

For reprint orders, please contact: reprints@futuremedicine.com



Personalized diagnostics: the struggle for position

For several major cancers, drug selection already pivots on biomarker results (e.g., trastuzumab for breast cancer, and gefitinib and erlotinib for lung cancer). Fast-paced advances in genomic and proteomic laboratory technologies could enable the widespread use of molecular testing before therapy selection in any field of medicine. This article describes two potentially large obstructions to such innovation. First, laboratory tests have traditionally been commodities with low prices, prices that matched the resources required to operate the laboratory technology itself. Assuming that the marginal costs of molecular laboratory technology will fall, there will be a widening chasm between estimated test revenue and the costs of innovative and definitive clinical trials, and regulatory approval for new tests. Without corrective action, even cost-saving laboratory tests could be in shortfall, because they will not be created through upfront investment. Second, it is argued that while diagnostic tests, drugs and surgical procedures should meet a fundamental standard for payer coverage ('will health outcomes be improved?'), molecular diagnostics could require different analysis pathways than those that are used to evaluate interventions.

KEYWORDS: biomarkers ■ business model ■ drug development
■ drug-test codevelopment ■ payer

Bruce Quinn

Foley Hoag LLP, 155 Seaport Boulevard,
Boston, MA 02210, USA
Tel.: +1 617 832 1291
Fax: +1 617 832 7000
bquinn@foleyhoag.com

The technologies that support innovation in healthcare delivery have evolved rapidly over the past 20 years. Advances in electronics and software have allowed for the development of highly accurate cardiac pacemakers, implantable cardiac defibrillators and imaging devices such as PET scanners. Advances in biomaterials have led to drug-eluting stents and innovative drug delivery systems. Recombinant DNA technologies have produced cloned proteins for valuable therapeutics, including human-sequence insulin, growth hormone, erythropoietin, glucocerebrosidase and others.

Among dozens of technological medical innovations, molecular diagnostics is the 'medical miracle' that arrives in the clinic most closely tied to its roots in a research laboratory. Basic steps, such as DNA or RNA extraction, purification, amplification and quantification, are fundamentally similar whether they are performed in a research laboratory as part of a doctoral thesis or as a widely used clinical genetic test. Admittedly, for clinical use, the molecular test may be industrialized – even converted to be run on a sophisticated cartridge-based laboratory platform – and clinical quality controls will be added, but the underlying chemistry that creates the clinical test is still readily recognized. From this viewpoint, it is hard to imagine a technology more easily transferred from the laboratory bench to the clinic than genomic testing.

However, the conclusion that the transfer of molecular diagnostics from bench to clinic is easy would be completely wrong. As other authors have noted, the safe, clinically effective and prudent adoption of molecular diagnostics for drug selection in clinical care will lead to fundamental shifts in drug-development paradigms, in regulatory approval processes, in business models and in clinical thinking [1,2]. To the extent that molecular diagnostics are most visibly paired with drug selection, and concerns that the international pharmaceutical industry has entered difficult times are rampant [3], much of the policy discussion has focused on a familiar set of applications where biomarkers are used during drug trials to select patient populations, to protect patients who are genetically vulnerable to adverse events, to tailor dosing for drugs with a narrow therapeutic index or to monitor therapeutic response [4].

As much as the introduction of biomarkers for patient selection alters Phase III trials, I argue that even greater challenges arise when we push personalized medicine models outside the pathway of a Phase III drug trial. For example, within a Phase III trial, the price of the biomarker test and the cost of its development, although significant, are dwarfed by the overall investment in bringing the index drug to market. To date, there has been relatively less discussion devoted to how new-generation

future
medicine part of fsg

biomarkers may be developed, commercialized and integrated into clinical care when they are created entirely outside of the familiar pipeline of pharmaceutical research and development. In order for such molecular diagnostics to meet both regulatory approval standards and the payer's standards of evidence, there will have to be substantial and rising levels of investment to produce clinical evidence. Such trials are difficult to fund without some degree of intellectual property protection and access to adequate reimbursement in the healthcare system. Neither is easily assured. The response to intellectual property issues for molecular testing has varied considerably from one country to another for molecular diagnostics [5]. Regarding reimbursement, in many countries, the laboratory test payment systems have, so far, treated diagnostics as commodity items based on an immunoassay or a PCR reaction for which payment is crosswalked to existing generic price categories. Although this approach has administrative simplicity, it can create technology voids due to perverse microeconomic incentives. These will occur if developers fail to invest at all in creating new diagnostics that could be predicted to have very high utility in healthcare delivery, but which would be automatically reimbursed at low fixed amounts that make it impossible to undertake the upfront investments in product development.

