THE FUTURE OF COVERAGE AND PAYMENT FOR PERSONALIZED MEDICINE DIAGNOSTICS

PMC
PERSONALIZED MEDICINE COALITION

EXECUTIVE SUMMARY

THE FUTURE OF COVERAGE AND PAYMENT FOR PERSONALIZED MEDICINE DIAGNOSTICS

EXECUTIVE SUMMARY

The past decade has seen incredible progress for the scope and impact of personalized medicine. Just ten years ago, tamoxifen and trastuzumab were among the few drugs that were routinely paired with a companion diagnostic test. Today, it seems that the FDA approves new precision-targeted therapeutics nearly every month, and the medical development pipeline includes hundreds of targeted therapies to come. Impressive as this is, drugs targeted using combination diagnostic tests are not the only channel through which personalized medicine can become a reality. Genetic tests are increasingly available to diagnose both children and adults with difficult clinical presentations, such as metabolic and mitochondrial diseases. Non-invasive ultrasensitive genomic prenatal testing for trisomy disorders such as Down's syndrome is also rapidly replacing traditional amniocentesis methods, and amazingly sophisticated cartridge-based kits running on platforms the size of a laser printer can now provide molecular identification of both viral and bacterial diseases. Other technological achievements include the development of multi-analyte assays with algorithms, which help physicians plan the management and in many cases reduce the overtreatment of diseases as diverse as breast, prostate, ovarian, and thyroid cancers. Dozens of papers have surfaced on the potential role of next-generation sequencing in providing a "real-time" genomic map of cancer that could be used to support clinical decision-making.

This has the potential to rapidly facilitate a much deeper understanding of the mechanisms of chemotherapy resistance and point to actionable, effective ways to limit it. The aging of the American population is a stimulus for a longer-term and more progressive approach to decision-making in medical care. Personalized medicine can control costs and improve outcomes, and the increased attention on the future of Medicare will demand creative approaches to accomplishing these goals.

This white paper summarizes our progress and outlines the challenges we will face in the coming years. We foresee three major challenges on the horizon:

- Imminent federal pricing of highly innovative molecular tests
- Inconsistent standards and paradigms for evaluating diagnostic, prognostic, and predictive genomic tests
- Lack of incentives for genomic medicine

Our call to action is to monitor, evaluate, and contribute to the debate in these three areas to ensure that personalized medicine is integrated into the development of Medicare policy in the years ahead.

The Personalized Medicine Coalition (PMC) would like to thank Bruce Quinn, M.D., Foley Hoag LLP, the principle author of this report, for his contributions, as well as those of our members who took the time to read and comment on this paper's many drafts.

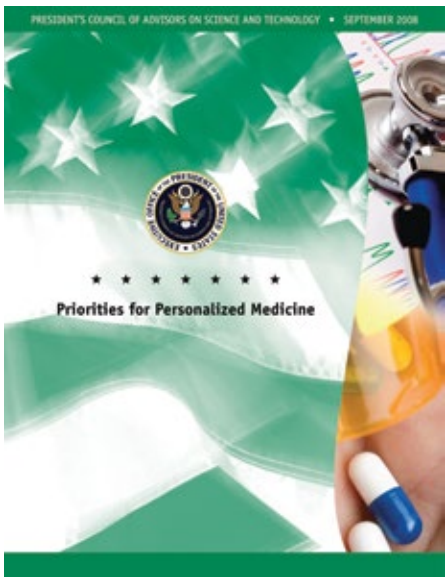
CONTENTS

A DECADE OF GREAT PROGRESS FOR PERSONALIZED MEDICINE

The Personalized Medicine Coalition was launched in 2004. Today, a decade later, it represents a broad spectrum of more than 230 stakeholders in the personalized medicine field. Its membership includes young innovator companies, academic medical centers and medical schools, and some of the largest companies in the pharmaceutical and in vitro diagnostic industries as well as patient, provider, and payer organizations.

In 2008, about midway through PMC's ten-year history, the President's Council of Advisors on Science and Technology (PCAST) issued a 77-page report on personalized medicine.[1] The PCAST viewed personalized medicine as a new endeavor that was exceptionally technology-driven and intimately patient-centered. As the Council wrote, personalized medicine arrives at the "convergence of scientific and clinical opportunity and public health need." Specifically, the report stated that personalized medicine holds unique promise to positively impact "the increasing cost of health care and the decreasing rate of new medical product development."

There is no question that every new advance in personalized medicine essentially aims to either reduce the cost of health care or provide value-added benefits for patients. But since 2008, the speed and breadth of the development of personalized medicine has exceeded even the promise envisioned by the PCAST. Just ten years ago, tamoxifen and trastuzumab were among the few drugs routinely paired with a companion diagnostic test. The panel expected that the pipeline of drug clinical trials underway by 2008 would soon reach approval and market access for patient care. Today, only six years later, new precision-targeted therapeutics are being approved by the FDA in steady succession, and hundreds of targeted therapeutics are now in Phase I, II, and III trials internationally. This is remarkable progress since President Bill Clinton and Prime Minister Tony Blair announced the near-completion of the initial human genome project in June 2000.[2]



PCAST report

PCAST's report also emphasized that drugs targeted by companion diagnostic tests are not the *only* forefront for personalized medicine. Other areas of genomic medicine moving forward at a record pace across the many areas of health care include:

New horizons for traditional genetic testing. Today, genetic tests are increasingly available for diagnosing difficult clinical cases, such as metabolic and mitochondrial disease. Broad gene panels of 10 to 30 or more gene candidates can be performed routinely and efficiently.[3]

Non-invasive maternal blood testing. In the past two years, non-invasive prenatal testing for trisomy disorders, such as Down's syndrome, has begun to rapidly replace traditional amniocentesis methods.[4]

Revolution in the molecular diagnostics of infectious disease. Genomic tests for infectious diseases, both bacterial and viral, can now be packaged as nearly complete test systems (including controls) that can rapidly diagnose a dozen or more viral causes of respiratory illness and identify potential hospital-acquired infections.

Multi-analyte assays with algorithms (MAAAs). Other technological achievements include the development of MAAAs, which help physicians plan the management and in many cases reduce the overtreatment of diseases as diverse as breast, prostate, colon, ovarian, and thyroid cancers. Several of these tests are FDA approved, which expands their potential as widely distributed platform-based kits.[5]

Next-generation sequencing in tumors and in cell-free circulating tumor DNA. Dozens of papers have focused on the potential role of next-generation sequencing in providing a genomic map of cancer.[6] This has the potential to expedite our understanding of the mechanisms of developing chemotherapy resistance and to create effective ways to limit tumor growth.[7,8] Depending on the indication, the analysis may be based on evaluating tumor tissue samples or circulating tumor-cell DNA.[9]

The previous examples are by no means exhaustive. There is now a range of journals devoted to personalized medicine and pharmacogenomics, including *Personalized Medicine*, *Pharmacogenomics*, *Journal of Personalized Medicine*, *The Journal of Molecular Diagnostics*, *Pharmacogenetics and Genomics*, *Genome Medicine*, and *Genetics in Medicine*. In addition to the content found in these specialty journals, breakthroughs in personalized medicine are covered in prestigious journals such as the *New England Journal of Medicine* and the *Journal of the American Medical Association*. The five previously described medical topics demonstrate that genomic medicine is having a broad and rapid impact on the delivery of health care, both in the U.S. and internationally.[10] In addition, *GenomeWeb*, an online news service for genomics and genomic medicine that is updated daily, has more than 120,000 subscribers and publishes more than 5,000 articles per year on personalized medicine.[11] Recent progress is truly phenomenal, and would have seemed like science fiction a decade ago.

PROBLEMS FOR PERSONALIZED MEDICINE ARE FORESEEABLE

Unquestionably the overall escalation of health care costs is a major societal issue. The drivers of rising health care costs include the relatively labor-intensive nature of most health care services and the increasing percentage of our population over 65. The *relative* costs of labor-intensive services like health care and education in society's budget rise disproportionately as other types of production become more mechanized and efficient,[12] and the *majority* of health care services are needed by the elderly. The percentage of our population older than 65—and even those older than 85—is rising rapidly and will continue to do so for the foreseeable future.[13]

While these trends make health care relatively costly, personalized medicine provides a *technology* that makes the whole system of health care more *efficient*. Personalized medicine can control health care costs by efficiently identifying the right drug for a patient—including, where appropriate, the right inexpensive generic drug. Personalized medicine can also reduce health care costs if molecular tests are performed on non-invasive samples, such as peripheral blood, or if a few cells are taken by fine needle aspiration, as these types of tests can eliminate the need for surgeries that can cost $20,000 or more. Personalized medicine can ultimately control costs and improve outcomes when we know the right cocktail of drugs to provide long-term disease control. This is well-established in the case of protease inhibitor and other anti-retroviral drug cocktails for HIV. For patients using these drugs, molecular diagnostics play a fundamental role in effective management. Initially, drugs can be selected based on molecular predictive profiling of the HIV virus affecting a particular patient; subsequently and on a monthly basis, the effectiveness of the drugs can be rapidly ascertained by measuring HIV viral load in peripheral blood. Although the use of molecular diagnostics is rising, it still represents only a fraction of the total cost of clinical laboratory tests.

