

THE DIGITAL FUTURE OF MOLECULAR MEDICINE: Rethinking FDA Regulation

Peter Huber
Senior Fellow, Manhattan Institute

Pharmacology is fast becoming an information industry. Biochemists can read every letter of life's core genetic code and determine the composition and structure of all its molecular progeny—the downstream proteins and other biochemicals that shape our health, for better or worse. They have the tools to design a drug that can control almost any molecular target. The power in nature's code and our mirror-image drugs resides in minuscule packets of material that technologies now in hand can read, copy, and manipulate. And these technologies are getting cheaper and improving even faster than their digital siblings.

But biochemists have arrived on the scene billions of years behind nature, which neglected to provide manuals that explain how all the molecular slivers of code that it has created fit together and interact. The search for a new drug is increasingly a search for information about how a molecule of our design will interact with different arrays of molecules that it will encounter in future patients and how those interactions will affect a patient's health. That search accounts for a rapidly rising fraction of the front-end cost and medical value of most drugs. Repeated again and again, with one drug after the next, the information acquired will end up in massive and very valuable databases. The analysis of the data using extremely powerful computers will expose the architectures and dynamics of countless molecular networks that make human bodies function well or badly and that the right drugs can control.

The private sector is already actively engaged in collecting and analyzing the data. Led by a rapidly growing group of companies as diverse as IBM, Myriad Genetics, and 23andMe, the digital community has grasped—far ahead of the FDA and much of the medical community—how fast molecular medicine can now advance by taking full advantage of the recent convergence of astonishingly powerful biochemical and digital technologies. Never before have two such powerful technological revolutions converged to advance a single objective of such universal importance. But unleashing the enormous power and economies of innovation on this last frontier of the information revolution will require fundamental changes in public policy.

The FDA has spent the last 30 years pondering how, if at all, molecular science might be shoehorned into the clinical trial protocols that Washington first used over 70 years ago and formalized in licensing rules developed in the 1960s. The regulatory system is now frozen in the headlights. Its drug-testing protocols cannot handle the torrents of complex data that propel the advance of modern molecular medicine. For all practical purposes, those protocols make it impossible to license most of the drugs and complex treatment regimens that are needed to control the biochemically complex disorders that these data torrents reveal.

Developed at a time when nobody could see or track the molecules that matter, the FDA's current testing protocols rely entirely on empirical studies and statistical correlations. They aim to guard, above all, against just one kind of error in the licensing process: selection bias. But modern pharmacology hinges on the scientific selection of the right drug-patient molecular combinations. The only practical way to work out most of the drug-patient science is to study how the drug actually performs in patients. And the first opportunity to do that systematically is during the drug-licensing trials.

As recommended in a recent report issued by President Obama's Council of Advisors on Science and Technology, the FDA should use its existing accelerated approval rule as a starting point for developing adaptive trial protocols to be used “for all drugs meeting ... an unmet medical need for a serious or life threatening illness” These protocols should promote the meticulous, data-intensive study of the drug's molecular performance during clinical trials. And they should use modern statistical designs to choreograph the adaptive trials needed to ascertain when a drug that provides only some degree of clinical benefit to some subsets of patients can become a useful component of complex molecular medicine.

Part 1 of this paper discusses the rapidly widening chasm that now separates modern pharmacology and the practice of molecular medicine from the drug-patient science developed and certified the Washington way. The chasm reflects obsolete policies and rules put into place to regulate ignorance, not knowledge; it reflects the dearth of molecular medical science, not the science itself or its efficient, orderly development. Part 2 discusses what it will take to unleash the full power of the precision molecular medicine that biochemical science, powered by digital technology, can now deliver.

FOREWORD

The Digital Future of Molecular Medicine: Rethinking FDA Regulation

by Dr. Andrew C. von Eschenbach, Chairman Project FDA

Former commissioner, U.S. Food and Drug Administration; director, National Cancer Institute

During my tenure as commissioner of the U.S. Food and Drug Administration, the federal agency tasked with evaluating the safety and efficacy of medical products that touch the lives of tens of millions of Americans every day, it became increasingly clear to me that a revolution in biomedical science augured the need for significant changes to assure the future success of the agency.

It made less and less sense to evaluate the effectiveness of promising new medicines through traditional clinical trials, in which a cross-section of the intended patient population is randomly selected to receive an intervention—and then compared with a similar population receiving a placebo or the standard of care. Neither population truly reflected the real world of diverse patients who will eventually receive the therapy, if it is approved.

