An idea, once born, can never more be undone.

Friedrich Duerrrenmatt, The Physicists

About 25 years ago, a fundamental change in the conduct of pharmaceutical research occurred: the baton was passed from the chemist to the biologist as the driver of innovation. Previously, the biologist – who was first and foremost a pharmacologist and physiologist – was relegated to the role of testing new compounds synthesized by chemists. More recently, empowered by rapid progress in cell and molecular biology based understanding of disease, biologists developed the ability to formulate rational hypotheses of disease-relevant mechanisms and began to present the chemist with putative cellular targets that then became the starting point for drug discovery – inverting the previous sequence. Commonly, these targets were derived from a mechanistic physiological understanding of how biochemical pathways affect certain phenotypes. In parallel, the development of tools to determine molecular structure generated great enthusiasm for ‘rational drug design’. As the limitations of this approach became disappointingly clear, the emphasis of drug discovery shifted to high-throughput, automated screening of as large and as diverse a collection of molecules as possible. This primary approach, initiated and driven by the biologist, is then followed by a dedicated, systematic effort to improve a ‘hit’ by the medicinal chemist. Highly potent, specific, and selective drugs thus developed, e.g. proton pump inhibitors, HMG-CoA reductase inhibitors, and serotonin re-uptake inhibitors, to name a few, have changed the face of medicine.

Despite these dramatic strides, a critical dilemma remains: our medicines do not work well for everyone. As physicians in clinical practice will testify on the basis of daily experience, a medicine can be well-tolerated and dramatically effective for Patient A, only to be ineffective or
Impact of genomics on healthcare

harmful in Patient B. This cumulative experience of unpredictability with respect to the use of pharmaceutical products shapes physician and patient attitudes enormously, and with good reason. An understanding of this fundamentally unsatisfactory situation also helps to explain the repeated outcry about the high costs of new medicines, which in public discourse greatly overshadows any acclaim for the pharmaceutical industry's dramatic successes. Certainly, in the current climate of cost containment, this 'hit-or-miss' quality of successful prescribing is no longer tenable. The most pressing questions facing the pharmaceutical industry at the end of the twentieth century are deceptively simple ones: what are the reasons that medicines work differently in different individuals, and how can we predict these differences? The answers will be found in pharmacogenetics.

Before grappling with these most fundamental issues, however, one would like to understand the reasons that biology-based efforts to produce effective medicines have led to this state of only partial success. The principal reason is a familiar one in the history of science: the confusion of association with causality. Biology-driven drug discovery is based on observed phenomena that imply and elucidate underlying physiological mechanisms without defining them precisely. The strong association of such a mechanism, or the response to a drug intervention with altered function in an organism does not ultimately provide a causal explanation for the alteration, though it permits the development of better hypotheses. Thus, one can measure improved glucose metabolism after administration of insulin without learning the cause of the insulin deficiency, and without appreciating the existence of factors modifying insulin sensitivity. Similarly, one can administer an antibiotic and observe a bactericidal effect associated with recovery of the patient from pneumonia, without appreciating the role of the immune system in the curative process. When disease-causing mechanisms are simple, this associative approach can suffice. However, when diseases are etiologically heterogeneous certain subgroups of disease-sufferers will respond to a given treatment better than others. In order to understand these differences, scientific techniques that allow a better understanding of causation are essential. Therein lies the promise of strategies that look for explanations at the level of the gene.

With the marriage of long-established principles of inheritance (Mendelian genetics) to newly-found tools of molecular biology, the discipline of molecular genetics was born. A critical shift in the practice and outlook of the biosciences occurred, enabling us to progress closer to a true causative understanding of biology than ever before. By the mid to late 1980s, single gene diseases began to be unravelled at an increasing pace, demonstrating clearly the power of this approach. The initiation of the Human Genome Project shortly thereafter endorsed this concept as one of the critical elements of advancement in the life sciences. The tools thus
created have increasingly empowered us to bypass the biochemical inter-
mediaries and to probe the ultimate substrate of biological variance: the
genome. In this way, we become increasingly able to understand the
differences between healthy and diseased individuals, and also the critical
differences among individuals who initially appear to have the same disease.
Association at this level becomes, in effect, synonymous with causation.

Will the advances in genetics help us find better ways of detecting,
treating, and preventing human disease? Will we be able to individualize the
prescribing of medicines in ways that take advantage of genetic variation?
We are early in the process of applying many of these new concepts and
tools, and much of the data generated so far remains anecdotal, in want of
replication and validation. Nevertheless, there exist major opportunities
and applications for molecular genetics in the process of drug discovery and
development, opportunities that offer a vision of more effective therapies
for individuals suffering from our most common diseases. This essay will
examine the characteristics of those diseases from a genetic perspective, and
describe the short- and longer-term importance of pharmacogenetic
strategies in developing improved, individualized treatments.

Mendelian versus complex, common disease

Molecular genetic techniques and approaches have proven most effective
in the area of the so-called ‘classical’ inherited diseases. These are illnesses
in which alterations in one gene are of overriding pathogenic importance,
and which are transmitted in a simple fashion that follows quite evidently
Mendel’s laws of inheritance, such as Huntington’s chorea, sickle cell
anemia, and cystic fibrosis, to name but a few. Thanks to advances in
methodology and technology achieved over the past two decades, the
concepts of ‘reverse genetics’ and ‘positional cloning’ have evolved from
pure theory to become standard genetic investigative techniques.
Powerful, unbiased ‘genome-screening’ techniques (which cover the
genome with narrowly-spaced, but otherwise randomly assembled
collections of markers) were applied to data-bases drawn from
genealogically and clinically well-characterized pedigrees of affected
families allowing the identification of disease-causing mutations in single
genomes. Early successes in this approach led to the discovery of the genes
underlying the cause of some of the most common single gene disorders,
such as cystic fibrosis and muscular dystrophy. These successes
dramatically underscored the remarkable power of molecular genetics to
find causative disease principles, and established the fact that a quantum
leap in biomedicine had occurred.