Several potential crises in molecular medicine are *foreseeable* today.[14] We believe that by articulating these problems clearly and continuing to organize stakeholders to seek solutions, we can ensure personalized medicine's continued progress.

WE BELIEVE THE MAJOR PUBLIC POLICY CONCERNS CAN BE ORGANIZED IN THREE CATEGORIES. THESE ARE:

1. Imminent Federal Pricing of Highly Innovative Molecular Tests

Between 2012 and 2014, there has been a wave of administrative changes in the "coding and pricing" paradigms for genomic tests. "Coding" and "administrative pricing" for laboratory diagnostic tests are handled in a unique way by the Centers for Medicare & Medicaid Services (CMS)' Medicare program and, by extension, most other payers. As this paper will describe, changes to molecular diagnostic reimbursements in calendar year 2013 were often rocky, and extended as far as a nearly complete cessation of federal payments for genomic tests in the first quarter of 2013.

However, even larger policy decisions will face CMS in future years. For many traditional genetic diagnostic tests, Medicare lowered its effective payments between 2012 and 2013. Although lower prices can reflect more efficient and more widespread technologies, it is crucial that reimbursement levels, which have been set unilaterally by policymakers, will not only ensure access to high quality tests but also continue to allow for the development of a pipeline of innovative tests that require substantial risk-based research. No entity in the U.S. is as influential as Medicare, which normally prices payments (other than for drugs) as closely as possible to the marginal cost of the provider. But innovators require enough revenue to both provide the product and recoup their development costs. Making wise policy decisions in this area is the single biggest challenge for the Medicare program and all its stakeholders. Under new legislation passed in April 2014, Medicare will establish new rules that will generally set Medicare payments equivalent to market rates for all laboratory tests. All stakeholders depend on a successful transition to this new payment setting system.

2. Inconsistent Standards and Paradigms for Evaluating Diagnostic, Prognostic, and Predictive Genomic Tests

Health technology assessments represent a rapidly growing area of international government policy, which has accelerated in the past several years as governments seek to contain costs. In the U.S., for example, the Agency for Healthcare Research and Quality (AHRQ) has seen a marked increase in its funding, and new entities, such as the Patient Centered Outcomes Research Institute (PCORI), have been established. Most of the paradigms for health care appraisal and technology assessment were developed for the purpose of comparing procedures or drug therapies. Appraisal methods are also established for "traditional"

diagnostic tests, where relatively simple statistics express the accuracy of diagnosis (e.g. true positive, false positive) against a single gold standard. Today, tests are evaluated for a much broader concept of their "clinical utility." Although most agree that genomic tests create a benefit by having an impact both on patient management and the delivery of treatments, many are still concerned that processes for "clinical utility" assessments are neither clear nor predictable. More objective and reliable standards for these evaluation processes need to become broadly accepted. Since many argue that reimbursement should be tied to value, value (utility for the patient) needs to be acceptably defined.

3. Lack of Incentives for Genomic Medicine

There will also be areas where genomic medicine could have a major impact on public health but unfortunately traditional funding, pricing, or reimbursement systems fail to provide enough incentive for its development. These areas include funding the education of physicians and patients in personalized medicine, funding allied professionals such as genetic counselors, and creating incentives to develop new tools that could revolutionize some therapeutic areas. For example, because of the complexity of human drug metabolism, in many cases the role of any individual gene in the metabolism of a single drug may have a limited but meaningful impact on clinical outcomes.[15] Initial attempts are being made to create new technologies, such as gene panel tests, for choosing among panels of generic drugs (e.g. statins and antidepressants).[16] To the extent that such tests or processes may be difficult to patent or protect from rapid copying, it remains uncertain how these areas of personalized medicine will attract sufficient investment capital and research and development funds to permit their commercialization.

1. Imminent Federal Pricing of Highly Innovative Molecular Tests

Although some predict that the U.S. health care system will turn away from fee-for-service payment in the next decade, the great majority of care is still delivered in this manner. Fee-for-service payment requires the use of procedure codes, diagnosis codes, and other information submitted on claims forms to determine whether a particular service should be covered and paid.

Claims coding is highly detail-oriented. Today, the details are changing fast, but a status quo lasted from the early 1990s through the end of 2012. Early in that period, tests that involved DNA or RNA measurement were rarely performed. From that quiet corner of the laboratory, molecular tests were billed to the patient's insurance through a simple, even trivial, coding system. There was a procedure code for DNA extraction, which might typically pay the laboratory $20. Then, if five types of DNA were amplified, there was a code for DNA amplification, which was paid at a rate of five times the unit of $20, or $100. A few similar process-oriented codes and a final reporting code could bring the total to between $200 and $300. This process was called "code-stacking." "Coding and payment" for molecular tests could be summarized concisely using this term.

CODE SET CHANGES

But in 2011, the American Medical Association (AMA), which controls the standard code set for the communication of outpatient and laboratory services between providers and payers, developed more than 100 new codes for genetic tests and deleted the former "stack codes" for molecular test processes (Figure 1). Medicare normally would have begun using the new test codes in 2012, but it faced a dilemma in whether genetic tests should be classified as (a) clinical chemistry laboratory tests, such as a glucose or thyroid hormone test, or (b) physician pathology services, such as examining a microscope slide containing a potentially cancerous cell to develop a diagnosis. By the end of 2012, Medicare determined that genetic tests are frequently managed by highly trained laboratory staff who are *not* physicians, and classified genetic tests as clinical laboratory tests in that they do not *require* a physician. Medicare released the genetic test code set for use in January 2013.

Representative Molecular Payment Codes	
Before 2013	**From 2013 Forward**
"Stack" Codes for DNA/RNA Processes • **83898** DNA Amplification Step • **83904** DNA Sequencing Step • **83912** Sequence Report & Interpretation	**Gene-specific Codes ("Tier 1")** • **81200** ASPA (aspartoacylase, Canavan disease) gene analysis • **81220** CFTR (cystic fibrosis gene), common variants • **81310** NPM1 (nucleoplasmin, acute myeloid leukemia), exon 12 variants *AND* **~500 genes classified in "Tier 2, Levels 1-9"** • **81400** Level 1 Molecular Procedure, e.g. single germline variant, single nucleotide polymorphism • **82404** Level 4 Molecular Procedure, e.g. analysis of single gene exon or > 10 amplicons • *Each "level" may list up to dozens of genes to be represented on insurance claims by a single Level code, e.g. 82404*

FIGURE 1: Medicare released AMA molecular test code changes, including those shown here, for use in January of 2013.

PRICING POLICY CHANGES

The stack codes for molecular tests were many years old and had fixed prices of about $20 per molecular step (such as DNA extraction or DNA amplification). The new codes represented single genetic test services that were likely to range widely in price, from less than $100 to more than $1,000. There was no one-size-fits-all price, nor did Medicare propose using only a few tiered prices (such as prices for low-cost, middle-cost, and high-cost genetic tests).

As mentioned earlier, Medicare sets nearly all prices for medical and surgical services through purely administrative formulas that attempt to set price to cost; the exception is drugs, which are paid by law at average market prices. For example, physician services are set by elaborate and granular calculations that allocate minutes of physician time, staff time, supplies, and equipment.[17] In the case of inpatient and outpatient hospital services, nearly automated methods set Medicare prices in proportion to hospital charges, but within a fixed budget.

What happens when a code is created for a new laboratory test? For nearly 30 years, Medicare has used just two administrative methods to set a federal price for a new clinical laboratory test. In the first method, the new test is assigned a price that is a crosswalk from an older, already priced test. In fact, prior to 2013, Medicare had almost always used the crosswalking methodology. When the agency determines that no *crosswalk* is available with which to set pricing, Medicare assigns the new test a new price according to the year-long gapfill method (Figure 2).