The decoding of the human genome and rapid advances in molecular biology were also making it clear that patients and their diseases that we had once considered homogenous—such as cancer and diabetes—were vastly different by virtue of a constellation of gene or metabolic dysfunctions that modern science could now identify. The historical conundrum of why drugs would work for some, but not for others, could not be explained by these historically “gold standard” clinical trials, but these “responders” could now be prospectively identified by specific “biomarkers.” Traditional trials could not—and were not designed to—take into account rapid advances in our understanding of the mechanistic causes of disease, rather than just clinical symptoms. In short, it is time to rethink what our gold standard should be.

In his new paper, Peter Huber tells the story of this biomedical revolution, and he maps out a path for guiding the agency into a new era of precision medicine that holds unprecedented benefits for patients and the American economy. By embracing new tools and technologies, the FDA can help unleash a new golden age of biomedical innovation.

Huber carefully explains why the FDA can no longer delay change and simply cling to the traditional way of evaluating the safety and efficacy of new medicines. There is talented leadership and staff at the FDA; but—like many large organizations committed to sustaining their core products, mission, and internal culture—it can be overtaken by rapid changes in market structures and technology. IBM, AT&T, and the “Big Three” automakers are only a few examples of once-successful firms that have had to adapt themselves to new technologies and new customer expectations—or risk obsolescence.

Federal agencies are no less immune to disruptive technologies than private firms, and Huber details how rapid advances in molecular biology and quantum leaps in information technology have progressed far beyond the 70-year-old double-blind, placebo-controlled trials that the agency (for the most part) still uses to evaluate new medicines.

Huber builds a powerful, well-argued case for regulators and researchers to “remove their blindfolds” and fully embrace the latest advances in molecular biology, adaptive clinical trial designs (which can shift patients toward more effective treatments as evidence accumulates), and powerful new statistical tools to identify and validate the biomarkers that will allow companies to match promising new drug candidates with the patients who are most likely to benefit from them and least likely to suffer serious adverse effects. Along the way, the FDA and the drug companies will also weed out unpromising or dangerous drugs much more quickly (and less expensively) than they can by using traditional clinical trial designs.

Suffice it to say that the FDA's current protocols are designed to gauge a drug's average effects for both safety and efficacy in clinical trials with "representative" populations. Good drugs are licensed, and bad drugs are relegated to the scrap heap, based on what is essentially a clinical popularity contest. The problem is that human biological diversity is much broader than regulators and researchers had assumed for much of the twentieth century. Matching the right drug to the right patient requires knowing just as much about the biochemistry of the patient as we do about the medicine.

For instance, cancer isn't a single disease; uncontrolled cell growth is driven by hundreds of different defects in cell metabolism and growth that vary widely among patients. Other common diseases, like diabetes, share common clinical effects (such as low blood sugar) but probably have myriad different biochemical causes. Patients are just as likely to vary in their susceptibility to serious side effects. Unless you test the right drug together with the right patients, you are often likely to draw the wrong conclusions about both drug safety and efficacy.

This mismatch between science and regulation has critical implications for patient health. The FDA's one-size-fits-all regulatory pathway has become breathtakingly expensive and time-consuming: it takes well over \$1 billion and a decade to develop a single FDA-approved medicine, according to recent estimates. These enormous sunk costs mean that some diseases will never be cured because it costs too much to develop drugs for them.

Some drugs that might work well in small populations are also abandoned because they work poorly or produce toxic side effects in large, untargeted populations. And the process of developing drugs for complex indications, such as neurological diseases, is so slow and unwieldy that it will take decades for researchers to match the right treatments for the right subgroups of patients.

Thalidomide, a case study that Huber discusses in depth, is the poster child for the tremendous complexity of molecular biology. Prescribed to pregnant women at a key juncture in fetal development, thalidomide produced horrific birth defects—forcing the drug's withdrawal in 1962. But it was returned to the market decades later as evidence accumulated that it could be used effectively to treat leprosy, AIDS, and several types of cancer.

Huber's key argument is that the best time to begin generating information about how a given drug interacts with a given patient's biochemistry is at the "front end," in small, biomarker-driven clinical trials that can then be used to license the drug for very specific uses in targeted patient populations. These studies will be ongoing and iterative and will both inform and be informed by information gleaned from large post-market databases of electronic health records that combine phenotypic and genotypic information.

