The Prospects for “Personalized Medicine” in Drug Development and Drug Therapy

J Woodcock

There has been much recent discussion about the advent of “personalized medicine,” but controversy exists over its exact definition; how, when, and whether it will be brought about, and what means could be used to measure its attainment. In fact, the concept of “personalized medicine” is a sort of shorthand used to represent the logical next steps in progression of medical science toward greater mechanistic understanding of health, disease, and treatment. This shorthand is attractive to the public because of its intuitive appeal, and irritating to the biomedical community because it glosses over the very real scientific and implementation challenges. This paper evaluates the trajectory and promise of these next steps for the currently problematic states of both drug development and therapy.

So what is “personalized medicine?” At its heart, it is nothing more than what medicine has always been at its best – the careful evaluation of the health (and health prospects) of an individual based on the best information obtainable about that person’s physical and mental state. The notion of “personalized medicine” stems from the recognition that now, or very soon, unprecedented types of information will be obtainable – information that will enable practitioners to parse states of health, disease, and treatment responsiveness into hitherto-unknown additional categories. Just as the ancient diagnoses of “dropsy,” “apoplexy,” or “rheumatism” have given way to multiple distinct disorders, currently understood diseases will be further refined by the application of new genetic, genomic, proteomic, imaging, etc, technologies. And, very important, it is hoped that new diagnostics may shed greater light on individualizing therapy, an area that has seen little progress outside of infectious disease.

Are these anticipated diagnostic advances fundamentally different from existing medical science in any way? Early twentieth century medical practitioners were masters of diagnosis – the art of the detailed history and physical examination. Cumulative experience led to syndromic descriptions that were written up in medical textbooks and diligently instilled by example on the wards. Few treatment options existed. Although advances in basic sciences have contributed numerous new diagnostic technologies during the past half-century, these diagnostics generally illustrate, rather than explicate, pathology. Much of medical practice remains semiempirical, i.e., trial and error: medicine is still more an art than a science. The last 30 years brought an explosion of therapeutic options that have been adopted based on empirical evidence from clinical trials. “Personalized medicine” represents the hope that the next decade will bring a comparable explosion of diagnostic tools that will provide mechanistic insight into pathogenesis and treatment response, and will help transform the scientific basis of therapy.

URGENT NEED FOR CHANGE
Currently, serious concerns exist about the state of drug development and the state of the health care system, and, to some extent, the problems identified in these sectors are linked. Although each struggles with unique issues, both sectors suffer the consequences of the immense variability of human disease and our current inability to make good predictions about the outcomes of interventions.

The “productivity problem” in the pharmaceutical sector has been extensively discussed and documented. Basically, pharmaceutical research and development investment has doubled, in real dollars, during the past decade. No corresponding increase in new chemical entities brought to market has been observed. In fact, the rate of submission of new chemical entities to regulatory bodies has decreased worldwide. Development costs per successful entry have increased markedly, particularly the costs of clinical trials. At the same time, the success rate in clinical phases of drug development has declined. These facts taken together mean that overall development costs per successful new chemical
entity are on a steeply rising trend. Although the actual costs are debated and obviously vary by individual case, it is estimated that development of a successful new chemical entity represents, on average, a $750 million to $1.1 billion research and development investment. This escalating trend in drug development costs is unsustainable, particularly given the state of health care. If this trend cannot be controlled, the very enterprise of drug development will be threatened.

Concurrently, in the United States, the issue of the affordability of health care is a major societal concern. The cost of drugs, representing an out-of-pocket expense to a segment of the population, has become emblematic of the overall affordability issue, leading to significant hostility toward the pharmaceutical industry. Insurers are increasingly concerned about the effects of highly expensive but useful new pharmaceutical interventions on their ability to manage their businesses. Public discussions about pharmaceuticals are most often centered on cost control issues.