In an unprecedented move, Medicare decided to use the *gapfill* method to price the 100-plus new genetic codes during calendar year 2013. CMS does not set the gapfill prices itself. Instead, the agency relies on ten Medicare Administrative Contractors (MACs) (Figure 3).[18]

Lab Pricing: Crosswalk Method	Lab Pricing: Gapfill Method
Based on public comment and central agency staff decisions, the new test code is assigned an existing price already used for an older, similar test code.	• Lists of codes to be priced by gapfill are forwarded to the ten MACs around January 1. • MACs assign prices by March/April, and report them to Medicare. • Public comment is accepted for 60 days. • MACs revise prices by September, and report them to Medicare.

FIGURE 2: Medicare assigns a price to a new laboratory test code according to one of the two processes described here, which are commonly referred to as the crosswalk and gapfill methods.

Just before the beginning of a calendar year, Medicare tells its MACs which lab test codes have been administratively priced (using the crosswalk method) and which codes will be initially unpriced (subject to the gapfill method). According to the gapfill process, the MACs assign prices to the specified laboratory codes early in the year, and should begin processing claims for those tests and codes submitted by laboratories within a month or less. By spring, the MACs are required to report their pricing decisions to Medicare, which posts the prices (MAC by MAC) for public comment. In September, the MACs submit final prices to CMS. CMS regulations then set the median of the ten MAC prices as the permanent, future national price for each code.

Medicare is in the process of going from more than two dozen Medicare contractors 15 years ago to ten or fewer in the near future. The final contractor map is shown below. Not all jurisdictions are in their final assignment yet. When new molecular test codes are gapfilled, each MAC jurisdiction selects and submits its proposed price to Medicare.

CONSOLIDATED A/B MAC JURISDICTIONS

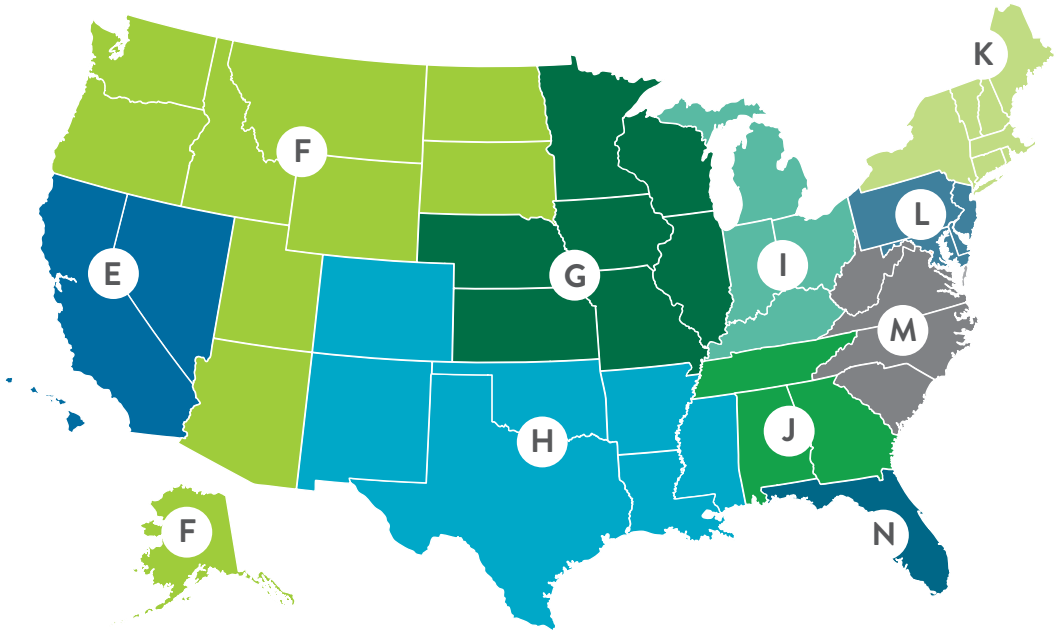


FIGURE 3: Medicare is in the process of going from more than two dozen Medicare contractors 15 years ago to ten or fewer in the near future. The final contractor map is shown above. Not all jurisdictions are in their final assignment yet.[19] When new molecular test codes are gapfilled, each MAC jurisdiction selects and submits its proposed price to Medicare.

CHANGE IN CODING AND PRICING: THE SYSTEM ALMOST CRASHES

In 2013, this process unfolded with a great deal of difficulty.[20] Many labs reported receiving no payments at all from their regional MAC in the first fiscal quarter of 2013. Then, payments began to be distributed erratically, and some tests were not paid for at all—they were rejected for payment.[21]

The unexpected tacit rejection of many tests for payment surprised the personalized medicine community because contractors normally make coverage decisions using a public process called "Local Coverage Determinations." According to this process, the MAC publishes the proposed coverage policy on its website and receives public comment. Because coverage denial decisions were made without this public comment process, organizations in the laboratory industry petitioned CMS, the Department of Health and Human Services (HHS), and Congress with complaints and demands for corrections.[22] The range of concerned stakeholders included PMC, the American Clinical Laboratory Association (ACLA), the Association for Molecular

Pathology (AMP), the College of American Pathologists (CAP), and others. Most of their concerns were that the problems endangered patient access to molecular diagnostics, many of which had become standard-of-care tests.[23] At one point, one of the nation's largest laboratories, LabCorp, stated that it had $40 million dollars in unpaid molecular tests.[24]

On behalf of its members, PMC raised its concerns to CMS. The organization stated that (1) there would be unintended consequences of the pricing process on the future of personalized medicine, (2) prices were being set with a lack of governmental transparency, and (3) inadequate payment levels could result in patient access and quality problems.[25] Final prices settled at levels that were typically somewhat lower than in the prior calendar year.[26,27]

PRICING CHALLENGES COULD GET WORSE

We believe that the pricing issues for molecular lab tests have several root causes, and that two of these causes in particular will have *greater implications in 2014 and 2015*, when multi-analyte and next-generation sequencing tests will be subjected to CMS' gapfill pricing process.

The *first root cause* is that gapfill rate setting is handed over to local physician medical directors who help guide claims processing for the Medicare program. Nationwide, only one or two of the several dozen Medicare medical directors are pathologists. Gapfill methods have been used since the 1960s to price physician services that lacked a code and a price. If a physician did a two-hour surgery in a field where a one-hour surgery was $1,000 and a three-hour surgery was $3,000, the two-hour surgery would likely be priced on a one-time basis at $2,000. This sort of pricing does not scale well for pricing genomic technologies like next-generation sequencing tests, which are based on, in some cases, tens of millions of dollars of upfront development and investment. For example, the Exact Sciences Corporation, which is bringing a molecular colon cancer screening test through FDA approval, has accrued $320 million in investments.[28] Such a test would be difficult to price based solely on the cost of the benchtop chemistry per test and the addition of a few minutes of labor. The MolDX program, specially contracted by CMS, devoted substantial resources to molecular pathology price-setting.[29] However, other Medicare contractors seemed to "copy" the MolDX prices, so the process differed from the original intentions of the gapfill regulations. Many in the community argued that if the MolDX contractor did not get pricing right, neither would any other contractor. The solution is hardly stable: If the MolDX program were discontinued, a giant "gap" in the "gapfill" process would open up.

WHAT IS THE MOLDX PROGRAM?

MolDX is a special program launched by Medicare as a subcontract to its West Coast jurisdiction (California, Hawaii, and Nevada, now called Jurisdiction E). The regional contractor, then Palmetto GBA, was able to hire and allocate staff to coverage and reimbursement issues for molecular diagnostics. The contract, now in its third year, provides coverage and reimbursement decisions for seven states (four east coast states, Jurisdiction 11, and three west coast states, Jurisdiction E). The program has a special website, www.palmettogba.com/MolDX. The project uses a panel of external expert reviewers to help in technology assessments based on publications and dossiers submitted by a test provider.

Some stakeholders have criticized the program for its lack of transparency about the evidence that was reviewed, the criteria used to assess the evidence, differences from the local coverage determination process as described in law and regulation, and lack of agreement with published studies, consensus guidelines, and expert opinion.[30]

The *second root cause* is that Medicare reportedly used an indirect cost level of 30 percent to assign final prices for genetic tests, after time and materials for a particular test were determined by something akin to "time-and-motion" accounting methods. This means that the cost of goods sold would be 70 percent of revenue. This indirect cost rate is probably low for the field of innovative molecular medicine. The largest publicly traded laboratories in the U.S. have indirect costs (defined as revenue minus costs of goods sold) that are closer to 50 percent. But even those laboratories would generate most revenue in classical mass-scale clinical chemistry tests such as hematocrits, sodium and glucose tests, and urinalysis, for example. Esoteric tests and research-based tests must therefore have a *higher than average level* of indirect costs, including intellectual property costs, investment at scientific risk (also known as development risk), and overhead for professional laboratory operations in the molecular field. Research-oriented manufacturers, as a benchmark, tend to report their cost of goods sold at around 20-25 percent of revenue. Even physician offices, according to Medicare's own data, have average *indirect* expenses of around 60-70 percent.