Companies such as IBM are already operating such databases, offering powerful tools for combating HIV (the EuResist database) and, in the not-too-distant future, cancer. Google and Amazon update their databases thousands of times every day based on precise algorithms that improve their ability to predict who is likely to be looking for what, and when, and why. Similar algorithms and computing platforms can be linked with molecular diagnostics to help researchers, regulators, and companies match new drugs with the molecular profiles of patients who are likely to benefit from them—or, conversely, who should avoid them.

It will not be simple or easy for the FDA to embrace these transformative tools. First, the FDA should take stock of how best to deploy its staff, expertise, and budget to respond to ongoing changes in basic science and product development. The agency will need to reform the clinical trial process and engage additional research partners to help validate new biomarkers, especially by collaborating with other federal agencies, such as the National Institutes of Health (which funds critical research in molecular biology), and with academic medical centers that can bring together the large distributed groups of patients, researchers, and hospitals that will be needed to run new, molecularly guided adaptive clinical trial designs.

To its credit, the FDA knows this and has already taken initial steps to embrace several of these tools. But the FDA's advance in embracing new models of regulation has been glacially slow and largely limited to just a few diseases such as cancer, HIV/AIDS, and some orphan drugs.

For precision medicine to flourish, Congress must explicitly empower the agency to embrace new tools, delegate other authorities to the NIH and/or patient-led organizations, and create a legal framework that protects companies from lawsuits to encourage the intensive data mining that will be required to evaluate medicines effectively in the post-market setting. Last but not least, Congress will also have to create a mechanism for holding the agency accountable for producing the desired outcomes.

The FDA, like any other large, bureaucratic organization, will find it difficult to change and to embrace new models of "doing business" until its "customer" (Congress and society at large) has clearly signaled that the "product" that the agency is delivering is no longer acceptable. Huber has done the agency a tremendous favor by drawing our attention to the need for such change in the agency's clinical trial protocols—and in a way that allows Congress and the agency to chart a clear path toward modernizing the agency's role and functions.

The process of creating a truly precise framework for molecular medicine will be the work of years, not a few months. But, if done as Huber suggests, it can become a self-advancing, self-correcting process that will put the patient at the center of decisions about how and when to deploy or remove new medicines in the battle against complex, life-threatening ailments. Medicine has always aspired to offer patients "personalized" treatments. Huber shows how it can become both personal and precise.

Previous FDA modernizations efforts—including Accelerated Approval, the Orphan Drug Act, and the Prescription Drug User Fee Act—have saved countless lives and helped establish the U.S.-based biopharmaceutical industry as the world's most innovative source of new medicines.

For the first time, we can see how medicine can attack the molecular roots of complex chronic diseases, rather than simply ameliorate them. For the millions of patients at risk of developing devastating ailments such as Alzheimer's, science holds the hope of fuller, more productive lives. For America, it means *trillions* of dollars in lower health-care costs spent treating chronic disease, better-paying jobs in a flourishing life-sciences industry, and a reenergized economy as life sciences transform every sector, from agriculture to energy and even defense.

But for all these technologies to reach fruition, we need the Food and Drug Administration—which has done so much for so long to keep our food supply safe and evaluate new medicines—to adjust and adapt to new challenges. Other stakeholders, from patients' groups to companies, also have their own critical roles to play in advancing medical progress. The goal of the Manhattan Institute's Project FDA is to encourage an ecosystem for U.S. medical innovation where many partners work seamlessly together to advance truly disruptive medical innovations. With Huber's paper, and the painstaking work of the many experts and scientists that he catalogs, this vision is one step closer to becoming reality.

CONTENTS

1	Part 1: The Fading Myth of the FDA's "Gold Standard"
	Regulating Ignorance
	Targeted Drugs
	Molecular Medicine
	The Crowd of One
	Validating Biomarkers
	The New Gold Standard of Drug Science
	Engine versus Experts
	Algorithms Replace Labels
	Disassembling Patients
13	Part 2: Precision Medicine and the FDA
	The PCAST Proposal
	Adaptive Trials
	Enlisting the Right People
	Bayesian Statistics and Alternative States of Nature
	Statistics Bounded by Molecular Reality
	The New Science of Molecular Crowds
29	Endnotes

ABOUT THE AUTHOR

PETER HUBER is a senior fellow at the Manhattan Institute writing on the issues of drug development, energy, technology, and the law. He is the author of the forthcoming book, *The Cure in the Code: How 20th Century Law Is Undermining 21st Century Medicine* (Basic Books, Fall 2013).