This article will focus on two underappreciated problems that could substantially impact on the commercialization and clinical integration of personalized medicine via molecular diagnostics in the next decade. First, the economic incentives and disincentives that occur when innovative diagnostic tests are priced by crosswalks to legacy generic prices are far from ideal. Granted, it can be easily foreseen that there could be scattered short-term benefits of low, fixed-fee pricing. However, I argue that the enforcement of very low market prices for innovative genomic tests essentially 'nationalizes' the benefit of a valuable healthcare product, and this takes place at the cost of *a priori*, large and perverse disincentives against the wider private development of useful and cost-saving tests for the healthcare system as a whole. I will discuss whether an alternative system where public investment in the research and development of the tests, balanced by commodity pricing of the resulting tests, is likely to be a good compensatory mechanism. Second, with a few exceptions, public and private payers making routine payment coverage decisions do not yet

have well-developed tools to assess the impact of novel diagnostics on the clinical care pathway [1,6]. Without such an assessment, it is impossible to make sound and insightful coverage decisions for novel diagnostic tests. I argue that there are several reasons for 'genetic (or genomic) exceptionalism' – that there are now predictable scenarios where current conventions for evidence-based medicine and health technology assessments may backfire and prove dysfunctional for genomic tests. Data for new surgical interventions or new drugs can only be found in prospective clinical trials, which have a favored position in levels-of-evidence hierarchies. By contrast, data for molecular diagnostic tests can be derived from large archives of specimens banked during clinical trials. Data from these specimens may identify a marker that strongly selects patients who respond, or entirely fail to respond, to a drug. These retrospective data make it highly probable that a test will be useful. Since this information destroys clinical equipoise, a prospective randomized trial (where marker-positive patients would be knowingly and prospectively randomized to a drug highly unlikely to work for them) may be unethical [7]. The result is a statistical 'valley of death' where there are 'too much' data for a randomized trial under the principle of equipoise, but 'not enough' data (i.e., enough prospectively randomized data) to satisfy the predetermined decision models of payers in their well-intentioned implementation of evidence-based medicine.

In short, we know that whenever investments are being considered for new medical devices and biotechnology products, including molecular tests, investors will weigh development risks that are purely scientific and technical, as well as risks posed by existing or future market entrants. My thesis is that molecular diagnostics are in a substantially worse predicament created by policy conventions that did not foresee the rich current opportunities for the creation of innovative molecular tests to reduce net healthcare costs. Innovators in clinical molecular diagnostics face additional uncertainty that has been brought about by delayed and unfavorable payer decisions in the face of retrospective data, and this uncertainty is amplified by the risk that the eventual market price of the test may be a low, fixed, legacy price that is administratively imposed rather than a rational value based on health system savings, quality-adjusted life-years (QALYs) or other socially prudent economic factors.

I make no claim that the concerns raised in this article are new. They follow from a collision of basic principles in microeconomics, evidence-based medicine and the financing of research and development. My own occupational background over the past 20 years has included work as a research-oriented medical school professor, as a physician executive in a global life sciences consultancy and as a senior medical director within the federal Medicare payment system in the USA. As a result, I believe the policy concerns raised in this article are substantial and that it will be worth the effort to develop forward-looking policy solutions.

Legacy payment systems are counterproductive for innovation

In the many national healthcare systems where some form of fee-for-service payment is used, healthcare insurance transaction systems became increasingly codified in the 1970s and 1980s. Examples of the necessary nomenclature include the Current Procedural Terminology vocabulary in the USA, which is managed through a special division of the American Medical Association (IL, USA), and the International Classification of Diseases system for clinical procedures. When these nomenclature and coding systems were being set up, laboratory tests were surely one of the least interesting aspects of the medical services that had to be classified. Laboratory tests fell naturally into discrete entities such as the 'glucose test' and the 'serum cortisol test'. These could easily be named, numbered and assigned discrete procedure codes.

Although payers and policy-makers have long had concerns about the overuse of diagnostic laboratory tests, the payments per test were quite modest. Regulatory approval of clinical chemistry tests was based mostly on their basic accuracy in measuring the analyte of interest. For screening tests, such as a hepatitis screening test used in blood banking, regulators required more voluminous data.

The industry structure is characterized by manufacturers of test kits and clinical laboratories that buy the kits, perform the service and are paid for the service. Manufacturers of test kits competed primarily on the price of the kit. Manufacturers were also encouraged to innovate in ways that would save money for the service laboratory, such as using more reliable test platform equipment with fewer breakdowns, or kits with faster cycle times or reduced sample preparatory work.

Although recent and comprehensive review articles are difficult to find, my experience at international laboratory industry conferences and in discussions with international laboratory vendors suggests that many countries have set low and relatively fixed prices for standard laboratory tests. This would certainly be expected in a commodity marketplace (e.g., serum glucose tests or serum cortisol tests).

The US laboratory payment system is worth reviewing briefly because it is the largest single healthcare market in the world and its payment rules or incentives have a corresponding influence on new medical product development. In the USA, the federal Medicare system set fixed prices for codified, familiar laboratory tests based on prevailing charges for the tests in an index year, 1984 (there have been small, intermittent revisions for inflation). The resulting US schedule of test prices is called the Clinical Laboratory Fee Schedule and it defines the test fees paid by Medicare, the US public health insurance program for citizens over the age of 65 years [101]. Most private insurance companies in the USA pay a price which is indexed to the Medicare fee schedule (e.g., 90 or 110%). In some health systems, routine laboratory tests are part of the capitated payment to primary care physicians for providing routine patient care. In other systems, such as the National Health Service in the UK and the US Department of Veterans Affairs (DC, USA) hospital network in the USA, laboratory test providers, such as hospitals, receive annual budgets, part of which are allocated to laboratory staff and equipment, so that individual clinical tests do not have a market price, and such systems will work to obtain the best-price test kits from vendors in the market.