In 2014, 2015, and 2016, there will be national pricing of multiple medical technologies that require enormous upfront investments to establish their clinical validity, which will often happen through large-scale clinical trials. These technologies have the potential to strain the accounting approach used thus far for gapfill genomic tests.

Multi-analyte assays with algorithms. MAAAs[31] require large-scale retrospective studies for molecular validation, as well as additional prospective clinical utilization trials. CMS declined to price MAAA codes in 2013 and continued to state, until late in 2013, that the agency would not allow use of these codes in 2014, either. This seemed especially puzzling in light of the fact that a range of these MAAA tests are used nationally, are placed in clinical care guidelines, are approved (when submitted for approval) by the FDA, and are paid for by Medicare's local MACs under a nonspecific code (e.g. a code for "other molecular test, not specified"). It was surprising that the same test would become unpayable on the first of the year simply because it was given a specific laboratory test code. These tests clearly require a great deal of upfront validation, investment, and research that is not required for a typical laboratory test run on a platform based on a pre-existing and mass-produced kit. Therefore, pricing such a test purely on the "cost of goods sold"—the final marginal cost of operations—cannot sustain personalized medicine products and services. The tests should be subject to gapfill rules, which allow Medicare to consider "prices of other payers" as well as "resources to perform the test," thus accounting for the resources needed to bring the test into existence in the first place.[32] The application of the gapfill rules to federal MAAA test pricing will play out for the first time in 2014.

Next-generation sequencing tests. Although next-generation sequencing tests can be highly efficient at the correct position in a patient care pathway, they require large upfront investments in capital equipment, test development, and bioinformatics. In many cases, they also require close interpretation and the personalization of the test results to the clinical problem presented by the patient. In short, simply assessing the raw chemistry cost in order to set the price of next-generation sequencing would fall far short of the actual costs of bringing the test into existence. So far, it has been difficult for CMS to allocate the costs of bioinformatics and sophisticated computer equipment, at least for its physician fee schedule (a pricing system familiar to its regional medical directors). Other capital-intensive medical equipment, such as radiotherapy machines and PET scanners, are assigned a useful life, and their investments are amortized for the time a patient has the dedicated use of one machine.[33] This cannot be done for computers, as the "uses of the computer for a single patient" are usually difficult to calculate or might be measured in milliseconds. Faced with this challenge, CMS would assign the next-generation sequencing bioinformatics computer to general overhead, which would be the same overhead rate as a paraffin block or urinalysis laboratory. Clearly, Medicare will have to specify and adapt new and more fit-for-purpose principles in pricing next-generation sequencing tests.

2. Inconsistent Standards and Paradigms for Evaluating Diagnostic, Prognostic, and Predictive Genomic Tests

Questions regarding data in support of the coverage, payment, and use of personalized medicine products and services are important. As evidence, in the past five years, there have been innumerable conferences, often sponsored by prestigious entities such as the Institute of Medicine (IOM), on the topic of evaluating genomic tests and bringing them into medical practice.[34] The number of academic policy publications on this topic number in the dozens. In April 2013, PMC and BIO co-sponsored a national forum entitled *Evidence, Coverage, and Incentives: A PMC/BIO Solutions Summit* that attracted more than 300 attendees.[35] Most paradigms for evidence evaluation use one or more of the following four frameworks:

1. BlueCross BlueShield (BCBS) Evidence Framework. This framework is used by all BCBS Technology Evidence Centers (TECs).[36] Five evaluations are conducted sequentially: (1) Does the technology have regulatory approval? (2) Does scientific evidence allow conclusions for the technology's effect on health outcomes? (3) Does the technology improve net health outcomes? (4) Is the technology as beneficial as established alternatives? (5) Is the improvement attainable outside investigational settings?

2. Levels of Evidence Framework. Although there are many variants of these frameworks, with some adding "meta-analysis of multiple randomized controlled trials" as the highest level, the U.S. Preventive Services Task Force guidance is typical—Level I: randomized, controlled trials; Level II: well-designed, controlled trials without randomization; Level III: cohort and case-control studies; Level IV: respected authorities and expert committees.[37]

3. Systematic Review Protocols. Systematic reviews by independent bodies were introduced to counter concerns that traditional expert reviews (or the reviews of medical specialist panels) could be biased based on the experience of the individual expert or the economic concerns of the panel. There were concerns, for example, that contrary evidence might be downplayed or omitted. Systematic reviews predefine the body of evidence to be searched (e.g. Medline and other archives) over a duration of years or a set of languages; all relevant articles are pulled by abstract,

and all articles meeting a predefined standard are collected and reviewed in full against systematic rules. In the U.S., the Agency for Healthcare Research and Quality (AHRQ) publishes guidelines for systematic evidence reviews and contracts with 11 regional centers that perform the reviews on a contractual basis.[38]

4. Centers for Disease Control (CDC)/"Evaluation of Genomic Applications in Practice and Prevention (EGAPP)" Diagnostic Test Paradigm. The National Office of Public Health Genomics at the Centers for Disease Control (CDC) established the "Evaluation of Genomic Applications in Practice and Prevention (EGAPP),"[39,40] an independent, non-federal working group.[41] Diagnostic test evaluations are broken into three cardinal parts: analytical validity, clinical validity, and clinical utility.[42] Using two of these levels, the FDA generally describes its evaluations as focused on analytical validity (the test performance on a chemical basis, e.g. precision and reliability) and clinical validity (the correlation between the test's analytical results and a clinical phenomenon). Insurers describe their focus as clinical utility (the impact of the test on health care, in clinical use), which is supported by the test's analytical validity and clinical validity. This common framework is used, for example, by the Medicare program under MolDX.[43]

The four paradigms described above are not exclusive, and they are used interdependently in various ways. Although there is no question that genomic tests for personalized medicine can only have an impact on patient care through their "clinical utility," there are concerns that a surprising number of unintended consequences can occur when paradigms developed primarily for the evaluation of direct interventions (such as surgeries or drugs) are applied to diagnostic tests. Other potential approaches are possible, as discussed in a recent Cochrane Collaboration methodologic review.[44] A few potential problems are highlighted here:

Levels of Evidence. Noted authority Sir Michael Rawlins, head of the U.K.'s National Institute for Health and Clinical Excellence until 2013, has decried the crystallization of "levels of evidence" as meaningful *in and of themselves*.[45] That is, Rawlins is concerned that "levels of evidence" in isolation "attempt to replace judgment with an oversimplistic, pseudoquantitative assessment," rather than evaluating evidence as "fit for purpose…the heart of making decisions." According to Rawlins, what we have called "levels of evidence" is important but does not reflect the "amount of evidence." "Levels of evidence" are actually a *trial classification scheme* reflecting the degree to which controlled trials provide greater confidence that an effect between two groups is real and not an artifact of confounding or bias. This should not be confused with something more important and more abstract: the confidence that a scientific conclusion is correct. For example, there is no doubt that the Huntington's gene causes Huntington's disease, or that the cystic fibrosis gene causes cystic fibrosis, despite the fact that these conclusions (which might merit a Nobel prize) are based on observational data and expert opinion that the observational

data is correct. Although the conclusions about Huntington's disease or cystic fibrosis are scientific facts, the reasoning that supports these facts would be classified within the basement "levels of evidence" in the standard paradigm.

Quality of Trials. But evidence evaluation bodies categorize both the trial design (e.g. a randomized trial) and other aspects of trial quality. These other aspects include the size of the trial, whether the evaluation metric is subject to bias, and whether placebo effects have been controlled for. Although all of these are important, diagnostic test trials may lack some of the ideal features. For example, an ideal trial is "double-blinded," meaning that neither the physician nor the patient knows which arm of the trial they are in. This approach is nonsensical for diagnostic trials. Imagine a PET scan trial with two arms: one with and one without a real PET scan. Patients in the control arm might be put through a simulated PET scan experience for a few hours. It doesn't make sense, however, to double-blind the trial: The *physicians* would have to receive a report for every patient in the two arms, but half the reports would be "fake" reports (for example, randomly stating the patient in the non-PET arm does or does not have cancer on his simulated PET report). Downgrading a diagnostic test trial because it is not double-blind is therefore illogical.