However, in as much as these ‘single gene diseases’ are devastating for
the affected patients and their families, their overall impact on public
Impact of genomics on healthcare

Table 1 Heritability estimates for some common complex diseases

<table>
<thead>
<tr>
<th></th>
<th>Heritability</th>
<th>MZ twin concordance</th>
<th>DZ twin concordance</th>
</tr>
</thead>
<tbody>
<tr>
<td>High blood pressure</td>
<td>0.60</td>
<td>0.6–0.8</td>
<td>0.3–0.5</td>
</tr>
<tr>
<td>Asthma</td>
<td>0.72–0.8</td>
<td>0.12–0.89</td>
<td>0.03–0.05</td>
</tr>
<tr>
<td>Type 1 diabetes</td>
<td>0.72</td>
<td>0.35</td>
<td>0.37</td>
</tr>
<tr>
<td>Type 2 diabetes</td>
<td>0.26</td>
<td>0.50</td>
<td>0.43</td>
</tr>
<tr>
<td>Type 2 diabetes + IGT</td>
<td>0.61</td>
<td>0.63</td>
<td>0.04</td>
</tr>
<tr>
<td>Rheumatoid arthritis</td>
<td>0.32</td>
<td>0.15</td>
<td>0.04</td>
</tr>
</tbody>
</table>

health is in fact extremely small. They pale in importance compared to more common diseases, such as ischemic cardiovascular disease, non-insulin dependent diabetes, osteoporosis, cancer, etc., that affect large segments of the population, and that exert the major toll of morbidity, individual suffering, and public health care spending.

We recognize today that these so-called ‘common complex diseases’ all have both external (environmental) and internal (genetic) contributors. The external contributors are amply documented in the annals of a century of epidemiological investigation: toxic exposures, exercise patterns, dietary habits have all been shown to modify disease risk. The internal contributors or genetic factors have historically been recognized primarily based on the observations of familial clustering of disease and aggregation of cases, in the absence of classic segregation patterns seen in single gene diseases. This clustering became apparent for many of these diseases as a result of epidemiological surveys, and then from more directed twin- and adoption studies (see Table 1). For many of these illnesses the establishment of a ‘positive family history’ has long served as an important parameter for individual disease risk estimation. Thus, for example, stroke occurs twice as commonly in men with a maternal history of stroke than in those without such a history; and hypertension is significantly more commonly shared among natural than among adoptive siblings and parent-child pairs. In contrast to the monogenic disorders mentioned above, these diseases are polygenic: the inheritance patterns are complex, involving the transmission of combinations of disease-susceptibility genes, each of which may contribute in varying degrees to the disease phenotype.

Today we appreciate that the manifestation of a common ‘complex’ disease is the result not simply of an additive accumulation, but of a more complex interplay between several environmental (e.g. diet, environmental agent exposure, behavior) and several genetic risk factors that may be categorical (e.g. permissive) or quantitative (e.g. additive, multiplicative) in nature. Thus, any one of these risk factors alone is commonly not sufficient to lead to the manifestation of disease at a given time during an
individual's life. This concept helps to explain the puzzling individual examples of heavy smokers who escape lung disease or alcoholics who remain free of cirrhosis. Genetic factors decrease or increase the risk of experiencing a disease, and thus act as either protective or predisposing. The cumulative effect of many counterbalancing or enhancing influences, both environmental and genetic, ultimately determines overall risk. This will be discussed in more detail below.

The complex pattern of inheritance that these diseases display, *i.e.* one that does not follow the simple rules of an autosomal dominant or recessive transmission, but rather presents itself as accumulation or clustering of disease in families, is at times referred to as non-Mendelian inheritance (a term that is probably more appropriately reserved for epigenetic inheritance). Paradoxically, however, the apparently unpredictable pattern of inheritance in common complex diseases is directly predicted by, and in perfect agreement with, Mendel's second law, describing the independent assortment of alleles; thus, a large number of contributing alleles, each and every one of which follow precisely Mendelian inheritance rules, but assorting randomly, create the familial complex transmission patterns observed. Theoretically, given full knowledge of the effects of all contributing gene variants on a particular phenotype, characterization of an individual's genotype profile should allow a reasonably accurate prediction of his/her disease risk that approaches the reliability possible in monogenic disorders.

However, given the anticipated multiplicity and complexity of genetic factors that, in concert with a similarly complex set of environmental variables, contribute to any common complex disease, it is clear that we will, even under the best of circumstances, be left with a considerable variance around any point-estimate of risk that we arrive at using even the most sophisticated approaches. This is primarily the consequence of the fundamental paradox we are encountering by recognizing disease heterogeneity, and trying to account for it by subdividing larger conventional diagnoses into smaller and smaller subgroups of patients: at the extreme end, this reduces sample size to unity – very much in accordance with the recognized uniqueness of each human being, but with complete lack of the kind of statistical data we traditionally use to test the likelihood that an observation is (biologically) true. Thus, meaningful data will only be obtainable on a limited number of those gene variants that are relatively common and that contribute at least modestly to the phenotype, leaving considerable "noise" emanating from more rare gene variants, and from those that individually contribute only minimally to the trait (although in sum they may significantly modulate its expression).