Huber most recently wrote *The Bottomless Well*, coauthored with Mark Mills, which Bill Gates said “is the only book I’ve ever seen that really explains energy, its history and what it will be like going forward.” Huber’s previous book, *Hard Green: Saving the Environment from the Environmentalists* (Basic Books, 2000), which was called “the richest contribution ever made to the greening of the political mind” by William F. Buckley, Jr., set out a new conservative manifesto on the environment which advocates a return to conservation and environmental policy based on sound science and market economics. In 1997 he authored two books, *Law and Disorder in Cyberspace: Abolish the FCC and Let Common Law Rule the Telecosm* (Oxford University Press), which is an examination of telecommunications policy, and (with the University of Pennsylvania’s Kenneth Foster) *Judging Science, Scientific Knowledge and the Federal Courts* (MIT Press). Previous books include *Orwell’s Revenge: The 1984 Palimpsest*, (Free Press, 1994), *Galileo’s Revenge: Junk Science in the Courtroom* (Perseus Book Group, 1991); and *Liability: The Legal Revolution and its Consequences* (Basic Books, 1988).

Huber has also published articles in scholarly journals such as the *Harvard Law Review* and the *Yale Law Journal*, as well as many other publications, including *Science*, *The Wall Street Journal*, *Reason*, *Regulation*, and *National Review*. He has appeared on numerous television and radio programs, including *Face the Nation* and *The NewsHour with Jim Lehrer*.

Before joining the Manhattan Institute, Huber served as an assistant and later associate professor at MIT for six years. He clerked on the D.C. Circuit Court of Appeals for Judge Ruth Bader Ginsburg, and then on the U.S. Supreme Court for Justice Sandra Day O’Connor. Huber is also a partner at the Washington, D.C. law firm of Kellogg, Huber, Hansen, Todd, Evans & Figel.

Huber earned a doctorate in mechanical engineering from MIT and a law degree from Harvard University.

ACKNOWLEDGMENTS

The author is grateful for the excellent research and editorial assistance received from Matthew Hennessey, Yevgeniy Feyman, Kaitlin Keegan, and Janice Scheindlin.

THE DIGITAL FUTURE OF MOLECULAR MEDICINE: RETHINKING FDA REGULATION

Peter Huber

PART I: THE FADING MYTH OF THE FDA'S “GOLD STANDARD”

No drug may be licensed until the FDA is convinced that it will perform safely and effectively in future patients. All such predictions hinge, of course, on both the drug's chemistry and the patient's; pharmacology is not a science of one hand clapping. So the FDA does not license drugs—it licenses specified drug-patient combinations: the license's implicit promise of future safety and efficacy applies only “under the conditions of use prescribed, recommended, or suggested in the labeling thereof.”

The FDA has plenary authority to police how that science is developed. The agency has played a large and valuable role in developing protocols for laboratory tests, particularly for drug toxicity.¹ But the 1962 amendments to the federal drug law demanded, above all, “substantial” evidence, derived from “adequate and well-controlled” clinical trials. Fifty years ago, the FDA started drafting elaborate rules and protocols that spell out how Washington oversees the development of drug science. If, in the FDA's view, the science that is developed in this way predicts future benefits for certain patients, the FDA licenses the drug accompanied by a label that delineates who they are. These protocols, it is often said, establish the “gold standard” for drug science.

To this day, the drug-licensing process thus remains anchored in protocols developed at a time when pharmacology aspired, but mostly failed, to target molecules that no one could see, and

to control biochemical processes that no one could unravel. Today's biochemists and doctors, however, have the power to diagnose and treat with molecular precision from the bottom up.

Regulating Ignorance

Statistical analysis of the clinical symptoms of crowds is what medical science uses to pluck the most primitive form of cause-and-effect understanding out of the depths of ignorance. In Victorian London, it helped ferret out the cause of the city's periodic cholera epidemics. A doctor, John Snow, made the right connection in 1853: after a particularly nasty outbreak of cholera in Soho, he saved an unknown number of lives by persuading parish authorities to remove the handle from the neighborhood's Broad Street water pump. Germ science and the isolation of the cholera bacterium still lay three decades in the future.