Concerns about "Bias." A key role of both journal peer reviews and evidence reviewers is to evaluate trials for a lack of bias or to estimate the maximum likely scale of bias. Diagnostic trials are frequently inefficient when the two arms of the trial are pivoted on a diagnostic test, a phenomenon that is understood by sponsors and experimentalists[46] but may not be perceived by non-specialist reviewers. For example, the highest-quality trial is to be both randomized and pivoted on the diagnostic test itself. Clearly, this is not efficient. Imagine a trial of the effective cancer drug trastuzumab, where 100 women are treated with the standard of care chemotherapy and 100 are given the relevant diagnostic, the HercepTest. Twenty of the 100 women in the HercepTest arm are positive and given trastuzumab, of which ten respond. The overall trial design results in comparison arms in which 80 of the patients in both arms are treated identically. This is unlike a drug trial, in which patients in two arms (getting Drug A or getting Drug B) are never treated identically. A better trial design, in which HercepTest-positive patients are either treated with Herceptin or not, should not be rejected as "biased" and "low quality" when this is the only trial design that is efficient and makes clinical sense.

Framing of the Question at Issue. While this point may seem obvious, framing the question at issue is crucial if the subsequent systematic evidence evaluation is to be meaningful. A question of framing could be applied to AHRQ's evidence assessment for gene panel tests used to identify the tissue of origin in a patient with a cancer of unknown primary origin. In a traditional

response to this patient, the whole hospital works to identify the cancer: diagnostic tests include a thorough history and physical, body scans (CT, MRI, and/or PET scans), biopsies, colonoscopy, blood tests, and elaborate tissue diagnostics (such as flow cytometry with large panels of antibodies, immunohistochemistry, electron microscopy, and other technologies). Several validated gene panel tests can now fingerprint a tissue of origin—based on RNA expression—in 80 percent or more of these cases. The "question at issue" is fundamental: Should these gene panel tests be handled as an entirely new mode of care that is unsafe and/or unnecessary unless evaluated for long-term survival outcomes, or are they a "better-faster-cheaper" approach to doing the diagnostic effort that the physician team and hospital were doing anyway? Although there is no clear answer to this question, the authors of a recent federal technology assessment did not even see the opportunity to ask it. This resulted in sharp disagreements among experts as to the overall relevance of the technology assessment.[47]

Debates over the "Definition of Clinical Utility." These debates are frequent in the policy literature and at current conferences. Clinical utility is typically defined as "the relevance and usefulness of an intervention in patient care."[48] At this level, the definition seems clear enough, and one similarly worded definition fills in as well as another.

But when stakeholders debate "the definition of clinical utility," they are more likely debating the boundary zones of an outcome that is an appropriate measure of clinical utility. In its national coverage decisions for diagnostic tests, the Medicare agency states that convincing outcomes are generally overall survival and similar, patient-relevant outcomes, whereas biomarker outcomes (such as laboratory tests and imaging results) are less convincing. As a generalization, this is appropriate, but this should not foreclose discussion over what metrics are "fit-for-purpose" in a particular setting. Rather than seeking a simple, short definition of clinical utility, the appropriate metrics for clinical utility are context dependent.

Biomarker Outcomes	Direct Clinical Outcomes	Subjective Clinical Outcomes
• Biomarkers • Imaging	• Survival • Unnecessary surgery avoided • Faster recovery • Effective drug chosen/ ineffective drug avoided • Major activities of daily living (walking, reading)	• Diagnosis (per se) • "Value of Knowing" • Impacts on caregivers

FIGURE 4: As described in this table, there are several kinds of clinical outcomes for diagnotic tests.

For example, sometimes biomarkers are appropriate outcomes (Figure 4). HIV patients may begin to fail on an antiretroviral regimen. If a diagnostic test helps select a drug that will rapidly reverse the rising HIV viral load, as opposed to a drug that will have no effect, that diagnostic test likely has "clinical utility," despite the fact that the measured patient outcome (of the new diagnostic test) is a biomarker—the HIV viral load. On the other hand, the position that being able to read (such as after cataract surgery) is an acceptable outcome but that knowing one's diagnosis (such as when facing a terminal disease) is not good enough deserves more articulate debate among a balanced group of stakeholders. Obtaining an accurate diagnosis has been a fundamental goal of medicine since William Osler's time.[49] Accurate diagnosis is the core work of pathologists and the major pivot point for a patient's care. We can make more headway in debates over the definition of clinical utility by realizing that we are not really debating a dictionary definition. Rather, we are debating what trial designs and outcomes should be relevant, where the gray areas lie, and how these factors are judged to play out when applied in a particular medical context.

Evaluation of Clinical Utility. One recurrent problem, then, is that the two-word phrase "clinical utility" does not give the test developer, the clinician, the patient, or the evaluation body a firm grasp of what to do next. A more operational definition would break the noun phrase, "clinical utility," which is only a label, into workable parts. With regard to the concept of "clinical utility," Felix Frueh, Ph.D., has suggested using the following six questions as a framework for communicating the utility of a test:

- Who should be tested and under what circumstances?
- What does the test tell us?
- Can we act on the information provided by the test?
- Will we act on the information provided by the test?
- Does the outcome change, in a way we find value in?
- Can we afford it? (Is it a *reasonable* value?)

Although these six factors underlying clinical utility may seem obvious, these or other frameworks would help stakeholders plan trials, frame evidence and expectations, and most importantly, clarify discussion and debate.[50] For example, when framed this way, the available evidence can be parsed into areas that the reviewer finds acceptable, in contrast to areas where the reviewer finds the evidence not yet convincing.

Building a Bridge from Clinical Validity to Clinical Utility. The most important contribution of a diagnostic test is that when it is inserted into a care pathway, the clinical outcome is better in some way than it would be otherwise. There will always be a mapping between clinical validity and clinical utility: The new test will bring clinical validity that is better in some way than the

information we had without the test (gain in clinical validity). This new information improves the decision and management pathway (thus, a gain in clinical utility). This relationship can be shown schematically in Figure 5:

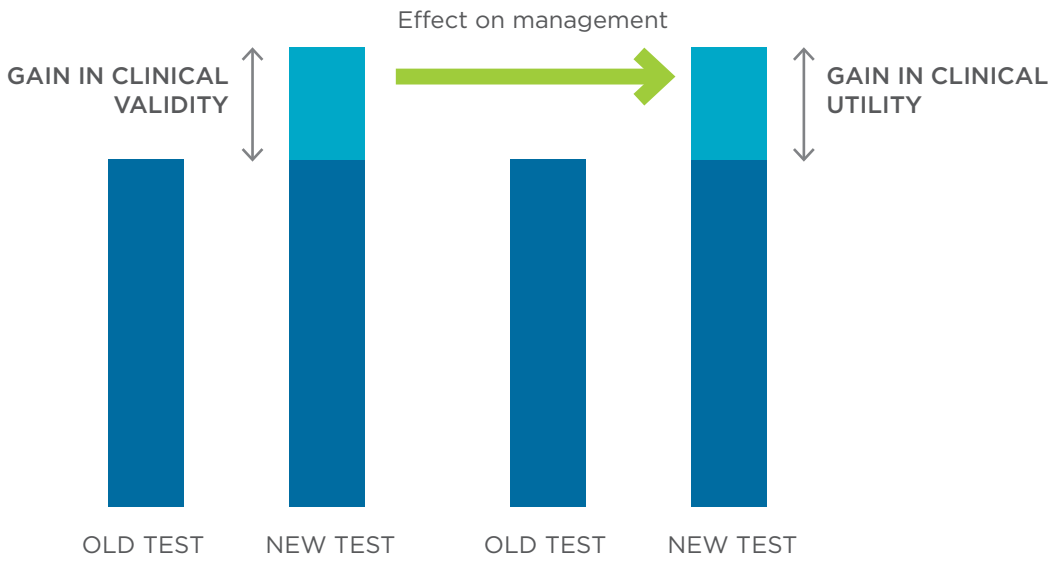


FIGURE 5: New diagnostic tests can provide information that is used to more effectively manage care.

Experience suggests that only context-specific discussion and debate can resolve the issues represented diagrammatically here. If the *clinical utility* is better, we have to be able to evaluate against what comparators, in what units, and with what uncertainty.[51]

Comparators. Drug trials are generally made *against one standard-of-care comparator drug or against placebo*, but the clinical utility outcomes for diagnostics potentially span a range of different management approaches to the patient's illness. The developer may want the most impressive and easiest comparator available, while the payer may hypothesize many realistic alternative management scenarios and prefer the one with the most extensive data (the largest populations, multiple types of populations, multiple care settings, and long durations).