Significant concerns about the US health care system are focused on quality in addition to affordability. These two properties are undoubtedly linked. Since the landmark report on health care quality, “To Err is Human,” from the Institute of Medicine, many health care quality improvement initiatives have been initiated. Most of the explicit quality initiatives related to pharmaceutical use have centered on preventing medication mix-ups, dosing errors, transcription problems, and the like (e.g., bar coding, e-prescribing systems, etc). Less attention has been paid to the most significant source of preventable harm related to pharmaceuticals: suboptimal prescribing (errors of both commission and omission). Efforts to improve the quality of prescribing have generally originated from earlier cost-control strategies developed for managed care formulary oversight. These strategies have broadened over time to encompass the assessment of the value of drug therapies and the promulgation of evidence-based best practices. Development of evidence-based therapeutic guidelines has been seriously constrained by the lack of what are termed “effectiveness” studies, i.e., outcome studies in the relevant settings and populations. This evidence gap has led to calls for additional requirements for drug approvals, or new types of required studies after approval. There is growing acknowledgement that the typical drug development program, although extremely costly to manufacturers, does not yield all the information needed to make treatment recommendations in the context of the health care system. Thus, the information needs of health care are perceived to be at odds with the need to contain costs during drug development. The problems of pharmaceutical productivity and the travails of the health care system are on a collision course.

LIMITATIONS OF THE CURRENT CLINICAL DRUG PERFORMANCE ASSESSMENT MODEL
The randomized clinical trial evaluation protocols currently used to assess the clinical performance of drugs represented a triumph over the anecdotal approach to medicine that was in general use before the second half of the twentieth century. Amendments to the Food, Drug and Cosmetic Act in 1962 created the requirement that drugs be shown to be effective before marketing in the United States. For most newly developed drugs, this meant that statistically significant positive results from at least two randomized, well-controlled efficacy trials needed to be submitted to the FDA as part of a marketing application. Randomization addressed certain biases and the effect of “random” (unexplainable) variability.

The randomized controlled trial (RCT) is a powerful tool that has provided the evidentiary basis for many of the advances of modern medicine. However, RCTs have, as currently executed, certain limitations. Although theoretically, one can answer, using an RCT, almost any question that can ethically be asked, one can answer only one or a few questions per trial. Unfortunately, there are an unlimited number of questions about the appropriate use of drugs and the outcomes of such use, and these questions evolve over time. At the same time, there is a decidedly limited universe of funding, patients, investigators, time, and resources available to conduct trials to answer these questions. And it is not always easy to know the right question to ask.

For example, one question frequently arising in health care debates relates to comparative outcomes. How do various drugs or regimens for the same condition compare, for example, with respect to safety, effectiveness, and tolerability? Finding the difference between two active therapies or regimens is usually much more challenging than simply showing that an intervention is better than a placebo. Therefore, advances in cardiovascular disease and cancer have been built on a series of very large randomized studies, each comparing the best treatment established in prior trials with a new intervention (e.g., adding another drug). This approach has been extremely successful, but it is very expensive and time-consuming. In some cases, by the time the trial was completed, the original question has lost its relevance (e.g., owing to superseding medical advances.) And resources have been lacking to apply this method to many diseases.

However, more fundamental, is this approach to comparisons always asking the right question? The usual interpretation of an RCT comparing two interventions in a population assumes that if a “superior” intervention is found, it will be the best choice for any given individual in that population. In fact, a treatment with a 10% advantage over a comparator could still be the wrong drug for many people. And a drug with a severe side effect may be the best treatment for people who are not at risk for that problem. The diagnostic question, “can we identify people for whom this is the right drug?” is fundamentally different from asking which treatment performs best in the population overall.

RCTs in many drug development programs often are less informative than desirable, owing to constraints imposed by the developers. For example, historically, assessment of the optimal dose range for new drugs has often been inadequate,
so that dosing recommendations frequently need to be changed after marketing. Phase 2 trials exploring dose usually employ a randomized, parallel arm fixed-dose design. In many cases, an adaptive design would allow exploration of a wider dose range and ultimately more experience with selected doses. However, sponsors, because of tradition, perceived regulatory disfavor, concern about the less structured approach, or other factors, do not usually conduct adaptive dose-finding trials. Additionally, one-size-fits-all dosing obviously is a great source of response variability (owing to polymorphic metabolism or other individual differences). Nevertheless, the impact of dose adjustment is usually not vigorously explored because the commercial success of a drug requiring such adjustment is deemed to be dubious.