The vaccines and antibiotics that followed worked wonders, but they owed their success to one brilliant trick—vaccines use biochemical fragments of the enemy microbe to fire up the human immune system—and lots of luck. Many of the early antibiotics were discovered by searching for microbes that had developed these molecules to kill their rivals. The first synthetic antibiotics were developed by chemists who happened to notice that some industrial dyes preferentially stained certain types of microbes; all the rest was intuition and guesswork. Insulin and estrogen, two pioneering drugs that tinkered directly with human chemistry, had likewise been designed by nature first. Most of the small number of other people-tuning drugs that emerged before 1962 were designed mainly by hunch, trial, and error—mostly error.

When prescribing the drugs of that era, doctors were guided almost entirely by clinical symptoms. Routine lab tests tracked only a few dozen infectious germs and a limited number of simple molecular “biomarkers”—blood-sugar levels, for example—that had clear, direct links to known diseases. Doctors knew little more about the molecular processes that made drugs perform well or badly.

The prevailing pharmacological model pictured magic-bullet molecules aimed at simple progenitors of disease. In this view, simple, clear lines separated disease and health. A discrete cause produced a discrete set of clinical symptoms—fluxes, fevers, lesions, or lumps—that uniquely defined the disease. The drug's story was the disease's, told in reverse. A single, straight line linked the drug to the root cause of the disease and the patient's return to health. A new drug did not need to be tested for long, nor did trials have to involve many patients. Medical science had scarcely begun to glimpse how one patient's chemistry can differ from another's, and it had little reason to suppose that such differences mattered much.

When a single visible cause is quickly and tightly connected to a visible effect that is easily tracked, simple statistical analyses are quite good at making the right connections. They correctly link a cluster of symptoms called “cholera” to a polluted well, or a bacterium, and the prevention or cure of cholera to the removal of the handle on the pump, or to the administration of a vaccine or antibiotic.

In implementing the 1962 federal drug law, the FDA accepted that view of things and expanded and standardized what Washington had begun doing in 1938, when the U.S. Public Health Service conducted a randomized trial of the pertussis vaccine in Norfolk, Virginia.² The FDA would scrutinize what the drug delivers up here—where patients ache and worry and clinicians diagnose and treat—not down there, where tetracycline (we now know) latches on to a specific receptor on the surface of the cholera bacterium. A good drug had to have the same effect in a large majority of patients suffering from the same, clinically defined disease because medical science lacked a way to distinguish patients whom the drug would help from those whom it wouldn't. The assumption—the blind hope, really—was that the FDA knew how to decide how many patients had to be tested, and for how long, to arrive at a robust statistical correlation and a label that would allow the drug to be prescribed safely and effectively to future patients.

The biggest worry was that wishful thinking by doctors or patients—“selection bias”—might

culminate in the licensing of drugs that did more harm than good—hence the randomized, “double-blind” clinical trials. To this day, Washington almost always requires and relies on the same kind of evidence—statistical comparisons of the health of two crowds—to decide whether a drug should be licensed. Typically, one crowd gets the real thing, the other a placebo; when a reasonably good treatment is already available, the comparison may instead be drug versus drug. Doctors track clinical symptoms. The newly healthy and the still sick, the living and the dead, vote the drug up or down.

These trial protocols, in short, are structured to regulate ignorance, not the systematic acquisition of reliable knowledge. They assume that the molecular science is impossibly difficult; the best we can do is search for strong statistical correlations linking a drug to its clinical effects. They do indeed set the gold standard—for dealing with blind ignorance. But when the cause-and-effect connections are complex, writing a good trial script requires information that only the trial itself can reveal.

If we were all exact biochemical clones of one another, testing a new drug in just two patients would suffice—one receiving the drug, the other a placebo. To expose how we differ in ways that affect a drug’s clinical performance, many more patients have to be tested, for a long time. But just how many, and for how long, depends on how many patient-side biochemical factors can affect the drug’s performance and how evenly or otherwise those factors are distributed among patients who will end up using the drug—biochemical facts that only extensive tests are likely to reveal.

Statisticians call this the “reference class problem,” or the problem of “external validity.” The relentless growth of FDA-mandated clinical trials since 1962 reflects the emperor’s own dawning realization that his wardrobe was furnished by Victoria’s Secret. Washington began losing confidence in quick, small clinical trials as science began to expose the slow, complex diversity of human chemistry. In the last decade, our newfound power to scrutinize everything down at the molecular level has exposed vastly more

biochemical diversity and complexity. And any molecular difference between two bodies might be the difference that allows the same drug to perform well in one body and badly in another.