An extremely important conclusion follows from these considerations: in the context of complex diseases a 'genetic risk factor' never takes on the kind of deterministic ('Mendelian') quality that the presence of a disease-
associated mutation represents for risk assessment in ‘classic’ monogenic diseases. Monogenic gene mutations can, in first approximation, truly be viewed as disease-causing (although their morbid manifestations are, as is well appreciated, not absolute, with considerable variance often observed among carriers of the same mutant gene within one family — and even more pronounced in different families — with regard to age of onset, severity of symptoms, and clinical course). In contrast, this will never be the case in complex disease. The very fact that we speak of ‘disease-susceptibility’ and ‘disease-predisposing’ rather than disease-causing genes bears witness to the far greater modulatory role that un-measured, un-measurable, or un-interpretable (due to missing genetic epidemiological data) factors play in these disorders, as discussed above. Thus, in complex disease, even the most comprehensive knowledge of relevant genetic and environmental factors affecting the likelihood of a disease will not fundamentally change the nature of diagnostic message — one of relative risk — delivered: thus, the much-feared ‘stigmatizing’ aspects of a single gene disorder diagnosis will not be encountered for the genetic variants contributing to a common complex disorder. Any one genetic factor will only shift the overall risk, but will never account for all of it, and even the integrative consideration of several genetic factors will not achieve this goal. It should be reassuring to patients that genetic testing in common complex disease will, therefore, not be ‘stigmatizing’, but will, in essence, deliver information that is qualitatively very similar to disease-risk assessments made today based on a patient’s environmental and behavioral profiles; information that, rightly or wrongly, does not encounter major sensitivity in the community. Thus, every smoker understands today that he or she faces a decidedly increased risk for a range of pulmonary diseases, yet, in the community, smoking is not perceived as a sign of impending doom or demise (of note, the relative risk attributable to smoking as a single predisposing factor for these illnesses is almost certainly manifoldly higher than that of even the most important common susceptibility-gene variant!). These facts, however, are not widely appreciated in either the lay or the medical community, and education about these issues is sorely needed.

Since all common complex diseases are of delayed onset, consideration of the time (or age) axis must also be regarded as an essential (permissive or contributing) domain in the consideration of disease causation, and it is indeed likely that risk-lowering interventions will first and foremost delay, but not necessarily altogether avoid, manifestation of an illness.
To understand how these individual susceptibilities to disease can be characterized, a brief discussion of single nucleotide polymorphisms (SNPs) – which are anticipated to be the overwhelming source of such susceptibility-conferring gene variants – will be helpful. Scattered throughout the genome of all species are single nucleotide changes that are the very substrate of human diversity. They occur in humans every 300–2000 base pairs along the genome; in principle, they may occur at any nucleotide, although for genetic epidemiological purposes those that are relatively common will be of greatest interest. The vast majority of these SNPs are functionally silent, occurring in non-coding or non-regulatory regions of the genome. However, some fraction of these SNPs does result in altered protein structure or expression. These (potentially) biologically functional SNPs are considered the essence and substrate of human diversity in both health and disease. Once these common, biologically (measurably) relevant SNPs have been identified, and, in particular after those that contribute materially to disease risk have been characterized, a SNP-based ‘genetic profile’, which may be viewed as an individual’s ‘fingerprint’ of relative (genetic contribution to) risk for various illnesses will become a feasible proposition. As such, SNPs are of course also to be expected to hold the clue to individual differences in drug efficacy and are a logical place to look for the explanations that we seek.

How will the use of genetic investigation and tools advance our approach towards these common complex diseases, in particular with regard to drug discovery and development? It is likely to have profound impact and to occur in a number of different areas — and on a number of different levels and timescales. Genetic investigation will affect all stages of the process, from target discovery, target validation and compound selection to clinical development and marketing of a therapeutic agent. We predict that the incorporation of genetic tools and considerations into this process will be one of the pre-eminent forces that will advance the practice of medicine towards a more individualized approach to health care. Ultimately, this will be done through the integration of the processes of diagnosis and treatment: after determining an individual’s personal disease profile, a therapy tailored to (the most) relevant individual characteristics can be recommended. The result will be more effective treatments and less dependence on trial and error in prescribing.

Methodological challenges in the understanding of common complex disease

If we contend that understanding the cause of disease should render drug targets that are superior compared with the more purely associative
ones we have pursued so far, then we should certainly expect the use of
genetic approaches to have a major impact on this level. The obstacles,
however, are inherent in the nature of complex, common disease itself.
Although the molecular genetics tool-kit for gene discovery has become
a commodity, its application to the dissection of these disorders is
fraught with major difficulty. The common complex diseases share four
characteristic features, each of which contributes its share of
methodological challenge. These characteristics are: (i) multifactorial
and ecogenetic nature; (ii) polygenic inheritance; (iii) genetic
heterogeneity; and (iv) continuous (as opposed to dichotomous) traits.