More generally, clinical trials in drug development programs and effectiveness studies are ordinarily designed to explore differences between treatments and controls, not among individual responses in the treated groups. As a result, trials of agents that are very effective in a subgroup of the population studied will frequently produce non-statistically significant results, or may result in overall treatment effects that are clinically insignificant, leading to the conclusion of minimal effectiveness. Agents that cause a specific side effect in a subgroup may be found too toxic for overall use. Of course, in the absence of diagnostic methods that reliably distinguish them, there is no way to study such subgroups. Herein lies the promise of “personalized medicine” to refine the empirical approach used in most clinical trials by incorporating powerful new diagnostics that can identify individual predictive characteristics and better control variability.

More efficient ways to evaluate drug performance during development and in the clinic must be created. Otherwise, the burden of uncertainty the unanswered questions – will continue to drive up costs, inflame controversy, and impede progress in both the health care and pharmaceutical development sectors.

**OPPORTUNITIES FOR IMPROVEMENT: PERSONALIZED MEDICINE**

The new diagnostic modalities that form the basis of the personalized medicine concept will be used to manage variability in numerous scenarios during drug development and use, including the following.

**Drug metabolism**

The inter-individual variations in drug exposure that result from polymorphisms in drug metabolizing enzymes have been known for many years. However, easily performed diagnostic tests to identify people with important metabolic differences have not been available until recently, and most are still being developed.

Drug developers generally seek to avoid candidates with polymorphic metabolism, but this is not always possible. Significant variability of drug exposure in the test population will increase the variability of response, decrease the power of trials to detect effectiveness, and lower the overall chance of success of a drug development program. Variability of this nature is particularly concerning in drug development programs such as those pursued in oncology, where drug dosing is often set at a determined “maximum tolerated dose”. In cases where slow metabolism generates a significant population with higher drug exposure at a given dose, the empirically determined maximum tolerated dose could reflect toxicities resulting from their exposure. Slow metabolizers could be overdosed and average metabolizers underdosed (in maximum tolerated dose terms) throughout development. However, developers are reluctant to pursue pharmacogenetically dosed trials owing to the aforementioned perceived market resistance to dose adjustment as well as possible difficulties in assay development and commercialization.

Although individualization of therapy based on drug metabolism could be important in selected drug development programs, it is a key opportunity for improving the outcomes, particularly the safety, of existing therapies. Numerous drugs on the market, including drugs with narrow therapeutic indices (e.g., warfarin), are metabolized by polymorphic enzymes. Genetic testing technologies are rapidly evolving, and it can be anticipated that assays reporting on multiple metabolizing enzymes will be available in the not too distant future. However, uptake by the clinical community will require “proof of concept” studies that demonstrate the benefit of testing in the individual case. Considerable questions remain about how this is to be accomplished.

**Definition of disease subsets**

Many “diseases” are clinical syndromes, defined observationally, that undoubtedly are made up of a collection of distinct pathogenic states. New genomic, proteomic, imaging, and other diagnostic modalities will provide better discrimination among these syndromes by providing more information on the underlying pathogenic processes, without necessarily providing full mechanistic, explanatory data.

Significant opportunities are presented by diseases with purely “clinical” diagnoses such as many psychiatric disorders, ADHD, and the like. Many of these disorders have highly variable and often poor responses to therapeutic interventions, potentially driven, at least in part, by differences in underlying pathology. Similarly, diseases with highly variable responses to therapy (e.g., asthma), not explained by drug metabolism differences, can be evaluated for predictors of treatment–response subsets.

Cancer diagnosis, staging, and therapeutic decision-making are active examples. Current approaches to treatment choice rely heavily on histology, tumor burden, and tumor distribution, but it is known that, for many tumors, additional factors significantly influence natural history. Additional diagnostic biomarkers such as gene expression assays are being explored as indicators of the underlying
tumor biology that may provide more accurate risk stratification both for standard therapy and in-trial designs. The above examples demonstrate how new diagnostics could decrease variability of treatment–response by identifying responsive subgroups, thereby increasing the size of the treatment effect, decreasing the size of needed clinical trials, and minimizing the number of exposed individuals who fail to benefit from therapy (in some cases, such individuals may contribute disproportionately to the adverse event experiences). These steps would result in an improved benefit/risk analysis and would increase the value of the intervention. In addition, identification of treatment–responsive subgroups may contribute to an understanding of disease pathogenesis.