Units. The units of clinical utility are numerous, as shown above (survival, surgeries avoided, faster recovery, etc.), but gray areas regarding the "allowable" outcome may exist.

Uncertainty. Finally, there are different types of uncertainty regarding the improved clinical utility that is being asserted. The first is simple statistical uncertainty: a result is 50 percent, plus or minus 5 percent. The second is pragmatic uncertainty, or external validity: Does the result reported in one trial apply to future and uncontrolled populations?[52] Third, there may be conceptual uncertainty about the result. For example, results are available from a ten-year-old patient cohort, which is the only way to have ten-year survival data, but perhaps today's drugs and treatment paradigms are different and evolving.

Gains in clinical utility are easier to have confidence in when they are clearly caused by quantified gains in clinical validity. How do we quantify gains in clinical validity? We again ask against what comparator, in what units, and with what uncertainty?

Comparator. There may be no simple gold standard. For diagnosing and staging cancer, PET scans may be better than CT scans, which may be better than X-rays. However, there is no perfect "gold standard."

Units. Tests that provide better prognostic validity are very important. The Onco*type* DX breast cancer prognostic test is one example. It provides a more accurate ten-year prognosis than prior measures, such as tumor size. However, it can be surprisingly hard to quantify the units for this gain in prognostic accuracy. Is the p value lower (p=.001 rather than p =.03)? What does this mean clinically? Is the correlation value—such as r2—lower? What does that mean clinically? Reclassification indices seem more concrete, but the finer points of their implementation are still debated.[53]

Uncertainty. In addition to statistical uncertainty, there may be pragmatic uncertainty. Does the cohort studied ten years ago match the cohort we'll see tomorrow? Do all reviewers agree that, conceptually, the statistical analysis shows that a test is predictive, prognostic, or both?

As indicated in the graphic above, through its effect on management, a gain in "clinical validity" precedes and is causal to a gain in clinical utility. When the magnitude of the gain in "clinical validity" is difficult to express, it lowers confidence in the overall story.

There is nothing wrong with debates about evidence. Such debates have been the soul of the scientific endeavor for hundreds of years. In 1650, near the end of his life, the eminent English scientist William Harvey, who discovered the circulation of blood, described the responses of his fellows.

"Scarce a day, scarce an hour, has passed since the birthday of the circulation of the blood that I have not heard something for good or evil said of this discovery. *Some* abuse it as a feeble infant, and yet unworthy to have seen the light; *others* again think the bantling deserves to be cherished and cared for; these oppose it with much ado, those patronize it with abundant commendation; *one party* holds I have completely demonstrated the circulation of the blood by experiment, observation, and ocular inspection, against all force and array of argument; *another* thinks it scarcely yet sufficiently illustrated—not yet cleared of all objections."[54] *(Harvey also observed that "no man more than 40 years old was found to adopt the doctrine of the circulation of the blood," just as Max Planck wrote, in 1948, that scientific truths do not convince their opponents, but their opponents eventually die.)*[55]

Rawlins asks us to undertake a rich and medically informed debate about whether studies are fit for purpose and whether an overall story for improved health care is convincing. Leading international regulators recognize and articulate that there must be a balance between evidence, benefit, and risk, and that excessive conservatism actually reduces public health.[56] The same may be true for coverage decisions. In the near term, better frameworks and better clarity of argument are probably the means to advance. Ultimately, clear use of these frameworks and evidence-based reasoning should guide the "convincingness" of the overall body of evidence. Rather than saying a test and its proposed application "lack clinical utility," we should say that we are not yet convinced of the clinical utility and explain why and what type of evidence would be convincing to us.

3. Lack of Incentives for Genomic Medicine

In two recent books, Bartfai and Lees discuss how impactful commercial incentives are on what pharmaceutical products are brought to market.[57] It is important for governments to help redress these imbalances and shift translational research to meet the most acute health care needs. Recently, academic authors have proposed a range of policy incentives and solutions to these issues,[58] some of which are reflected in recently proposed legislation.[59]

The landscape for intellectual property in diagnostic tests has been shaped by two Supreme Court decisions. The first case, Prometheus v. Mayo, decided March 30, 2012, invalidated a patent that related measured drug levels to an algorithm that could guide dosing of the relevant drug. The court ruled that this patent failed to exceed a standard of obviousness. The second case, Association for Molecular Pathology v. Myriad Genetics, decided June 13, 2013, ruled that patents on a gene per se were not valid because the gene was a product of nature, which was not, of itself, patentable. Both cases hinged on particular issues that would not underlie investments in a majority of new genomic tests. However, these cases do provide a springboard for discussion on the incentives that exist for investing in and producing new innovations in molecular testing.

While Bartfai and Lees focused on incentives and incentive gaps with regard to drug development, these topics may also be relevant to genomic tests. We are starting to see areas where genomic medicine could have a significant impact on public health, but where traditional funding or pricing and reimbursement fail to provide enough incentives. These areas include funding the education of physicians and patients in personalized medicine, providing funding for allied professionals such as genetic counselors, and offering incentives to develop new tools that could revolutionize some therapeutic areas. For example, it seems clear that because of the complexity of human drug metabolism, there are cases in which the contribution from an individual gene for a single drug has only limited impact on clinical outcomes.[60] As an alternative to the one-gene-one-drug approach in pharmacokinetics, the first attempts are being made to create new technologies, such as gene panel tests for choosing generic drugs (e.g. statins and antidepressants).[61]

It remains uncertain how investment capital and research and development funds can be channeled into these potentially important and cost-effective areas of personalized medicine. Since one of the most important mechanisms that exists in free markets is price, and list prices are meaningless unless they are accepted by payers, payers play a major role. Strong evidence for clinical validity and clinical utility should therefore be accompanied by a system that ensures that resulting tests can be brought through translational research and into the delivery of health care.

NEXT STEPS

We close this report by contrasting the public positions of the two most important federal agencies for the market entry of personalized medicine into the health care system: the FDA and CMS. In October 2013, the FDA released a cross-agency, 60-page document entitled, "Paving the Way for Personalized Medicine: FDA's Role in a New Era of Medical Product Development."[63] The agency discusses its initiatives to advance personalized medicine in other forums as well, such as the *New England Journal of Medicine*.[64] The FDA is closely involved with efforts costing $250 million to facilitate rapid advancements in genomic medicine, which are being conducted by the NIH and multiple industry and academic stakeholders.[65]

In contrast, CMS' main public documents in 2013 were the range of cost-cutting proposals discussed in this white paper, as well as a strikingly negative technology assessment on genomic testing last spring.[66] Why do the two agency's public positions seem to differ so greatly?

We sympathize with the position of the Medicare agency, which must create public policies protecting the fiscal integrity of the Medicare program in an era of great financial stress,[67] whereas the FDA has the focused mandate to review products one at a time against a standard of risk versus benefit. But the differences in communication are striking. For example, when the FDA convenes a public advisory panel on a device or drug, the FDA staff members are highly engaged in presenting data, viewpoints, and positions on the topic. In contrast, when the Medicare agency convenes a coverage advisory board, the agency adopts an almost invisible profile during the meeting.

This white paper noted that one of the fundamental principles of Medicare's payment system is to reimburse the final provider of a service with an amount close to its cost to deliver the service, and no more. For example, an oncology drug may cost $50,000 to purchase (more than its cost of production), but Medicare would pay only a few percent above that purchase cost to the oncologist or clinic that provides the drug and bills Medicare for it. Similarly, a pacemaker may have a purchase price of $10,000 or more above its cost of production, but Medicare would pay the hospital only a bit more or less than the cost of providing the pacemaker to a patient.

This bottom-up costing method can work well enough if there is a drive for innovation and if the original innovators can eventually recoup their risk and investment during the reimbursement process. But if the agency attempts to set payment to marginal costs as soon as something is invented and introduced, there is no opportunity to recoup R&D costs. And if real costs are systematically underestimated—through the application of routine laboratory "indirect costs" to next-generation sequencing or by omitting the costs of bioinformatics investments and software—then a real barrier to innovation is created.