The multifactorial nature, with external (environmental and
behavioral/life-style related) factors being as important, or more so, than
the aggregate of the genetic factors, makes careful accounting, or
controlling for these variables essential, while at the same time diluting the
magnitude of phenotypic variance ascribable to genetics. As described
above, the relationship of genetic and environmental factors may be not a
simple, additive one, but a complex (ecogenetic) interaction, requiring
careful statistical assessment of interaction terms. For example, while
excess sodium consumption will modestly elevate almost everyone’s blood
pressure as a basic, physiological consequence of fluid retention, it can do
so much more dramatically in the presence of gene variants that render an
individual ‘sodium sensitive’. Of course, depending on the relative
contribution of nature (genetics) versus nurture (environment) — which
varies along a sliding scale from disease to disease — discovery of the
genetic component(s) will be more or less difficult (as well as less or more
meaningful for disease management); thus, it maybe anticipated that the
genetic underpinnings of insulin-dependent diabetes will be more readily
accomplished than those of non-insulin dependent diabetes.

Likewise, the polygenic nature of common complex diseases means
that disease causation attributable to genetic contributors represents the
aggregate effect of several genes, making the effect for most single
contributing genes rather modest and difficult to detect. In addition,
consideration must be given to the interplay of genetic factors that may
be more complex (epistasis) than is captured by linear modeling.

Genetic heterogeneity means that multiple genes in different combina-
tions may contribute to an apparently identical clinical presentation. For
example, there might be a dozen gene variants that contribute to
osteoporosis. While one patient may carry the first eight, but not the
remaining four, the next patient may carry the last six, thus sharing only
2 genetic risk factors with the first one. Clearly, this genetic heterogeneity
of what appears to be a single disease entity, adds statistical artifact
commensurate with the degree to which the set of genetic factors involved
in disease causation differs from one person or one family to the next.
Lastly, it is important to bear in mind that common complex diseases are characterized by quantitative, continuous traits that do not offer the luxury of simple dichotomous categorization. The reason for this becomes clear when we consider the additive (or more complex interactive) influences — either in the direction of disease predisposition or prevention — that a set of genes and their respective variants may have on a clinical phenotype. Let us assume that there is a dozen gene variants that tend to raise blood pressure (BP) and another dozen that tend to lower BP. Person A might have variants that increase the activity of 8 BP-raising genes and only 3 genes that lower BP, resulting in elevated blood pressure compared with the population average. Person B might have 5 BP raising genes and 5 BP lowering genes, and map to the apex of the Gaussian curve. And person C might carry 2 BP-raising and 6 BP-lowering gene variants and be assigned to the lower tail of the distribution. Using this example, it becomes already clear that there are, strictly speaking, no self-evident criteria for any of the common complex diseases to declare ‘normal’ or ‘abnormal’, ‘diseased’ or ‘healthy’ status, and when we use these categorizations (as we pragmatically must) in the practice of medicine, they are based on more or less arbitrary definitions and cut-off values in comparison with large samples. By binning quantitative data into categorical, a great deal of information is lost, and much potential error introduced. For example, a measurement error of 2 mmHg on a poorly calibrated sphygmomanometer that reads 141 mmHg instead of 139 mmHg can categorize a patient falsely as hypertensive and thus radically change his or her clinical status from control to case, when in a quantitative analysis the effect of this error would have been hardly noticeable. Limiting case or control status, if one must carry out a study based on dichotomous variables, to more extreme tails of the distribution is somewhat helpful, but presupposes the availability of originally quantitative data and also ignores a large part of the spectrum of clinical variance. Similar examples as drawn here from blood pressure/hypertension abound, e.g. for serum glucose or cholesterol, for atherosclerotic changes in vessels, or for neurofibrillary tangles in aging brains. In evaluating the importance of these conditions, painstakingly precise measurements are essential to capturing the relevant phenotypical information, and analyses based on sloppy measurements are obviously worse than useless.

Considered together, these four considerations help explain why these disease entities are called ‘complex’. To the biometrician, the elements enumerated above amount to a statistical nightmare, where modest signals are pitched against a multitude of sources of noise. How might one tackle this problem, in order to identify new potential targets for drug development?
Long-term impact: new target discovery

One way of approaching this dilemma is to use reductionist approaches, tried and tested by experimentalists in all branches of science to reduce noise, at the potential cost of failing to capture the truly relevant components of a phenomenon. The use of inbred animal models of disease represents such an approach — it allows an investigator to reduce both environmental and genetic heterogeneity, the former by tightly controlling the environment, the latter by working with inbred disease strains in which all individual animals share exactly the same set of disease-contributing genes. Most importantly, of course, programmed breeding protocols can be implemented, allowing the creation of very large and powerful ‘pedigrees’ that are impossible to find in any human population. The price of this methodological reductionism is the uncertainty of whether any extrapolation from the experimental to the human condition is justified. Still, novel disease mechanisms and pathways thus identified may provide valid new targets for drug discovery. Among human populations, a similar scenario of reduced complexity can be found among so-called founder populations, or genetic isolates, where an expanding population can be traced back to a more or less recent, small original ancestry, thus limiting the degree of genetic diversity (i.e. heterogeneity for a given disease or trait) present in the sample. Recent efforts to find asthma-related genes among the settlers of Tristan da Cunha, a remote island in the South Atlantic, as well as long-standing genetic investigations in the Finnish-Karelian, the Icelandic, and the Amish populations, are prominent examples of this approach.