**Targeted therapy**

Modern drug discovery focuses on cellular targets for drug intervention. Identified target molecules are usually believed to be participants in the pathogenesis of the condition of interest or in its potential mitigation. In many cases, the target molecule itself may be polymorphic (e.g., B2 adrenergic receptor), have variable expression, or be subject to somatic mutations that affect expression or function (cancer). In such cases, assays that can delineate the status of the target (e.g., at the level of receptor expression, genetic polymorphism, gene expression, somatic mutation, etc) in individual patients may provide significant predictive value for treatment response. This has been demonstrated for trastuzumab (Herceptin) and imatinib (Gleevec) in cancer therapy. However, clinical development of most targeted therapies is unaccompanied by the use of assays to evaluate target status.

In particular, current investigational cancer drugs frequently target aberrant cellular pathways, but usually are not developed in concert with diagnostic assays. This empirical approach requires the new therapy to achieve a positive benefit/risk analysis in a defined population of cancer patients independent of target status, which can be a difficult challenge. For example, the cancer drug gefitinib (Iressa), an epidermal growth factor receptor-targeted therapy, did not confer a survival advantage in a trial of non-small-cell lung cancer despite its association with documented dramatic tumor responses. Conversely, Genentech, developer of trastuzumab, states that development of that drug without incorporation of a diagnostic assay would have required a trial of 11,000 patients to achieve a statistically significant result. Such a trial (if feasible) would likely have resulted in a much smaller estimate of treatment effect, decreasing the perceived value of the therapy. Although the benefit of trastuzumab for diagnostic-test-positive individuals would have remained the same, the perception of treating physicians and patients would have been much different, the ratio of adverse events to successful treatment much higher, and the value of treatment lower.

A number of real and perceived obstacles inhibit pursuit of diagnostic-directed targeted therapies. The original concern of developers stemmed from the “blockbuster” model of drug development: seek indications that will establish a large market. Narrowing the market share by targeting a subpopulation was deemed highly undesirable from a business perspective. However, the current commercial imperative is to lower or control costs by improving the clinical development success rate. Non-targeted therapies that are marginally effective may turn out to have a substantial benefit when paired with the appropriate diagnostic. It may be possible to explore, in the same series of trials, the efficacy of targeted and non-targeted approaches. Such development programs could salvage some new drugs and improve overall success.

Technical obstacles also exist. It is often not clear what assay should be used to assess the target. Additionally, drug “targets” are usually part of cellular pathways or networks that are influenced by multiple factors occurring downstream and upstream of the identified molecule. The state of these molecules or pathways may be determinative of drug response in a given individual, reducing the predictive value of a single-target assay. Probably, the most serious technical barrier, however, is timing. It is not clear how identification, selection, and testing of an analytically valid diagnostic can be achieved in time for use in pivotal therapeutic trials. This would require planned pursuit of a co-development strategy early in the drug development program, a time when many developers are still hoping that the agent will have the widest possible market.

The availability of a biomarker for target status could rapidly improve both development and clinical practice in a given disease area. The development and use of HIV therapies provide useful examples. Treatment resistance can develop through a variety of mechanisms as the virus is exposed to the selective pressure of an antiviral regimen. Assays to evaluate responsiveness (or resistance) to therapy have been developed and these guide modern antiretroviral therapy. The existence of a variety of assays allows developers to rapidly screen new agents directed at a treatment-resistant virus, and physicians to identify patients who are candidates for standard or investigational therapy.

**Monitoring treatment response**

It is clear that not all, and often a substantial minority of, people with a given condition respond favorably to any particular therapeutic intervention that is known to be effective in that condition. The *therapeutic trial* is a pragmatic response to this reality, wherein drugs are prescribed in a serial fashion until a favorable response is achieved. Many current therapeutic guidelines recommend that physicians follow specific drug-prescribing sequences.