As occurs with other innovative technologies, the newest generation of molecular diagnostics, such as the *Oncotype DX*[®] test developed by Genomic Health (CA, USA) and the *MammaPrint*[®] test developed by Agendia[™] (Amsterdam, The Netherlands), do not fit well into the legacy approach to product coding and payment. Nor is their development linked to a much larger investment in developing a drug candidate. Now, to those medical scientists focused purely on the creative intellectual and technologic challenge of pushing the boundaries of genomic technology forward, the coding and payment conventions we mentioned may seem too trivial (or dull) to discuss. However, small gaps can disable a system, such as an elevated freeway built with no on-ramps, or

a computer that has no connection between its hard drive and its motherboard. Similarly, the dusty payment conventions that begin in mundane coding systems can present substantial problems for investors, whether the investment decisions channel the internal investments of large firms or whether the decisions are made by venture capital investors looking at small companies.

There are risks lurking behind the fixed laboratory test prices, because prices have many functions. Prices may reflect the marginal costs of delivering commodity products efficiently. Prices may signal a ‘tug-of-war’ on the value created by the product and how it will be split by the vendor (provider surplus) and the end user (consumer surplus). Prices signal what products buyers would like brought to market (e.g., consumers seem willing to pay for US\$200 iPods (Apple, CA, USA) in a world where \$40 CD players provide similar sound quality). When prices are frozen for a long time, a complex array of market signals break down. Eastern European countries under socialism, such as the Soviet Union and Hungary, had systems of fixed administered prices, as did the USA during World War II, in the early 1950s and briefly during the Nixon administration in the early 1970s. However, without natural price movements, in the long run, the fixed prices must be paired with increasingly elaborate central planning and coordination systems, with ‘shadow prices’ and other control mechanisms, which fight an uphill battle to keep the diverse forces of production in balance [8]. Very few countries maintained a system of fixed prices and increasingly compensatory and complex central planning for long without collapsing.

Returning to the clinical laboratory and assuming a largely fixed-price fee schedule, we consider a general case under legacy pricing and then some special cases.

■ General case

In the general case, one would expect the fixed-price laboratory fee schedule to eventually result in surpluses of some listed tests and shortages of others. This is inconspicuous, however, because most laboratories will provide a large market basket of tests, and while they may prefer to run more profitable tests (e.g., a test with a frozen \$10 price and a current underlying \$2 test kit), the mechanisms for a laboratory to control physician test orders are relatively weak and as long as the net basket of tests is profitable, the full range of tests will still be supplied. However, the

introduction of new tests is dependent on the projected income for tests, for the clinical laboratory as well as for the test kit manufacturer. Thus, new tests whose price would be ‘underwater’ relative to development costs will probably not be produced. This is true regardless of the clinical value of the test in individual patient care, or indeed, even if the test could save the health system money in the long run.

Under the fixed-price system, where new tests are fitted into existing price categories, we can consider two special cases where innovation will still occur. These are a large marketplace where the legacy price will exceed the average test cost, and a new marketplace where the legacy price will exceed the average test cost.

■ Large marketplace & legacy price exceeds average cost

A new laboratory test, the troponin test, was widely introduced in the mid-1990s to rule in and rule out patients with chest pain that might be caused by an acute myocardial infarction. The clinical studies to assess the value of troponin were relatively straightforward. Only approximately a third of patients presenting with new onset chest pain in the emergency room have an acute myocardial infarction, and the value of troponin lay in identifying these patients a few hours earlier than prior tests could do. The clinical trial was prospective and conceptually straightforward, run in a superiority trial against an accepted ‘gold standard’. In retrospect, we see that the relatively clean clinical decision for which the test was used, the relatively modest study populations required and the very rapid collection of data in each patient signaled a test that could do well under a legacy pricing system. The research costs in this case could probably be borne either by government-funded research or by the test manufacturer industry.

■ New marketplace & legacy price exceeds average cost

Point-of-care testing is becoming more common through the introduction of small, reliable devices (as have long existed for glucose self-testing) to analyze biomarkers at the site of care, such as methicillin-resistant *Staphylococcus aureus*, *Streptococcus* (streptococcal pharyngitis), cholesterol and hemoglobin A1c. Although the costs of engineering these improved and smaller technologies are not minor, there are no significant clinical trial research costs because the tests can be evaluated primarily on *in vitro*

accuracy. If the new smaller devices and tests can be operated below the legacy price for the tests of interest, these new products can do well under a legacy pricing system. However, the driver for developing and commercializing a new test on the point-of-care platform need not be the clinical value of the test, or cost savings when other services are replaced by the test, but simply whether the new platform test can be performed at the legacy cost of the laboratory test in question.

Under other circumstances, innovation will be stymied, as described in the following section.