An alternative approach towards dealing with very complex data sets lies, conceptually, in collecting massive amounts of data that provide sufficient statistical power to cancel out noise and to reveal the signal. Large genome screens carried out in extensive collections of nuclear families in the search for common complex disease genes represent an early, and still relatively modest, example of such an approach. Presently, available off-the-shelf technology already allows significant scale-up of such approaches. Limiting factors to this approach are processing cost per sample and, much more significantly, the difficulty of identifying and ascertaining sufficiently large proband cohorts. Such ‘brute force’ approaches will also require enormous capacity for data handling and analysis. Novel biomathematical approaches that are capable of handling very large and complex datasets, such as clustering or principal component analytical methods, will play a major role in our ability to interpret the data generated and transform it into meaningful information. This is particularly true for the envisioned application of very large sets (10s to 100s of thousands) of random genetic markers (rSNPs) in large case-control cohorts for genome-wide association...
Genetics in drug discovery and development

(linkage disequilibrium) studies; for this to become reality, however, SNP processing technology will need to advance by orders of magnitude upwards (to something like $10^6$-$10^8$ genotypes per day), paralleled by an equally significant decrease in cost, to fractions of a cent per genotype.

While intelligent arguments can be made in support of the primacy of 'causal' targets over those that are only associative or simply deduced from function, it should be also be emphasized that this does not by necessity imply that only individuals in whom a particular causative disease mechanism is operative or predominant will respond to a medical intervention aimed at that mechanism. Rather, by finding a gene (and, in due course, the associated additional elements of the biological pathway or signaling cascade) that, if dysregulated, has the potential of inducing the morbid phenotype of interest (at least in a subgroup of patients with the disease), one generates important new functional and mechanistic-pathogenetic knowledge about the disease. Based on our current experience, there is a very good likelihood that a drug targeting such a mechanism — if it normally participates in the regulation or homeostasis of a disease-relevant phenotype or trait — will be at least partially effective also among individuals in whom a different pathomechanism underlies the same disease, thus broadening the pharmacological arsenal available to any patient suffering from this disorder.

The use of causative targets will not necessarily accelerate, in an individual case, the time needed for discovery and development of a new drug, since most of the steps involved in proceeding from target to marketed medicine remain unchanged. Indeed, we may anticipate that a gene or its encoded protein, even if identified as 'disease-contributing', may not necessarily be a tractable drug target at all, and that other elements or members of the pathway will first need to be elucidated. However, it is safe to assume that the drug discovery process, as a whole, will become more efficient as putative drug targets are identified by genetic approaches. These targets and/or associated pathways come with the intrinsic validation of being disease-relevant. They may provide enhanced knowledge about the structural effects of genetic polymorphisms and thus can guide compound selection. For these reasons, the rate of failures and 'killed' projects is likely to decrease as more such targets are used. Overall, therefore, the drug discovery process will become more expeditious, more efficient, and thus more economical.

**Mid-term impact: target validation and compound selection**

On a more intermediate-range time scale, the deployment of molecular genetic approaches has the potential of significantly contributing to the
drug discovery process by improving the likelihood of selecting compounds that will ultimately prove successful and by providing guidance for the clinical development process. In addition, there is certainly value in validating the target as disease-relevant, by showing an association between a gene variant and disease prevalence or incidence, even at the relatively late stage of clinical development.

An important mid-term goal is clearly the screening of active and contemplated drug targets for the presence of genetic variants. If genetic variants are found, testing for potential impact of such variants on disease, disease susceptibility, and on chemical tractability will be the logical next steps. Screening the targets for SNPs is thus widely considered a useful, perhaps essential strategy. Polymorphisms found in the target, particularly if they affect the translated amino acid sequence, should be assessed with regard to their possible impact on target structural confirmation, and its possible effects on target-ligand binding/interaction. In addition, even if they are a priori 'silent', they should be followed-up with genetic epidemiology studies to compare the prevalence of the respective variants in cases (patients with the disease in question) and controls (since they may be in linkage disequilibrium with non-captured variants affecting gene regulation, or they may directly affect mRNA stability and or translational efficacy).

If such an association between a molecular variant of the target and the disease is indeed found, at a sufficient, predetermined level of statistical significance and/or magnitude of effect, this information may have important consequences both with regard to the chemical screening strategy as well as to the design of clinical studies. Presumably, a compound that selectively targets the disease-associated molecular variant of the target could be designed and tested. The above-described strategy would certainly validate the drug target as a disease relevant one, and presumably would increase the likelihood of successful completion of the development efforts. Thus, a small molecule with high, selective affinity (if the mutation affects binding) to the target variant of interest, but not to the variant found in non-affected probands, might be pursued, and a clinical development strategy may be considered that preferentially or exclusively enrolls carriers of the disease-associated target.

In contrast, if no association between the target's genetic variant(s) and disease prevalence is found, and assuming that there is at least a possibility that the mutation affects drug binding, it might be wise to conduct secondary screens with assays employing the molecular variant(s) of the target to ensure equally optimized pharmacodynamic characteristics for all variants and to avoid unwelcome surprises later on during clinical development in the form of pharmacogenetic phenomena that increase response variance, thereby decreasing study statistical
power, and in turn increasing development cost, and potentially fragmenting the future market of the drug.

Carried out in a more comprehensive fashion, this ‘frontloading’ of target polymorphisms may also include other members of the targeted pathway, as well as other ‘candidate’ genes considered to be potentially interacting with the mechanism of action of the drug, or its metabolism.