This approach works adequately when the response is easily assessed, the disorder is not marked by spontaneous fluctuations and, most importantly, lack of immediate success does not have serious consequences. However, in diseases such as major depression, cancer, transplant rejection, and so forth, timely, effective intervention may be of the essence. Thus, there is interest in diagnostics that can rapidly
evaluate early response to treatment. In some cases, existing biomarkers may, upon careful assessment, play this role. New technologies such as functional imaging or proteomics may also generate new biomarkers for early assessment of treatment response.

In addition to improving outcomes in the clinic, diagnostic tools to assess treatment response could be important in improving development success. Such tools would be useful not only in early studies of clinical activity (to help decision-making), but also in efficacy trials for early identification of non-responders. Minimizing exposure of probable non-responders could substantially improve the benefit/risk analysis for many drugs.

For the patient, early assessment of individual response makes great sense. Absent a marker to predict response probability in advance, ascertainment of response status as early as possible can reduce exposure to potentially toxic treatment, increase the chance of ultimate therapeutic success, and decrease the time, effort, and costs of undergoing treatment.

Prevention
Disease prevention, embodied in public health measures and vaccination, has had astounding success in improving health over the last century. Successes in cardiovascular disease during the past decades have spurred interest in using pharmacologic therapy for prevention in other disease areas, but little progress has been made. Preventive drug trials require that at-risk subjects, many of whom will not ever experience the outcome being studied, be exposed to a drug for many years. Initiating such a trial requires a persuasive argument that the foreseeable harms from the intervention would be significantly outweighed by the overall benefits. Additionally, if the outcome of interest takes 5 or 20 years to occur on average, the trial may take too long and be impractical. To address this problem, many have advocated the use of biomarkers as surrogates for treatment efficacy, the use of which would shorten trial duration. However, this approach not only is contingent upon the validity of the biomarker, but it also leaves unanswered the question of long-term safety.

The best hope for making significant progress in preventive pharmacologic therapy is risk stratification based on new diagnostic biomarkers. The biomarkers would be used in a so-called “enrichment” design that would enroll patients into a trial or trial stratum based on biomarker status. In a multistratum trial, timed interim analyses of strata could be carried out. In general, identification of a high-risk group can provide a more compelling rationale for intervention; in many cases, the high-risk group will progress more quickly to an outcome, thus shortening the needed duration of trials; and, if an intervention works, the benefit/risk analysis has the best chance of being favorable. A finding of benefit and adequate safety in a high-risk group could pave the way for trials in lower-risk individuals. Diagnostic biomarkers to stratify high-risk groups would require an evidentiary base far short of that needed for a surrogate endpoint for approval.

Drug safety
No area is more emblematic of the limitations of the current empirical approach to drug testing than drug safety assessment. Given that drug safety has long been, and is currently, an area of great controversy and attention, it is interesting that relatively little scientific progress has occurred in this area.

Drug discovery naturally focuses on desired pharmacologic activity. As candidates are developed, they are screened for undesirable characteristics, for example hydrophobicity or structural element similarity to known carcinogens or teratogens; however, the major focus is on interaction with the intended target. Although the in vitro and animal safety evaluations of candidate drugs conducted before human testing have been largely harmonized worldwide during the last decade, preclinical safety evaluation has undergone little scientific evolution. Standardized animal safety testing has the dual objectives of determining a safe starting dose for humans and establishing the major organ targets for drug toxicity. These animal safety tests are observational, not explanatory. Occasionally, when severe toxicity is encountered, further testing may be pursued to evaluate applicability to humans, for example, comparative metabolism studies. Early human safety studies are aimed at finding tolerable doses and eliminating candidates with toxicity deemed excessive, given the proposed uses of the drug. The design of the subsequent drug development program is usually driven by the need to demonstrate drug efficacy. The safety profile of the drug is generated by extensive toxicity evaluations during patient exposure in the various efficacy trials. In recent years, regulators have asked for special safety studies or evaluations in, for example, patients with renal or hepatic disease, or the elderly: most of these studies are directed at dose adjustment recommendations. Very recently, an additional study of drug effect on cardiac repolarization has been needed: this is also an empirical evaluation, the interpretation of which is difficult and controversial.