■ Underserved marketplace & legacy price is below the projected average cost

Using the US system as a representative one, there is a laboratory code for ‘other immunologic test’ 83520, paying a fixed price of approximately \$20. Currently, there is no good monoclonal antibody test to detect recurrent ovarian cancer. The performance of one legacy test, called CA-125, is poor. A new monoclonal antibody test that detected recurrent ovarian cancer would probably be priced at approximately \$20 using code 83520. Since there are approximately 20,000 women per year in the USA who develop ovarian cancer, and approximately 5000 who go into remission, twice-a-year testing for this population would be 10,000 tests at \$20 or a dollarized annual market size of \$200,000 (for the purpose of demonstration, I am omitting the prevalent population and multiplying only by the incident population, but as a counterbalance, I am assuming 100% utilization in the incident population). Assuming the marginal cost of the \$20 test is \$5, and the cost of service delivery another \$5, the potential margin on the new test is \$10 per test or \$100,000 per year – too little to justify an upfront clinical trial investment of tens of millions of dollars in test development. The net present value of the test will be weighed up against the upfront investment to create it. Using a 10-year investment window and a 5% discount rate on future returns, the net present value of the test is approximately \$800,000, which could be further reduced to \$400,000 for a 50% risk of technical failure during test development. \$400,000 is the ‘net present value’ and is much less than the \$10- or \$20-million investment. Meanwhile, a regulator who is required to set the government price based on only the marginal cost of the test kit (\$5) might think the

legacy price of \$20 is already too high, not too low. However, it will not matter, because the new test outlined in this example will not be developed by private investment. Whether private investment in such a test is irrelevant and should be replaced by upfront public investment is a very important consideration that I discuss in the section ‘Potential resolutions to pricing & coverage issues’.

This problem could multiply across other diseases that might be addressed by innovative molecular tests. For example, imagine ten diseases, each with 100,000 patients, and in each of which €2000 of medical costs could be saved by a novel molecular test. The test has a marginal cost of €100, but to repay research and development the market price of the test would need to be €500. If all ten tests were developed, annual test costs would be €500,000,000 and annual savings would be €2,000,000,000, for a net annual savings of €1,500,000,000 (€1.5 billion). However, we will leave this hypothetical country behind, and assume that an astute regulator signals that each of the ten tests will be priced by his government at its marginal cost, €100. Only one test is developed, by an innovator that rapidly goes bankrupt but leaves its patent to the government, which allows any laboratory to provide the test for €100. In this country, only one disease is tested and treated, with test costs of €10,000,000, annual savings of €200,000,000, and thus net annual savings in this country of €190,000,000 ... instead of €1,500,000,000. In addition, of course, the 900,000 patients with the other nine diseases still struggle with their afflictions while incurring their annual €2,000 of medical costs.

Innovative diagnostics present distinctive problems for payer decisions

Like the rationale behind legacy coding and pricing of diagnostic laboratory tests, the rationale for the practice of evidence-based medicine is not difficult to understand. Too much of medical care has been done ‘just because’ that is the way it is done. Services may be undertaken that are counterproductive for the patient, or at best, a toss-up for effectiveness. Practice patterns may vary considerably from one portion of a nation or a province to another. It would be better to have evidence for the wiser choice – should the patient with low-grade prostate cancer have a prostatectomy, for example, or is watchful waiting a better bet?

At least in the setting of medical coverage policies (will a private health payer or a government health payer cover Service X at all?), there are two major components of evidence-based medicine. The first phase is a systematic review of the literature. The second phase is a judgment phase when the data are weighed, conflicts resolved (or at least discussed) and the government or private payer determines whether the service should be covered or not.

The strategy for the first phase of this work – for undertaking the literature review systematically and for extracting and summarizing the data – will follow an explicit set of algorithms and rules. The range of publications that can be considered may be restricted by a set of languages; the years of publication that will be searched are stated; keywords and search terms are defined and then searched in an international medical database, usually Medline; as a supplement, bibliographies of available papers are searched for relevant citations; and so on. Like the steps in a chemistry student's laboratory notebook, the literature search is good science because it is replicable. The health technology reviews can be collated and performed by individuals who are skilled in basic epidemiology but do not necessarily have great clinical expertise in the subject matter under review.

There is nothing wrong with the now well-established approach to systematic literature reviews that I have just described. No longer will a large trial be ignored in silence in the expert review written by one elder authority who disagreed with its result. If data are dropped, and a large proportion of clinical publications may be dropped, it will be because they fall short on the objective levels of evidence scale (scales that have meta-analyses of randomized prospective trials at the top, and results of observational trials, or worse, expert opinion, at the bottom). The employment of these levels of evidence filters to classify the clinical trials can itself be validated by interobserver agreement studies.

However, what happens after the systematic literature review is undertaken and after the results are set in a collation of (preferably tabular) data? What happens next is drastically less well defined. Whether the work product will be a clinical policy guideline or a medical coverage decision of a payer, a great deal of judgment is usually involved. Like the mathematical concept that a sum is only as good as the factor with the least significant digits, the final two-part policy document may have a plethora of significant digits – so to speak – in

its upfront systematic literature review, but it may hardly be 'accurate' to a factor of two, or ten, in its conclusions.