**Short-term impact: pharmacogenetics**

‘Idiosyncratic’ individual differences in drug response have long been recognized as a major limitation of successful and safe medical treatment. In the absence of identifiable factors that would allow the prediction of a specific response these ‘differences’ have been a source of major frustration for patients, for physicians, for third party payers, and for the pharmaceutical industry. Given the, so-far, only possible trial-and-error approach to finding the appropriate drug exposes the patient to delays in receiving appropriate treatment for his or her condition, increases the workload on the physician, and represents wastefully spent resources on ineffective treatments for the health care payer. The industry sees large investments wasted because of adverse events that cannot rationally be avoided, and because the dilutional effect of non-responders may prevent a clinical trial from reaching the statistical significance needed for regulatory approval. The field of pharmacogenetics seeks to understand the genetically encoded inter-individual differences among patients that are at least one important source of these differential responses. A better understanding of the nature of these differences, and the ability to test for them, are, therefore, increasingly being viewed as not only sensible, but as a prerequisite for selecting the appropriate drug for the patient on an individual level.

Pharmacogenetic phenomena can conveniently be separated into two principal groups, those that affect the metabolism of a compound (pharmacokinetics) and those that affect the activity, and thus the efficacy of a medicine (pharmacodynamics). The former have of course been recognized for many years, based on biochemical observations and assays, and are considered routinely during the development of a drug. Thus, phase I trials are commonly carried out in appropriate, well-defined cohorts with representation of specific metabolic phenotypes, e.g. fast and slow acetylators. The recognition of a substantial number of variants of enzymes that play important roles in drug metabolism based on DNA polymorphisms opens the way to a more comprehensive, more precise, and more sensitive examination of these interactions than has been possible on the level of protein biochemistry. Understanding
the interaction between certain metabolizing enzyme variants provides not only the opportunity to avoid adverse events, but also to recapture efficacy in relevant patients by adjusting dosage and/or choosing optimized formulations.

The recognition of genetic variants’ role in affecting drug response or efficacy is of more recent date, and while conceptually attractive and clearly to be expected, remains at this point mainly anecdotal and in need of replication and validation. For example, the acetyl cholinesterase inhibitor, tacrine, has been reported to show differential efficacy dependent on ApoE genotype (as well as a superimposed sexual dimorphism) in patients with Alzheimer’s disease, and pravastatin has been reported to provide differential benefit to patients with coronary heart disease stratified by cholesteryl-ester transfer protein (CETP) genotype. Given the complexity of the disorders for which such examples have been reported, the clear-cut, categorical results shown seem surprising, and additional data will certainly be required. However, it is of some interest to speculate that, in the presence of very powerful environmental stressors, which are orders of magnitude stronger than those encountered in the natural environment, the effect of certain genetic factors may be greatly enhanced, and thus become much better discernible. Viewed from this angle, pharmacological agents could certainly be seen as such ‘non-physiological stressors’. If this concept were indeed to emerge as a common principle, then we may ultimately be confronted with significant pharmacogenetic effects far more often than would be expected based on differential, modest contribution of individual gene variants to complex disease entities.

Many pharmaceutical companies are today seriously considering, or already engaged in, the collection of a DNA sample from each of their patients participating in phase I, II, or III trials, to retain the option of investigating potential associations between genetic variants and positive (efficacy), negative (lack thereof), or adverse clinical responses. Currently, these approaches are restricted to a limited number of likely or suspected ‘candidate genes’ (generally, these are disease-related and/or drug-specific). The ongoing efforts to develop a comprehensive, densely spaced, genome-wide SNP map may allow us in the future to conduct screens for pharmacogenetically active genes as whole-genome, unbiased searches.

What are the consequences of discovering a pharmacogenetic interaction? On a conceptual level, we may be able to avoid exposing side-effect-prone individuals to the drug, or, in the case of a pharmacokinetic phenomenon, to adjust the dose accordingly. Regulatory authorities are beginning to recognize this and have issued early language encouraging pharmaceutical companies to develop rational risk-stratification approaches based on pharmacogenetic
testing. This information may allow them to receive regulatory approval that would otherwise not be granted. Likewise, we may avoid prescribing a drug to patients who are unlikely to respond, restricting it to a genetically defined group of responders. Defining such a subpopulation of responders might thus allow the ‘rescue’ of a drug that would – in a non-stratified population – not cross the statistical threshold of efficacy required for regulatory approval.

Commonly, concerns are raised that such a scenario may not be economically viable: by ‘fragmenting’ indications into smaller (sub-)segments, the critical mass of revenue that justifies and enables an expensive drug development program may not be reached. Enrolling patients stratified by genotype will slow down recruitment, and the need to provide a diagnostic along with the pharmaceutical will further complicate the logistics of marketing and distribution. However, it is extremely difficult to sustain the argument that a responsible company should market products to individuals for whom they will be harmful or ineffective, if the means exists to predict these results. Furthermore, on a practical level, the effort to produce medicines with superior therapeutic efficacy and fewer side effects will likely be repaid by improved patient adherence to therapy (compliance), by higher market segment penetration, and by longer lasting treatment. Moreover, if the genotype-specific response is one tied to genotype-specific disease etiology, then one may speculate that predisposition testing for this genotype may extend the size of the target audience to include less symptomatic, earlier cases, and perhaps even asymptomatic individuals (prevention). It stands to reason that a substantially more effective drug would also command a premium price; under those circumstances, the companion diagnostic could, of course, also generate additional revenues. In addition, if a specific responder subgroup is known and targeted, then restricting enrollment to this group should allow significant savings by allowing much smaller trials to be carried out, provided that safety requirements are met. Consideration of pharmacogenetics and the performance of appropriate studies is likely to become very quickly as much standard part of a drug development process as detailed pharmacokinetic and drug–drug interaction evaluations have become by now; and it is not unlikely that regulatory authorities will one day require such data to be submitted. Indeed, in as much as it should be in every company’s interest to identify important pharmacogenetics phenomena as early as possible to retain the opportunity to adjust regulatory submission and marketing strategies in time, there is utility even in the two extreme outcomes of this policy: if sufficiently complex pharmacogenetics effects are noted to make the marketing of a drug not feasible, killing the project at the earliest possible timepoint is the only logical and money-saving consequence; and demonstrating complete absence of any measurable
pharmacogenetics effects will serve as a strong argument for maximally broad marketing of the drug.