The FDA’s conventional trial protocols deliberately lose all such details in the crowd, collapse biochemically complex phenomena into misleadingly simple, one-dimensional, yes/no verdicts, and will often reject good drugs that many patients need. They test too many of the wrong patients, and they develop the selection criteria for prescribing the drug to the right patients much too slowly, if at all. Today’s gold-standard molecular medicine is anchored in biochemical facts that the FDA takes pains to keep out of the sight of doctors conducting the front-end clinical trials, and it uses reams of empirical data that no drug company could collect and disseminate without risking prosecution for the promotion of off-label medicine.

Targeted Drugs

Science learns how to make consistently reliable predictions only by mastering the fundamental mechanics of cause and effect. Drugs are molecules that interact with other molecules in ways determined by mechanistic biochemical rules. The science is complex because drugs operate in the extremely complex biochemical environments of human bodies. But the rock-solid science that we are seeking is, ultimately, chemistry—precise, logical, and deterministic.

Drug designers have understood this for decades. The modern tools of “structure-based” drug design were first used successfully in the 1970s. The details are hard, but the idea is simple: hold a molecular blueprint of the disease up to a mirror, and you will see in the reflection molecular blueprints for one or more antidotes. With a promising molecular target in hand, drug designers now rely heavily on raw computing power to analyze the structure of the target and design mirror-image molecules. Alternatively, designers enlist the immune system of a genetically engineered laboratory animal to design antidotes—monoclonal antibodies.

Thalidomide, the notorious sedative that caused thousands of birth defects in the countries where it was licensed and, though never licensed in the United States, spurred the enactment of the 1962 drug-law amendments, would end up bridging the old era of pharmacology and the new. In 1964, shortly after it had become the most reviled drug in history, Jacob Sheskin, an Israeli physician, admitted to his ward a frantic woman suffering from the excruciatingly painful skin lesions and mouth ulcers that often develop in the later stages of leprosy.³ In an attempt to calm her down, he prescribed some left-over thalidomide that he happened to find on his shelf. Overnight, to his astonishment, her skin lesions and mouth ulcers were dramatically reduced. Dr. Sheskin's colleagues were skeptical; they couldn't imagine how a sedative could help treat a bacterial infection. To convince himself, Dr. Sheskin went to Venezuela, where leprosy was common, and conducted successful clinical trials. But he still had no clue as to why the drug worked, and medical science didn't have the tools to find out.

By the late 1980s, it did. Thalidomide doesn't attack the leprosy bacterium; it alleviates symptoms that develop when the infection sends the human immune system into overdrive. Researchers at Rockefeller University tracked the connection to a human protein called tumor necrosis factor, one of three intercellular signaling molecules (cytokines) that thalidomide suppresses. TNF plays important roles in the communication system that the body uses to fight both germs and cancerous human cells. But when engaged in a losing battle, the body sometimes produces too much TNF, which can then cause painful lumps and lesions on the skin. TNF overloads can also cause wasting syndrome, a common condition in the late stages of AIDS. Doctors treating AIDS patients grasped the implications and began prescribing thalidomide to treat ulcers and weight loss. Other doctors were soon investigating the drug's effects on a variety of skin disorders and other inflammatory conditions, as well as autoimmune diseases such as lupus and rheumatoid arthritis.

Meanwhile, other drug designers had begun designing precisely targeted drugs from scratch. In the 1970s,

three researchers at Squibb set out to tame a protease enzyme that snaps proteins apart in the process of manufacturing a hormone that helps control our blood pressure. In 1981, the FDA approved captopril, the first of the now widely used ACE inhibitors.