At the end of a typical drug development program, observed toxicities are added up, their frequency calculated, and, if any are severe, strategies to mitigate them are considered. There is rarely time within a typical development program to test mitigation strategies. After drug approval, great efforts are expended on drug safety both by manufacturers and regulators. These efforts usually focus on detection, enumeration, label warnings, and evaluation of ongoing benefit risk.

The emerging science of safety
New scientific tools provide tremendous opportunities to modify the safety assessment paradigm. For example, off-target effects can be evaluated by cell-based in vitro assays or assays that compare gene expression patterns with those of known agents. The tools of “systems biology” seek to
combine these and other sources of knowledge into patterns that can provide a broader view of probable drug effects. Additionally, numerous new biomarkers are being developed that will be more sensitive and specific for organ toxicity than existing methods. These biomarkers can be used to bridge animal and human toxicity findings, to closely monitor subjects for early signs of organ damage, and to develop a mechanistic understanding of safety problems.

Why do only a few of the people exposed to a drug get certain adverse events? Known reasons include vulnerability to an exaggerated pharmacodynamic effect (potentially mediated by receptor differences); differences in genetic susceptibility (e.g., G6PD deficiency) and drug–drug interactions. However, many adverse events are deemed “idiopathic” (a term in practice considered synonymous with “inexplicable”). In fact, most rare or uncommon adverse events have a predictable incidence in a given population. This certainly suggests that there are underlying, scientifically explicable root causes and that, therefore, these events might be preventable. The most immediate opportunity in this area is a consequence of the recently improved understanding of the human genome. It is now possible to investigate the genetic contribution to important adverse events.

The use of new biomarkers and safety tools, combined with active surveillance tools in the health care environment, has the potential to revolutionize approaches to drug safety, moving from detection and enumeration to prediction, prevention, and active management of drug toxicity in the future. To ensure the best outcomes for patients, drug safety must be “built in” to the development and use processes. Enumerating harm is not enough.

Towards personalized medicine
Diagnosis is the foundation of medicine, and new diagnostics – imaging technologies, in vitro assays, and others – will provide the foundation for personalized medicine. It is easy to cast a skeptical eye on any of the scenarios outlined above, but taken together they offer unmistakable opportunities. Nevertheless, multiple barriers remain in the way. Most important is the bias toward intervention that pervades the health care system (including the reimbursement system) and, to some extent, the research community.

Another significant problem is a misunderstanding of the role of biomarkers in drug development. The use of biomarkers as described above has become confused with their use as “surrogate end points” for drug regulatory approval. These uses are completely different, and yet discussions of biomarker use so often degenerate into arguments about “validation” of surrogates. Drug and diagnostic developers imagine that the nearly insurmountable evidentiary barriers to the use of surrogate end points apply to other regulatory uses as well.

A final barrier exists in commercial development. The lack of major financial incentives for diagnostic product development often impairs the rigorous pursuit of evidence needed to understand their clinical value. Drug developers are unsure about the value of novel markers during the risk- fraught clinical development process. The conceptual framework and regulatory pathways for drug diagnostic co-development need to be articulated.

Because the benefits of better understanding will accrue to many parties and because no one entity has all the data, resources or mandates to accomplish what needs to be done, there is increasing interest in using public–private partnerships and other consortia to develop clinical data relevant to biomarker use in pharmacotherapy. As discussed in the article by Lesko in this issue and recently, FDA is spearheading some of these efforts through its Critical Path Initiative. FDA is working with the National Institutes of Health and its institutes, with the pharmaceutical, biopharmaceutical and diagnostic industries, and with private foundations to establish novel partnerships. In the European Union, the Innovative Medicines Initiative is proposing to study safety and efficacy biomarkers. 8 Other efforts are underway worldwide.

These efforts cannot come too soon. There is an urgent need to improve the efficiency, information yield, and predictability of drug development; the safety and effectiveness of pharmaceuticals; the quality of health care and treatment outcomes for patients. These improvements can simply not be achieved by continued near-exclusive use of empirical methods of generating clinical knowledge.

“Personalized medicine” is the future. The only remaining question is how soon it will come about.

CONFLICT OF INTEREST
The authors declared no conflict of interest.

© 2007 American Society for Clinical Pharmacology and Therapeutics