The current emphasis on cost-effectiveness for new therapies will dictate that every effort be made to produce medicines which are as useful as possible for their targeted recipients. Thus, pharmacogenetics will find itself positioned at the very core of the development of new medicines, and companies that fail to recognize this will have to learn the hard way.

Towards individualized, integrated health care: medical, ethical, economic implications

In the broadest of terms, the consequences of using genetically driven approaches towards target discovery and drug development will result in an increasingly ‘individualized’ approach towards health care, logically so because genetics is, after all, the door through which we gain access to inter-individual differences. As we are learning to consider these differences as important elements in the approach to the patient, and in our ability to deliver effective and successful health care, we will invariably arrive at a more and more ‘custom-tailored’ practice of medicine. Conceptually, this is nothing new: ever since the days when blood-letting was practiced as near-universal therapy, an increasingly differentiated and sophisticated approach towards classifying separate clinical entities, i.e. differential diagnosis, has been the driving force behind all progress of medicine. The incorporation of genetic data simply adds yet another level of sensitivity and specificity to the diagnostic armamentarium that is in everyday use by the practicing physician today, albeit with certain special features. Inherent in this development is an increasingly important role of diagnostic tests in the practice of medicine: as therapy tends to target more and more specific subsets of disease entities, and of the overall patient population, increasingly precise diagnostic tests that help discern these subcategories are the obvious prerequisite. If the prescription of a drug requires knowledge of a certain biochemical or genetic marker, then only a fully integrated ‘package deal’ will be viable in the market, providing the diagnostic, the therapeutic, and —importantly — the specialized information needed for patients, physicians, and other health care personnel to be able to use and implement such a sophisticated, highly personalized approach.

Genetic information remains a breed apart from all other kinds of medical data, with regard to the trepidation with which the general public views it. This reserve, which has its roots in the historical
association of eugenic movements with the academic discipline of genetics, as well as in the fear that genetic data will be misused by insurance companies, employers, etc. to discriminate against individuals ‘branded’ as being at risk for certain diseases, represents a major stumbling block for both genetic research and the use of genetic data in the clinical setting. Requirements for confidentiality and limitation of using the data, even in the research setting, only in the precise context specified by the probands, place major constraints on our ability to generate meaningful data. While genetic data differ from other health care data in their prospective and prognostic information content, and in the potential implications their knowledge or disclosure may have for blood relatives of the tested probands, these characteristics have real meaning only in the case of ‘classical’ single gene disorders. In contrast, the role of genetic factors in common complex disease, as we have seen, is a much less discrete one, providing, therefore, never more than probability estimates of associated disease risk. Insurers have, for many decades, been asking applicants about their family medical history, and have thus had access to an integrated parameter of heritable risk whose reliability far exceeds that of any of the presently available genetic tests in complex disease. This practice never led to an outcry of protest. Clearly the need for education, information, and teaching about genetics is large and evident, to put the risks in perspective and illuminate the potential benefits of our growing knowledge.

An aspect that currently occupies many genetic epidemiologists’ thinking should not go unmentioned: great emphasis is frequently placed on the importance of measuring SNP allelic frequencies among different ‘ethnic groups’ and ‘races’. Although there is clear evidence for differential allele frequencies among individuals from different ethnic backgrounds (a consideration that used to be considered important in forensic testing until sufficiently complex panels of markers had been designed), a focus on these differences harbors the great danger of re-defining race on the basis of genetics, and not, as we have come to appreciate, on the sociocultural and socio-economic level (in as much as the increasing intermixing of populations from different backgrounds should make a biological definition of race increasingly meaningless, anyway). Indeed, genetics may be viewed as a unique opportunity to move (certainly in the field of biomedicine) away from ethnic/racial stereotypes: understanding the genetic variant underlying sickle cell disease has allowed us to use that parameter, rather than skin color, as a (much more accurate) predictor of disease risk; along the same lines, antihypertensive treatment targeting sodium-sensitive disease will be applied to carriers of the relevant gene variant, once found, based on a genetic diagnosis irrespective of ‘race’, and no longer, as is current practice – for lack of molecular-genetic definition of the trait – in a blanket, first-line approach to hypertensive
African-Americans. Thus, there is great cause for optimism that the appreciation of the much more profound genetic diversity that is bound to be present at the inter-individual, as compared to the inter-ethnic level will relegate the use of ethnic stereotyping to the closet. This will, however, require a very detailed knowledge of disease contributing genes; the reason that at least in gene discovery and pharmacogenetics trials ethnic stratification is currently still to some extent defendable relates to our inability of accounting on a genotype level for 'genetic background' and/or 'modifier genes', with ethnicity (if representative of a relatively old and homogeneous ethnic group) serving as a integrative parameter for such differences. The perceived need for characterizing random mapping (in contrast to biologically relevant) SNPs for allele frequencies among different ethnicities is more difficult to understand, as we already know that only about 20% of such SNPs show significant differences across ethnic groups; given the large number of SNPs currently being generated, these could simply be ignored.