In a recent article that discussed under what conditions randomized control trials (RCTs) are – and are not – the correct gold standard, Grossman and McKenzie quote an anonymous colleague who stated that, "The fact is, you've got to read the bloody thing and think about it – and there aren't too many other rules" [9] (also see [10,11]). RCTs help establish a key factor for medical evidence, in that the index intervention should be the only difference between the test group and the contemporary comparison group. The comparison group in a prospective blinded RCT represents a population that is closely and randomly interwoven with the test group. However, all the other factors that determine an intellectually sound and clinically useful trial are up for grabs – the sites of service, the entrance and exclusion criteria, the outcomes measured, and so on. Should the comparison group receive a placebo or an active control? There are numerous pros and cons, and this can be a very complex, problematic and absolutely critical question. It is fundamental to the interpretation of a trial and its application to medical policy, but the clockwork rules of the systematic literature review, by themselves, tell us nothing at all about the correct answer, nor does the bald fact that the experimental design under review had been an RCT.

Since genetic tests for hereditary diseases became available some years ago, policy experts have discussed the issue of 'genetic exceptionalism' – whether genetic tests require fundamentally different standards than other laboratory diagnostic tests. For example, the genetic status of a grandparent and a grandchild may indicate the genetic status of the parent in between who refused to be tested. Thus, this is an 'exceptional' ethical consideration not shared by most diagnostic tests, like blood pressure, glucose or a chest x-ray.

I believe that genomic tests on archival materials raise at least two problematic issues that are a form of genomic exceptionalism for the evidence-based medicine and policy decisions that are built on the retrospective archival methods. There seems to be no accurate place for some genetic studies in the conventional levels-of-evidence hierarchy. For example, few in the scientific community, physician community or patient community would debate whether the cystic fibrosis gene causes cystic fibrosis, or the huntingtin gene causes Huntington's disease.

However, in the filter of evidence-based medicine by study design, these conclusions are based on a very low level of evidence – observational data, supported by an even worse level of evidence, expert opinion (that the observational data are correct).

In molecular diagnostics, there are increasingly situations where retrospective molecular studies of archived patient samples provide convincing data that a relationship exists between the outcomes of the molecular test *in vitro* and events in the clinic, such as drug response. Breast cancers that are negative for the Her2/neu protein are unlikely to respond to trastuzumab therapy. Colorectal carcinomas with a mutation that constitutionally activates the EGFR/KRAS pathway simply cannot respond to EGFR-targeted monoclonals at the cell surface [12]. We will probably face many more situations where retrospective studies on archived specimens show a very strong relationship to an important clinical outcome.

So what are the problems? First, when this retrospective data is substantial, a RCT may become difficult to conduct. Given the retrospective molecular data suggesting that a drug is very unlikely to work for a patient with a certain genomic marker, a trial that randomizes the patient to that drug after testing for the marker may be considered unethical owing to the violation of clinical equipoise between the two arms of the trial. The study sponsor may be unable to obtain institutional review board approval for a prospective RCT design. Or, even if the trial design is approved for (hypothetical) consenting and informed patients, actual enrollment may be extremely difficult to obtain and the trial may languish with few patients. Second, the whole situation creates an impasse if the standard for a quality trial, as enforced by payers, will be a RCT that no one can now undertake. The subjective character of the resulting decision-making is evidenced by regulatory differences between the USA and Europe, where European approval for erlotinib based on *KRAS* mutation status appeared much earlier than the corresponding approval in the USA under the US FDA, and where the US delay hinged on uncertainty regarding the use of retrospective genomic data.

Potential resolutions to pricing & coverage issues

Some of the problems described earlier can move toward a resolution if we reframe the questions that we ask, but we will also see that there is a

relationship between pricing and coverage decisions. As economists have discussed for a couple of centuries now, the right price is difficult to define in the abstract. Is the right price for a product its marginal cost? Is it triple the marginal cost? Is the right price ‘what the market will bear’? – this raises problems when the unit of valuation is pegged to the price of a human life-year. Is the price determined by alternatives in the healthcare treatment pathway and their cost? Is it determined by entrants to the marketplace? If I remember my coursework in economics, finance, marketing and strategy correctly, each of these serial hypothetical questions can be answered by ‘maybe’.

I propose that one streetlight on the dark road of health economics for molecular tests is that the healthcare system should be designed to pay an amount that will bring products to market that have high clinical utility. The clinical value may be defined in QALYs – under that metric, all things being equal, a \$1000 test that extended your life by 1 year would probably be a good deal. Some molecular diagnostic tests may be significantly cost saving although lifespan or QALY is unaffected. In those cases, it may be economically efficient for society to share savings, when savings are created, with those that developed the cost-saving test. For example, if a \$1000 test saves \$20,000 of chemotherapy, then the test will be ‘self-funding’ for the healthcare system, if priced at the \$1000 cost. Foreseeable objections are that this will impose excess costs on society and direct excess profits to the luckiest of the test developers. However, I ask those who hold this objection to slow down. The goal is not to pay more than will bring the highly useful product to market – rather, the suggestion is to pay enough to bring the product to market and avoid a ‘valley of death’ where prospective entrants will refuse to invest and enter a market because they can see that the expected investments and the projected returns on the test yield a negative value to the investor; and as in the example here, absent the test society’s healthcare costs are higher than they would be with the test brought into the market.