Gleevec was another early triumph of structure-based design. The first solid molecular link between a cancer—chronic myelogenous leukemia (CML)—and a flawed human gene had been discovered in 1960. The culprit is a corrupted version of a gene that codes for one of our many kinase enzymes. Scientists at the company now called Novartis developed computer models of the enzyme, used them to design various structures that might latch on exclusively to the CML-kinase binding pocket, synthesized them, tested the most promising ones, and got to Gleevec. It worked astonishingly well. Medicine now has “tools to probe the molecular anatomy of tumor cells in search of cancer-causing proteins,” the National Cancer Institute exulted when the license was issued in 2001. Gleevec is “proof that molecular targeting works.”⁴

A new disease called AIDS surfaced a month after the FDA licensed captopril. Soon after HIV was isolated, biochemists found the gene for a protease enzyme that the virus uses to assemble its protein shell, manufactured the enzyme itself, worked out its three-dimensional structure, and identified a key point of vulnerability. Then they designed the first HIV-protease inhibitor (saquinavir), which completed a lightning-fast trip through the FDA in 1995. Other drugs targeting other aspects of HIV's chemistry soon followed. As the National Academy of Sciences would observe in 2000, the remarkably fast development of HIV-protease inhibitors had a “revolutionary effect on modern drug design.”⁵

Molecular Medicine

The formerly blind doctors now have keen molecular vision, too. In early 2012, scientists at Stanford University described how they had spent the previous two years tracking DNA, RNA, cell proteins, antibodies, metabolites, and molecular signals—some 40,000 biomarkers that yielded billions of data points—in the body of geneticist Michael

Snyder, the team's senior member, to create the first-ever "integrative Personal 'Omics' Profile": an "iPOP."⁶ Though Snyder had no family history or conventional risk factors, the data revealed a genetic predisposition to type 2 diabetes. Later in the study, the data tracked the onset of the disease in what has been described as "the first eyewitness account—viewed on a molecular level—of the birth of a disease that affects millions of Americans." Then the iPOP team watched the diabetes markers revert to their normal state in response to treatment.

The technologies for designing and mass-producing the diagnostic biochemicals that power the iPOP, along with many less ambitious molecule sniffers already on the market, have been mastered. Arrayed on chip-size, micro-electro-mechanical laboratories, sniffers are now becoming complete bio-scanners that can, for a few dollars a whiff, search a cheek swab or a drop of blood for hundreds—and soon, thousands—of genes, proteins, and other biomarkers. Sensor chemicals on the surface of plastic or paper cards mounted in a breathalyzer can detect lung cancer, tuberculosis, and 100 other biomarkers associated with other disorders. More complex sequences of assays are now performed by automated banks of compact diagnostic machines that can quickly and cheaply diagnose germs, genes, and biochemical imbalances of every kind—in as many specimens of bodily fluids or tissues as anyone cares to supply.

The problem for medicine in general, and the FDA in particular, is that our keen molecular vision is revealing a great deal of complex biochemical diversity down there. Most diseases, as defined by their clinical symptoms, can't be tracked back to a single, magic-bullet cause. The clinical symptoms that once defined "the disease" are, in fact, produced by complex sequences and webs of biochemical cause and effect, configured in different ways in different bodies. There are many more biochemically distinct disorders at the bottom than clinical symptoms reveal to doctors up at the top. And drug-safety issues are almost always complex because side effects can potentially involve any part of the complex biochemistry of all the different bodies in which a drug may land.

The question, then, is how we develop the science that can reliably predict when, if at all, a drug can be safely and effectively prescribed to some patients when its performance may be determined by its interactions with different combinations of molecules in patients who are suffering from what looks, superficially, like the same disease.

The Crowd of One

The iPOPping of Michael Snyder began when he was, by all clinical appearances, perfectly healthy, and it thus established a biochemical baseline for his personal clinical health. The early genetic scan, however, revealed a genetic propensity for high cholesterol, which he already knew about, and for diabetes, which came as a surprise. He then watched his cholesterol level drop sharply when he started taking a cholesterol drug. After his blood-sugar level suddenly jumped on day 301 of the tracking, he watched aspirin, ibuprofen, exercise, and a low-sugar diet wrestle it back down.

For Michael, the patient, that might have been enough; but for Professor Snyder, the scientist, there was more to learn. Analysis of the iPOP data also revealed how his RNA was activating different genes at different points of the study. As the patient recounts, "we generated 2.67 billion individual reads of the [relevant RNA molecules], which gave us a degree of analysis that has never been achieved before.... This enabled us to see some very different processing and editing behaviors that no one had suspected. We also have two copies of each of our genes and we discovered they often behave differently during infection."