Medicare policies are powerful. In 2014, Medicare is rolling out a range of new policies with an impact on personalized medicine that remains to be seen. One of these is that the cost of clinical laboratory tests will be bundled into the base payment for office visits and procedures undertaken in the hospital outpatient setting.[68] This is concerning, for example, when an HIV-positive patient has an outpatient office visit—for which the hospital is paid about $92—but requires several hundred dollars of molecular tests. The agency remains uncertain about the valuation of bioinformatics, whether this service takes place in the context of a MAAA or in the context of complex genomic tests, such as next-generation gene sequencing panels. Some of this uncertainty will be resolved by 2017, with the April 2014 passage of the "Protecting Access to Medicare Act." Under this new law, routine clinical chemistry tests and genomic tests will eventually be paid by Medicare at the average of commercial insurance market prices. The law also requires laboratory coverage decisions to be made by a public process for draft and final local coverage decisions.

Although this white paper has focused on federal policies, such as those of the Medicare agency, the payment process in the U.S. for commercial payers is also complex and cumbersome.[69,70] The previous head of the Medicare agency articulated the need to try to foresee unintended consequences in technology.[71] We need to take this advice while trying to foresee unintended consequences of our policy decisions.

"Innovation is happening broadly across the country. The promise of personalized medicine and innovation is amazing, and we're already seeing dividends."

Patrick Conway, MD

Deputy Administrator for Innovation and Quality & Chief Medical Officer, Centers for Medicare & Medicaid Services at the Tenth Annual State of Personalized Medicine Luncheon hosted by PMC March 15, 2014

In contrast to options for retail goods or business-to-business purchases, consumers do not have much say in their choices about health care and treatment options in the future. Clearly, someone who will get leukemia five years in the future would want his or her physician to have access to a diagnostic test to pick the right, as opposed to the wrong, chemotherapy. It is clear that if such a test and the corresponding drug are brought to the FDA, they will get regulatory approval. Payers and policymakers should also want such a diagnostic test to be invented and commercialized. Medicare can lower prices when it perceives a sector's margins are too high, and it can test whether its beneficiaries have adequate access to services by measuring whether enough physicians are taking new Medicare patients. In contrast, Medicare administrators can't do patient surveys to measure whether cost-saving and effective diagnostics are not brought to market because investors were too scared of the agency's pricing power.[72]

It may be difficult to have a successful payment system for laboratory tests without considering value. The marginal cost of the benchtop chemistry is, from the builder's perspective, only one small part of bringing the test into being, which includes much broader efforts in research, publications, and the creation of a delivery channel. Similarly, from the public and payer perspective, the marginal cost of the benchtop chemistry is only a small part of the test's value to the health care system, which is measured in years of life extended, quality of life, and the avoidance of drugs that are not helpful. The payer evaluation system has generally not been focused on addressing these critical questions: what should be developed, what should be brought to the health care system, and how can clear signals be used to incentivize this? Because personalized medicine is so innovative and evidence-based, reimbursement needs to go beyond the marginal cost of the benchtop chemistry and consider the other factors described in this paper. The development and laboratory economics of bringing about cost-savings via molecular diagnostics can be substantial for the originating provider, even when the health care system as a whole benefits.[73] Laboratory medicine gets its value in real life—in the context of ongoing health care, diagnoses, prognoses, and patient care choices.

We need a reimbursement system that encourages the improvement of patient care through medical innovation. Patients expect no less.

REFERENCES

1 http://www.whitehouse.gov/files/documents/ostp/PCAST/pcast_report_v2.pdf

2 http://transcripts.cnn.com/TRANSCRIPTS/0006/26/bn.01.html

3 Smith RJH et al. (2014) Deafness and Hereditary Hearing Loss Overview. *GeneReviews*. At: http://www.ncbi.nlm.nih.gov/books/NBK1434/

4 Bianchi DW et al. (2014) DNA sequencing versus standard prenatal aneuploidy screening. *New Engl J Med* 370:799-808.

5 Two examples include the Vermillion Ova-1 test for ovarian cancer risk and the Nanostring ProSigna test for breast cancer prognosis.

6 Frampton GM et al. (2013) Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing. *Nature Biotechnol* 31:1023-31.

7 Raguz S & Yague E (2008) Resistance to chemotherapy: new treatments and novel insights into an old problem. *Brit J Cancer* 99:387-91.

8 Werner HMJ et al. (2014) Cancer systems biology: a peek into the future of patient care. *Nat Rev Clin Oncol* 11:167-176

9 Bettegowda C et al. (2014) Detection of circulating DNA in early- and late-stage human malignancies. *Science Trans*. Med. 6:224ra24.

10 Flores M et al. (2013) P4 medicine: how systems medicine will transform the health care sector and society. *Personalized Medicine* 10:656-576.

11 www.genomeweb.com, data for CY2013.

12 Baumol WJ et al. (2013) The Cost Disease: Why Computers Get Cheaper and Health Care Doesn't. Yale Univ. Press. Cf. also http://en.wikipedia.org/wiki/Baumol's_cost_disease.

13 http://www.aoa.gov/AoARoot/Aging_Statistics/future_growth/future_growth.aspx

14 We may need more of this perspective in health care; even though the phrase "foreseeable problems" has 40,000 Google hits, it has only nine Pubmed hits. Pubmed contains over 23 million records (suggesting that only 0.00004% deal with foreseeable problems). The related concept, "unintended consequences," usually characterized after the problem has appeared, has 1.5 million Google hits but still only 1,100 Pubmed hits (0.004% of the 23 million Pubmed articles).

15 Limdi NA & Veenstra DL (2010) Expectations, validity, and reality in pharmacogenetics. *J Clin Epidemio*l 63:960-9.

16 Hall-Flavin DK et al. (2013) Utility of integrated pharmacogenomic testing to support the treatment of major depressive disorder in a psychiatric outpatient setting. *Pharmacogenetics Genomics* 23:535-48.

17 There are federal tables that list granular supply and time inputs for thousands of physician services, often with 5-20 inputs per service priced.

18 https://www.cms.gov/Medicare/Medicare-Contracting/Medicare-Administrative-Contractors/MedicareAdministrativeContractors.html

19 https://www.cms.gov/Medicare/Medicare-Contracting/Medicare-Administrative-Contractors/Downloads/Consolidated-AB-Map-Vision.pdf

20 In perspective, it was not as much difficulty as the notorious failed launch of Federal Health Exchanges in October 2013, but it seemed like a lot of difficulty at the time.

21 Gottlieb S (2013) Medicare Has Stopped Paying Bills For Medical Diagnostic Tests. *Forbes* (3/27/2013), at http://www.forbes.com/sites/scottgottlieb/2013/03/27/medicare-has-stopped-paying-bills-for-medical-diagnostic-tests-patients-will-feel-the-effects/

22 See e.g., "Low 2013 molecular rates may bankrupt some labs," free online at: http://www.darkreport.com/PDF/Low-2013-Molecular-Test-Rates-May-Bankrupt-Some-Pathology-Labs.pdf

23 For a nine-association consensus statement, May 2013, see: http://www.amp.org/about/Press_Releases/ConsensusStatement_ProfSoc(5102013).pdf

24 *Laboratory Economics*, February 2014, p. 7.

25 http://www.personalizedmedicinecoalition.org/sites/default/files/files/PMC_Final_CMS_MolDX_Gapfill_LTR_070313.pdf

26 The American Clinical Laboratory Association complained to CMS before the gapfill year started (11/30/2012), early in the gapfill year (3/27/2013), and after final prices were posted (10/30/2013). See: http://www.acla.com/sites/default/files/Letter%20from%20A%20%20Mertz%20to%20M%20%20Tavenner%20re%20Gapfilling%20-%20Final.pdf http://www.acla.com/sites/default/files/Marc%20Hartstein%20GapFill%20Ltr%203-27-13.pdf http://www.acla.com/sites/default/files/ACLA%20Comments%20on%202013%20Final%20Gapfill%20Prices%20(2).pdf

27 Medicare's 2014 prices for all laboratory tests, including molecular tests with CPT codes, is available at: http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/ClinicalLabFeeSched/clinlab.html

28 Exact Sciences, 2013 Annual Report, February 2014, http://investor.exactsciences.com/secfiling.cfm?filingID=1047469-14-1578&CIK=1124140

29 www.palmettogba.com/MolDX This program represents a recently created special subcontract for molecular test coverage and pricing rules, managed by the Medicare contractor Palmetto GBA.

30 See, for example, letter from AMP, ACMG, ASCP, ASHI, CAP dated October 30, 2013; available at: http://www.amp.org/publications_resources/position_statements_letters/documents/2013/MolDX%20Coverage%20Letter%20and%20Attachments%2010302013%20FINAL.pdf

31 For additional administrative background, see http://www.ama-assn.org//ama/pub/physician-resources/solutions-managing-your-practice/coding-billing-insurance/cpt/about-cpt/maaa-codes.page

32 Hornberger J (2013) Assigning value to medical algorithms: implications for personalized medicine. *Personalized Medicine* 10:577-588.

