discovery, and Environmental Health*', Wiley.

Vermillion 2013, 'Vermillion Reports Fourth Quarter and Full Year 2013 Results'. *Vermillion*. Available at: http://phx.corporate-ir.net/phoenix.zhtml?c=121814&p=irol-newsArticle&ID=1906975&highlight= [Accessed April 14, 2014].

Wagner, J 2008, 'Strategic approach to fit-for-purpose biomarkers in drug development.'. *Annual Review of Pharmacology and Toxicology*, 48, pp.631–51.

Wagner, J, Prince, M, Wright, E, Ennis, M, Kochan, J, Nunez, D, Schneider, B, Wang, M, Chen, Y, Ghosh, S, Musser, B & Vassileva, M 2010, 'The Biomarkers Consortium: practice and pitfalls of open-source precompetitive collaboration'. *Clinical Pharmacology & Therapeutics*, 87(5), pp.539–542.

Wagner, J, Williams, S & Webster, C 2007, 'Biomarkers and surrogate end points for fit-for-purpose development and regulatory evaluation of new drugs.'. *Clinical Pharmacology & Therapeutics*, 81(1), pp.104–7.

Wagner, J, Wright, E, Ennis, M, Prince, M, Kochan, J, Nunez, D, Schneider, B, Wang, M, Chen, Y, Ghosh, S, Musser, B & Vassileva, M 2009, 'Utility of adiponectin as a biomarker predictive of glycemic efficacy is demonstrated by collaborative pooling of data from clinical trials conducted by multiple sponsors.'. *Clinical Pharmacology & Therapeutics*, 86(6), pp.619–25.

Wagner, P & Srivastava, S 2012, 'New paradigms in translational science research in cancer biomarkers'. *Translational Research*, 159(4), pp.343–353.

Walker, EG, Brumfield, M, Compton, C & Woosley, R 2013, 'Evolving Global Regulatory Science Through the Voluntary Submission of Data: A 2013 Assessment'. *Therapeutic Innovation & Regulatory Science*.

Weiss, RL 2012, 'The long and winding regulatory road for laboratory-developed tests.'. *American Journal of Clinical Pathology*, 138(1), pp.20–6.

Williams, PM, Lively, TG, Jessup, JM & Conley, BA 2012, 'Bridging the gap: moving predictive and prognostic assays from research to clinical use.'. *Clinical Cancer Research*, 18(6), pp.1531–9.

Wilson, C, Schultz, S & Waldman, S 2007, 'Cancer biomarkers: where medicine, business, and public policy intersect'. *Biotechnology Healthcare*, 4(1), pp.33–42.

Woodcock, J 2010, 'Assessing the Clinical Utility of Diagnostics Used in Drug Therapy'. *Clinical Pharmacology & Therapeutics*, 88(6), pp.765–774.

Woodcock, J, Buckman, S, Goodsaid, F & Walton, M 2011, 'Qualifying biomarkers for use in drug development: a US Food and Drug Administration overview'. *Expert Opinion on Medical Diagnostics*, 5(5), pp.369–374.

Woodcock, J & Woosley, R 2008, 'The FDA critical path initiative and its influence on new drug development.'. *Annual Review of Medicine*, 59, pp.1–12.

Yu, H 2014, 'FDA weighs in on Drug Development Tools'. *BioClinica*. Available at:
http://www.bioclinica.com/blog/fda-weighs-drug-development-tools [Accessed April 25, 2014].

Zineh, I & Lesko, L 2009, 'Pharmacogenetics in medicine: barriers, critical factors and a framework for
dialogue'. *Personalized Medicine*, 6(4), pp.359–361.