Many factors in a competitive market will tend to drive costs downward over time. Even the most capitalist of monopoly suppliers will acknowledge that in a free market economy, both the buyer and the seller act as alert and intelligent businessmen. Furthermore, the supplier knows that, as the price rises, the marketplace of purchasers becomes smaller, so that the costs and risks of development are also amortized over a smaller population.

Since the marginal costs of molecular testing (and I include both nucleic acid and proteomic testing) will continue to drop, it is likely that the fast followers will often be able to enter a diagnostic market place with alternative tests that cost less to develop because a basic biologic principle had been established. This is a general feature for many modern technologies – patents or other intellectual property protection should be high enough to encourage innovation but not so high as to eliminate competition. Barriers to entry can be reduced, and competition thus encouraged, by government encouragement of precompetitive research or precompetitive standardizations that apply across many companies in an industry.

It is very important to discuss another fundamental approach to lowering test prices. This approach entirely shifts the sources of investment from private to public funds, and then releases the results as ‘generic’ molecular tests that can be reproduced at relatively low cost. In my experience, this approach is very popular with academically based researchers. However, for that audience, I would like to draw on the model that for other modern technologies, including healthcare technologies, this is not usually the way innovations are commercialized. For example, no country to my knowledge has been seen to invest \$800 million in a Phase III drug trial in order to create a generic drug on which neighboring countries will have a free ride. This could have happened – one or more new statins might have been developed in the past decade entirely with government funds in one or another European country – but this just does not seem to occur. In addition to the free-rider problem, there may be a second form of concern. There may be a consensus in many countries that highly risky technology investments are usually best chosen and managed by individuals directly bearing the incentives and risks.

By analogy, I would predict that few countries are likely to repeatedly invest \$10 million, \$20 million or \$30 million to develop a long series of clinically validated molecular tests that are then released for unrestricted use anywhere. However, there probably will be continue to be good examples where relatively small trials conducted with public funds provide very important tests. When this occurs, there will be a resulting multitier price dynamic in diagnostics as there is in other healthcare treatments (where we find, for example, a marketplace in which relatively resource-intensive

intensity-modulated radiation therapy for prostate cancer competes head-to-head with watchful waiting). Without question, we will gain on the pricing problem if we recognize its major outlines and debate the best solutions. To date, even authors who are highly concerned with the progress of research from the laboratory to the regulator, from the marketplace introduction to the impact on public health, have put little emphasis on the devils hidden in the microeconomics [13]. However, chasing these microeconomic and policy devils will be particularly important for tests that are not directly embedded in Phase III drug trials.

When a simply cheaper and equivalent alternative is not available, in order to pay what is necessary to trigger prudent investors and to bring highly useful diagnostic tests to market, we need to be able to tell what is highly useful and what is not. We will need professionals at payer entities who can do this. In the free market we spoke of, this is the requirement for the buyer to be a savvy businessman as well as the seller. I suspect that 30 years ago, being a medical officer at a private or public health plan was a very low-profile job and not likely to be a career goal for the ambitious. While we remember the great surgeons and medical scientists of recent decades, prominent deans of a few medical schools and leaders in public health, it is hard to find any record whatever of the great health insurer medical directors of the 1960s or 1970s. However, we are in the midst of a sea change. Sean Tunis, who was the Chief Medical Officer of the US Medicare program in the early 2000s, is a well-known author on technology evaluation and comparative effectiveness studies and now heads a research institute on these topics in Baltimore (MD, USA) [14,15]. Sir Michael Rawlins, Head of the NICE in the UK, is also an internationally respected medical policy author, and his 2008 Harveian oration stands as a substantial contribution to the field of medical policy decision-making [16]. So payment or payer entities are being ‘resourced’ now at their leadership levels with professional staff in a way that simply did not exist 40 years ago. International examples of diverse policy evaluation institutes are provided [102–106].

One key area that has seen recent and encouraging growth is the ‘thought capital’ of how we assess the external validity of clinical trials [17–19]. Most of these factors apply both to prospective randomized trials and to the retrospective, but controlled, trials and to other designs, such as

nested-control studies [20]. Close and explicit examination of the clinical problem at hand may help us to drill down on whether existing studies, placed in a clinical context, are adequate or not for a coverage decision [21]. We need to ask, what facts or assumptions are needed to bridge between the results of the test (e.g., what is the tumor's *KRAS* mutation status?) and the clinical question at hand (e.g., 'is this tumor resistant to therapy X?'). In this case, I would suggest that the three key questions are: whether the test is accurate (molecular tests usually are); next, whether the clinical population was relevant (it was as well selected as the Phase III trial, because the population was the Phase III trial); and finally, whether the result is likely to be due to chance (a greater risk with genome-wide association studies, and a risk that should be neither over- nor under-weighted). None of these three key questions is particularly better answered by a randomized prospective trial.