33 To give an oversimplified example, assume a machine has a $100,000 cost and a five-year life. CMS amortizes the cost at about $20,000 per year and assumes the machine is used about 2,000 hours per year, so a cost of $10 is allocated for each hour of its use. CMS uses complex rules with additional adjustments, but this simplified calculation captures the general principle. This approach breaks down for computers that are used for milliseconds or seconds.

34 For a comprehensive review, see: Parkinson D et al. (2014) Evidence of clinical utility: an unmet need in patients with cancer. *Clin Cancer Res* 20:1428-44.

35 Evidence, Coverage, and Incentives: A PMC/BIO Solutions Summit. Summary of Proceedings, April 17, 2013. Freely accessible here: http://www.personalizedmedicinecoalition.org/events/PMC-BIO_Evidence_Coverage_and_Incentives.

36 http://www.bcbs.com/blueresources/tec/

37 http://en.wikipedia.org/wiki/Evidence-based_medicine

38 http://www.ahrq.gov/research/findings/evidence-based-reports/overview/index.html

39 http://www.cdc.gov/genomics/gtesting/EGAPP/

40 Teutsch SM et al. (2009) Evaluation of Genomic Applications in Practice and Prevention initiative: methods of the EGAPP working group. *Genet Med* 11:3-14.

41 Ibid.

42 Somewhat different nomenclature has been proposed by EGAPP itself (Teutsch, 2009) and used by Institute of Medicine (Michael CM & Ball JR, eds., Evaluation of Biomarkers and Surrogate Endpoints, Institute of Medicine, 2010.) However, by 2014, the terms "analytical validity, clinical validity, and clinical utility" have remained dominant in most discussions in the U.S. among both stakeholders and policymakers.

43 www.palmettogba.com/MolDX; see website sections on test evaluations.

44 Cochrane Library (2011). Cochrane Methods, *Cochrane DB Syst Rev Suppl* 1:1-40. At: http://www.thecochranelibrary.com/SpringboardWebApp/userfiles/ccoch/file/Cochrane%20Methods%20September%202011.pdf

45 Rawlins M (2011) Therapeutics, Evidence, and Decision-making (pp. 4-6). CRC Press. See also: Rawlins M (2008) De testimonio: on the evidence for decisions about the use of therapeutic interventions. *Lancet* 372:2152-61. Dr. Rawlins is currently President of the Royal Society of Medicine.

46 Tajik P et al. (2013) Trial designs for personalizing cancer care: A systematic review and classification. *Clin Cancer Res* 19:4578-88.

47 http://www.cms.gov/Medicare/Coverage/DeterminationProcess/downloads/id90TA.pdf

48 FDA, 2013, "Paving the Way for Personalized Medicine" (62pp.)

49 Osler was physician-in-chief of Johns Hopkins Hospital from 1889-1905.

50 In Frueh's example, the three part summary statement that diagnostic tests have analytical validity, clinical validity, and clinical utility is "unpacked" into six questions. For comparison, the three part summary statement that "movies have a beginning, middle, and end" can be unpacked into questions such as, "Who is the protagonist? Is he sympathetic or unsympathetic? What main challenge does the protagonist face? What character development does he undergo?" and so forth.

51 These three questions - against what comparator(s)?; in what units?; with what uncertainty? are fundamental to pharmacoeconomics evaluations, but apply to defining the gain in clinical validity and the gain in clinical utility as well.

52 This is Question 5 of the BCBS TEC questions shown earlier, and always requires subjective judgment, as patients studied outside of a trial are just that - uncontrolled.

53 See e.g. Pepe MS (2011) Problems with risk reclassification methods for evaluating predictive tests. *Am J Epidemiol* 173:1327-35 and references therein.

54 Stern BF (1941) Society and Medical Progress. Princeton.

55 Planck M (1948) Wissenschaftliche Selbstbiographie, Barth Verlag, Leipzig.

56 Eichler HG et al (2013) The risks of risk aversion in drug regulation. *Nature Rev Drug Discov* 12:907-16.

57 Bartfai T & Lees GV (2005) Drug Discovery: From Bedside to Wall Street, Academic Press; and Bartfai T & Lees GV (2013) The Future of Drug Discovery: Who Decides What Diseases to Treat? Academic Press.

58 Roin B (2009) Unpatentable drugs and the standards of patentability. Tex L Rev 87:503. Outterson K (2010) *Tex L Rev* 88:45-55.

59 E.g. MODDERN CURES Act of 2011. http://en.wikipedia.org/wiki/MODDERN_Cures_Act_of_2011

60 Limdi NA & Veenstra DL (2010) Expectations, validity, and reality in pharmacogenetics. *J Clin Epidemiol* 63:960-9.

61 Hall-Flavin DK et al. (2013) Utility of integrated pharmacogenomic testing to support the treatment of major depressive disorder in a psychiatric outpatient setting. *Pharmacogenetics Genomics* 23:535-48.

62 Hamburg MA & Collins FS (2010) The Path to Personalized Medicine. *New Engl J Med* 363:301-4.

63 http://www.fda.gov/downloads/scienceresearch/specialtopics/personalizedmedicine/ucm372421.pdf

64 See, e.g., Hamburg MA & Collins FS (2010) Perspective: The path to personalized medicine. *NEJM* 363:301-304. See also: Collins FS & Hamburg MA (2013) Perspective: First FDA authorization for next-generation sequencer. *NEJM* 369:2369-2371.

65 *Darky Daily* (Trade Journal, 3/5/2014) National Institutes of Health Creates Partnership with Big Pharma to Improve Development Success of New Drugs and Diagnostics and Speed FDA Clearance http://www.darkdaily.com/national-institutes-of-health-creates-partnership-with-big-pharma-to-improve-development-success-of-new-drugs-and-diagnostics-and-speed-fda-clearance-305#axzz2v6alpn3Z See also: *Cancer Letter* (2/21/2014): NCI Launches M-PACT Next-Generation Trial As Group System Nears March 1 Transition.

66 Genetic Tests for Cancer Diagnosis. May 1, 2013; see http://www.cms.gov/medicare-coverage-database/details/medcac-meeting-details.aspx?MEDCACId=67

67 Reducing Waste in Health care. Health Policy Brief / Health Affairs (2012). https://www.healthaffairs.org/health-policybriefs/brief.php?brief_id=82

68 White paper, Quinn B (2013). Evolving challenges for value-priced LDTs. At: http://www.foleyhoag.com/publications/ebooks-and-white-papers/2013/september/evolving-challenges-for-value-priced-ldts

69 On the costs of health administration, see Kahn JG et al. (2005) The Cost Of Health Insurance Administration In California: Estimates For Insurers, Physicians, And Hospitals. *Health Affairs* 24:1629-39. Morra D et al. (2011) US Physician Practices Versus Canadians: Spending Nearly Four Times As Much Money Interacting With Payers. *Health Affairs*, http://content.healthaffairs.org/content/early/2011/08/03/hlthaff.2010.0893.abstract

70 *Dark Report* (Trade journal). New Blue Card policies cause labs to go unpaid. *Dark Report* (2/3/2014).

71 Berwick DM (2008) Taming the technology beast. *JAMA* 299:2898-2899; esp. refs 5-9 therein.

72 Enthoven A & Fong K (2006) Medicare: Negotiated drug prices may not lower costs. http://www.ncpa.org/pub/ba575 (Similarly, *Wall Street Journal*, 11/13/2006, as "Pelosi on Drugs.") See also: Blair RD & Harrison JL (2010) Monopsony in Law and Economics. Cambridge University Press, 2nd edition.

73 Davis JC et al. (2009) The microeconomics of personalized medicine: today's challenge and tomorrow's promise. *Nature Rev Drug Discovery* 8:279-86. (See also: http://www.mckinsey.com/insights/health_systems_and_services/the_microeconomics_of_personalized_medicine)

MISSION

The Personalized Medicine Coalition, representing innovators, scientists, patients, providers and payers, promotes the understanding and adoption of personalized medicine concepts, services and products to benefit patients and the health system. The Coalition's mission is to educate policymakers and the public about the power and potential of individualized health care and raise the profile of personalized medicine so that both patients and the health system will benefit from improved clinical care and increased overall value.

PMC
PERSONALIZED
MEDICINE COALITION

1710 Rhode Island Ave., NW
Suite 700
Washington, DC 20036

P: 202.589.1770
pmc@personalizedmedicinecoalition.org