Although questions surrounding data confidentiality and protection of patient privacy tend to dominate current discussions about ethical and societal aspects of genetic testing, it is important to consider that, in order to derive benefits from genetic information, complete confidentiality cannot and must not be the goal. Prescription of any drug that shows important pharmacogenetic interactions requiring genetic testing will immediately disclose, by inference, the recipient's genotype to the pharmacist and the insurer. The presumably confidential diagnosis of diabetes mellitus is revealed in a similar way in current practice by the filling of a prescription for insulin. Arguably (or maybe not even?), it would make much more sense to disclose the genetic risk/predisposition, once genetic profiling is possible, to the pharmacist and insurer by submitting a prescription for a drug that will prevent the onset of the disease. Rather than treating genetic information as a special form of medical information, standards of confidentiality and individual choice regarding the use of all types of medical data need to be established and/or redefined. Ultimately, society will benefit most by simultaneously acting to protect not so much the privacy of medical information, but the way it is used, thus paving the way for patients to derive the benefits of the progress made in pharmacogenetics and other areas of medicine. This will be best accomplished by enacting appropriate laws that regulate the use of genetic, as well as of all other potentially sensitive medical information, and to make any use of this information outside the terms thus established unlawful and prosecutable. Lawmakers and ethicists in many countries are recognizing the need for such a legal framework and a multitude of efforts are currently underway to draft such legislation; coordinating these activities on an international level would be of great importance to
avoid a Babylon of non-aligned rules that will make the use of genetic and other medical data both for research and for medical practice very difficult.

Whereas the prospective or predictive dimension of genetic testing brings a number of ethical and legal considerations more sharply into focus than other diagnostic modalities, it also offers the opportunity to shift medical intervention from therapy to prevention, and from diagnosis to prediction. As we develop a more comprehensive understanding of the contribution that genetic risk factors confer to the overall risk of experiencing a wide variety of health problems, we will be able to offer a custom-tailed program geared at a continuity of health maintenance. Thus, at the outset, predisposition screening would profile an individual’s genetic risk for a range of diseases. In many instances, preventive counselling will already be possible and indicated at this stage, with an emphasis on risk reduction by environmental and behavioral modification. In a follow-up phase, the subject could be monitored pragmatically in a targeted, focused fashion with appropriately graded scrutiny — directly proportional to the determined risk for disease, avoiding the kind of much more dogmatic approach used today that, in spite of all its benefits, is fraught with inefficacy. This strategy will result in a more cost-effective health maintenance program with a better chance of identifying disease at earlier clinical stages where outcomes are likely to be better. Thus, viewed in this fashion, the use of genetic testing and information is supported and justified on the level of both macro- and micro-ethics.

A repeated caveat is necessary lest the reader incorrectly believes that genetic information yields the perfect ability to predict disease. For complex disease, unlike for single gene disorders, an understanding of genetic contributors to disease-susceptibility enables us to be better prognosticators. However, disease risk still will be expressed in terms of statistical probabilities, meaning that small numbers of patients designated at low risk will experience disease while some patients will still fail to be identified. However, the added forecasting precision that is achieved by effectively shifting the basis for preventive strategies from the large-scale epidemiological (population) to the much smaller-scale genetic (individual) level will, if applied comprehensively and consistently, provide a medically and economically brighter future of health care for all participants.

Outlook

A bold picture of emerging paradigm shifts in medicine, pharmacology, and the role of the pharmaceutical industry has been presented. How likely will it become a reality within our lifetime? Let us not forget that
while the technical basis for these developments is in place today, and while conceptually the changes outlined appear logical, two major hurdles will need to be overcome: the first is the creation of the knowledge base that is required to carry out profiling for genetic risk — an enormous and dauntingly difficult task — and the second relates to acceptance of these new approaches among the general public and the patients. Both issues are indeed likely to be closely linked: whereas an increasingly powerful database will provide clear evidence of conferring benefits sooner, this will support acceptance of the idea of genetic testing among patients. The experience with patient advocacy groups for single gene disorders shows that efforts to find the causative gene usually find strong support, being recognized as the first, essential step towards treatment, cure, or prevention. We may ultimately witness patients emerging as the driving force behind the development of this vision of integrated and individualized medicine, in an almost radical shift from their currently displayed skepticism, since they clearly have so much to gain.

If we are to be successful in making the use of genetic information a valuable and useful tool, we need to embark on this quest as a coordinated and interdisciplinary effort at a large scale. Only a functional collaboration among clinicians, geneticists, epidemiologists and biomathematicians, as well as lawmakers, bioethicists, sociologists, clergy, teachers, and patient advocates will ultimately succeed in creating the knowledge base and popular understanding needed for progress in medicine to take place. While the tasks in the field of science loom large, those in the field of societal dialogue are no smaller challenge. However, the very significant benefits both to patients and to society that individualized medicine will bring make it critically important for all concerned to work together to allow the growing understanding of the molecular basis of disease to help patients around the world.

Acknowledgements

I wish to thank Drs Lyn Singer-Lindpaintner, Jonathan Knowles and Christina Dahlstroem for their contribution in discussing and editing the contents of this manuscript.

References

4. Slutsky AS, Zamel N. Genetics of asthma: the University of Toronto Program. Am J Respir Crit Care Med 1997; 156: S130–2