On the other hand, if we try to bridge the results of a warfarin pharmacogenetics test (e.g., what is the patient's *CYP2D6* status?) and the clinical question, 'what should the starting warfarin dose be?', we quickly encounter a bramble bush of thorny questions, including:

- What is the patient's age?
- What is the patient's diet?
- What are the patient's risk factors for bleeding?
- What is the frequency of international normalized ratio monitoring?
- What is the relationship between international normalized ratio laboratory value and stroke?
- What is the patient compliance with medication?
- What is the patient compliance with international normalized ratio testing?

The value of warfarin genetics in routine clinical settings may vary along so many diverse dimensions that only a prospective and randomized trial with contemporary controls is likely to be useful for medical policy-makers.

There will probably be a transition period between the growth of innovative thought capital in the academic literature and its application in the day-to-day world payer decision-making. I note that, for a number of years, one textbook on evidence-based medicine has segregated study of therapeutic trials, in the first half of the book, from the analysis of diagnostic test trials, in the second half [22].

However, to date, separate standards for therapeutic and diagnostic test trials are found at few healthcare technology assessment agencies. In the USA, for example, Medicare and the Blue Cross Blue Shield Association (IL, USA) private insurers [107] do not appear to have different guidelines for the evaluation of diagnostic tests, but the more academically oriented CDC (GA, USA) do [108]. The important thing is that the problems of diagnostic tests and clinical trial extrapolation are increasingly recognized in the academic literature now, which is a prerequisite for eventual transfer to the worldly practice of medical policy-making [109].

Conclusion & future perspective

Molecular diagnostics provide insight into the individuality of human diseases, and this transition will, eventually, markedly transform medical practice. However, some of the problems facing the adoption (or commercialization) of personalized medicine diagnostics are only now emerging clearly. Since it appears to be difficult or ultimately dysfunctional to nationalize all aspects of medical research for drugs, devices and diagnostics (by paying for all research and development through the government and providing all medical devices and products through generic manufacturers at marginal cost), there will need to be a rational marketplace where investments, risks and rewards are distributed in a way that benefits the healthcare system as a whole. It will be increasingly difficult to create diagnostics based on substantial, costly and risky clinical trials if payments for the resulting tests are pegged to legacy generic values. In some cases, the result will be so dysfunctional as to block development of cost-saving tests, which makes no sense at all. Insightful discussions about optimal public policies will improve our approach to test commercialization and coverage and bring advanced molecular diagnostics to those medical problems where they are most useful and most needed.

Financial & competing interests disclosure

The author has no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

Executive summary

Payment systems for diagnostic tests

- Legacy payment systems for diagnostic tests were developed in the 1970s and 1980s, and are creating problems that were unforeseeable.
- Pricing of novel diagnostic tests at legacy generic prices will create shortages of useful and efficient products.
- Policy-makers are necessarily penalized by creating pricing rules that prevent the development of products that do not yet exist.

Evidence-based medicine meets genomic diagnostic tests

- Levels of evidence rules favor prospective controlled trials, but molecular diagnostics may be fully validated based on only observational studies (e.g., identification of the cystic fibrosis gene).
- Rules for the selection of studies to be included in literature reviews may disregard genomic test data that is not prospective.
- Genomic tests create an 'exceptional' situation based on the potential for many validated studies using archival samples – this is unfamiliar based on familiar trial design for novel drugs and surgeries.
- The ability to undertake powerful retrospective trials may yield a 'winner's curse' if payers are unwilling to accept retrospective data, even when it makes prospective trials unethical.

Potential resolutions

- It is unlikely that all research and development for molecular diagnostics will be nationalized, although that would allow for generic tests to be delivered at commodity prices.
- Tests delivered at value-based prices will attract competitors with lower costs.
- Policy-makers should look 'two steps ahead' to inform wise price-setting policies for molecular diagnostics.
- Elevation and discussion of these problems will be crucial to resolving them.

Bibliography

Papers of special note have been highlighted as:

▪ of interest

▪▪ of considerable interest

- Burken MI, Wilson KS, Heller K, Pratt VM, Schoonmaker MM, Seifter E: The interface of medicare coverage decision-making and emerging molecular-based laboratory testing. *Genet. Med.* 11(4), 225–231 (2009).
- Mckinnon RA, Ward MB, Sorich MJ: A critical analysis of barriers to the clinical implementation of pharmacogenomics. *Ther. Clin. Risk Manag.* 3(5), 751–759 (2007).
- Goodman M: Pharmaceutical industry financial performance. *Nat. Rev. Drug Discov.* 8(12), 927–928 (2009).
- Chackalamannil S, Desai MC: Personalized medicine: a paradigm for a sustainable pharmaceutical industry? *Curr. Opin. Drug Discov. Devel.* 12(4), 443–445 (2009).
- Gold ER, Carbone J: Myriad Genetics: in the eye of the policy storm. *Genet. Med.* 12(4), S39–S70 (2010).
- Striking discussion of how different countries have developed diverse policy approaches to genetic patents.
- Deverka PA: Pharmacogenomics, evidence, and the role of payers. *Public Health Genomics* 12(3), 149–157 (2009).
- Djulbegovic B: The paradox of equipoise: the principle that drives and limits discoveries in clinical research. *Cancer Control* 16(4), 342–347 (2009).
- Expands on this article's discussion of clinical equipoise and contains an excellent bibliography.
- Kornai J: *Overcentralization in Economic Administration (2nd Edition)*. Oxford University Press, Oxford, UK (1994).
- Grossman J, MacKenzie FJ: The randomized controlled trial: gold standard, or merely standard? *Perspect. Biol. Med.* 48(4), 516–534 (2005).
- Kristiansen IS, Mooney G: *Evidence-Based Medicine in Its Place*. Routledge, London, UK (2004).
- Best comprehensive book on the potential pitfalls that are found among the advantages of the evidence-based medicine approach.
- Ho PM, Peterson PN, Masoudi FA: Evaluating the evidence: is there a rigid hierarchy? *Circulation* 118(16), 1675–1684 (2008).
- Good presentation of reasons why the evidence hierarchy should not be taken as an automatic good.
- Siena S, Sartore-Bianchi A, Di Nicolantonio F, Balfour J, Bardelli A: Biomarkers predicting clinical outcome of epidermal growth factor receptor-targeted therapy in metastatic colorectal cancer. *J. Natl Cancer Inst.* 101(19), 1308–1324 (2009).
- Khoury MJ, Gwinn M, Yoon PW, Dowling N, Moore CA, Bradley L: The continuum of translation research in genomic medicine: how can we accelerate the appropriate integration of human genome discoveries into health care and disease prevention? *Genet. Med.* 9(10), 665–674 (2007).
- Luce BR, Kramer JM, Goodman SN *et al.*: Rethinking randomized clinical trials for comparative effectiveness research: the need for transformational change. *Ann. Intern. Med.* 151(3), 206–209 (2009).
- Garber AM, Tunis SR: Does comparative-effectiveness research threaten personalized medicine? *N. Engl. J. Med.* 360(19), 1925–1927 (2009).
- Rawlins M: De testimonio: on the evidence for decisions about the use of therapeutic interventions. *Lancet* 372(9656), 2152–2161 (2008).
- Outstanding article discussing the gaps between evidence review and the ultimate decisions of payment agencies.
- Mercer SL, Devinney BJ, Fine LJ, Green LW, Dougherty D: Study designs for effectiveness and translation research: identifying trade-offs. *Am. J. Prev. Med.* 33(2), 139–154 (2007).
- Example of the increasingly academic and intellectually rigorous approach to extrapolating from clinical trials to clinical practice. This will be increasingly important when the evidence is from retrospective genomic trials.
- Dekkers OM, von Elm E, Algra A, Romijn JA, Vandenbroucke JP: How to assess the external validity of therapeutic trials: a conceptual approach. *Int. J. Epidemiol.* 39(1), 89–94 (2010).
- Example of the increasingly academic and intellectually rigorous approach to extrapolating from clinical trials to clinical practice. This will be increasingly important when the evidence is from retrospective genomic trials.
- Rothwell PM: Factors that can affect the external validity of randomised controlled trials. *PLoS Clin. Trials* 1(1), E9 (2006).

- **Example of the increasingly academic and intellectually rigorous approach to extrapolating from clinical trials to clinical practice. This will be increasingly important when the evidence is from retrospective genomic trials.**
- 20 Anderson GL, McIntosh M, Wu L *et al.*: Assessing lead time of selected ovarian cancer biomarkers: a nested case–control study. *J. Natl Cancer Inst.* 102(1), 26–38 (2010).
- 21 Simon RM, Paik S, Hayes DF: Use of archived specimens in evaluation of prognostic and predictive biomarkers. *J. Natl. Cancer Inst.* 101(21), 1446–1452 (2009).
- 22 Riegelman RK: *Studying a Study and Testing a Test (5th Edition)*. Lippincott Williams & Wilkins, PA, USA (2005).
- **Excellent textbook on evaluating the medical literature, and distinctive in its segregation of therapeutic trials from diagnostic test trials.**
- **Websites**
- 101 Clinical Laboratory Fee Schedule: overview www.cms.hhs.gov/ClinicalLabFeeSched/01_overview.asp#TopOfPage
- 102 National Institute for Health and Clinical Excellence (NICE) www.nice.org.uk
- 103 Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen www.iqwig.de
- 104 Australian Government Department of Health and Ageing: Review of Health Technology Assessment in Australia (HTA Review) www.health.gov.au/internet/main/publishing.nsf/Content/hta-review
- 105 Agency for Healthcare Research and Quality www.ahrq.gov
- 106 Center for Medical Technology Policy www.cmtpnnet.org
- 107 Blue Cross Blue Shield Association: Service Through Research www.bcbs.com/blueresources/tec
- 108 Centers for Disease Control and Prevention: Evaluation of Genomic Applications in Practice and Prevention (EGAPP™) www.cdc.gov/genomics/gtesting/egapp/index.htm
- 109 NICE Diagnostics Assessment Program www.nice.org.uk/aboutnice/whatwedo/aboutdiagnosticsassessment/diagnosticsassessmentprogramme.jsp
- **Commendable effort to develop a distinct technology assessment program fitted for diagnostic tests.**