The researchers suspected a possible link between a viral infection and Snyder's blood-sugar surge 12 days after its onset, and they zeroed in on about 2,000 genes that were fired up during that period and another 2,000 that throttled down. They found among them links involving inflammatory proteins and antibodies, among them an auto-antibody that targets a human insulin receptor. The data pointed to "unexpected relationships and pathways between viral infection and type 2 diabetes." As one of Snyder's

colleagues notes, an analysis of this kind reveals how a patient's complex control systems interact with his own chemistry and the environment and thus point to how medicine "can best target treatment for many other complex diseases at a truly personal level."

In the iPOP world, it isn't just the medicine that gets personal; the science does, too. The science that describes the biochemical structure and dynamics of the disease and determines the efficacy and safety of the antidotes still involves a comparison of two or more patients, but they have the same name. "In a study like this, you are your own best control," says Professor Snyder. "You compare your altered, or infected, states with the values you see when you are healthy."

The development of this personal science does, however, build on a large body of knowledge previously acquired from other patients, and the data gleaned from Snyder's body will help refine how other diabetics use iPOP technology going forward. As Snyder notes, researchers with access to such data should be able to converge on a much smaller number of variables that can predict future blood-sugar health and track the rise and fall of diabetes and other diseases. But the picture that will likely emerge from this bottom-up, data-extravagant science isn't likely to please the crowd doctors. There are probably "many reasons why someone is at risk" of type 2 diabetes. "Diabetes is really hundreds of diabetes, and they just have one common characteristic, which is a high level of glucose." Different patients therefore require different treatments. "Some respond to metformin [a drug that suppresses glucose production in the liver], some don't. Some respond to anti-inflammatory medicine, some don't." And with diabetes, as with many other diseases, the key to effective prevention or treatment is "to catch it earlier."

The in-depth study of individual patients is the starting point for exposing such details. Prescribing one or more drugs and watching what happens in some larger group of biochemically similar patients is the surest way to pin down the causal connections. High blood sugar is a proximate cause of the clinical symptoms of diabetes; the best proof is supplied by

treatments that alleviate the symptoms by controlling the sugar. An inflammatory protein may be an antecedent cause, disrupting the insulin chemistry that ordinarily controls the sugar; medicine has been studying this possibility for some years and will confirm it by testing anti-inflammatory drugs in present or prospective diabetics. Using a drug to verify the link between a biomarker and a clinical effect is the molecular version of removing the handle from the pump. Or it may take several drugs, used simultaneously or sequentially, to establish that the complex cause that underlies the disorder can be beaten only with a complex treatment.

There is no practical substitute for this approach; biochemically complex diseases don't have a single handle. By studying patients alone, researchers are rapidly exposing many promising targets that have clear statistical associations with the many intractable disorders that we still face. Drug designers have the tools to create molecules that will control many of those targets, and lab tests often confirm that the precisely targeted drugs perform as expected. Yet the drugs often don't perform as hoped in clinical trials. It's the patients and their diseases that aren't cooperating. The magic molecular bullets work one on one but fail to consistently deliver the hoped-for clinical effects in FDA-scripted trials.

Given what we now know about the biochemical complexity and diversity of the environments in which drugs operate, the unresolved question at the end of many failed clinical trials is whether it was the drug that failed or the FDA-approved script. It's all too easy for a bad script to make a good drug look awful. The disease, as clinically defined, is, in fact, a cluster of many distinct diseases: a coalition of nine biochemical minorities, each with a slightly different form of the disease, vetoes the drug that would help the tenth. Or a biochemical majority vetoes the drug that would help a minority. Or the good drug or cocktail fails because the disease's biochemistry changes quickly but at different rates in different patients, and to remain effective, treatments have to be changed in tandem; but the clinical trial is set to continue for some fixed period that doesn't align with the dynamics of the disease in enough patients.

Or side effects in a biochemical minority veto a drug or cocktail that works well for the majority. Some cocktail cures that we need may well be composed of drugs that can't deliver *any* useful clinical effects until combined in complex ways. Getting that kind of medicine through today's FDA would be, for all practical purposes, impossible.

For a drug to perform well, we need to select the patients to fit it. Ideally, the in/out selection criteria will span all the patient-side molecules that will affect a drug's performance, in all the different combinations that occur in different patients. But most of the time, we don't know what all or even most of those biomarkers are—and we won't find out until we test the drug or drug cocktail in enough patients to expose them.

The FDA doesn't know, either—and it doesn't want biomarkers involved in the licensing process until it does. That is the biggest obstacle that now stands between us and the future of molecular medicine.

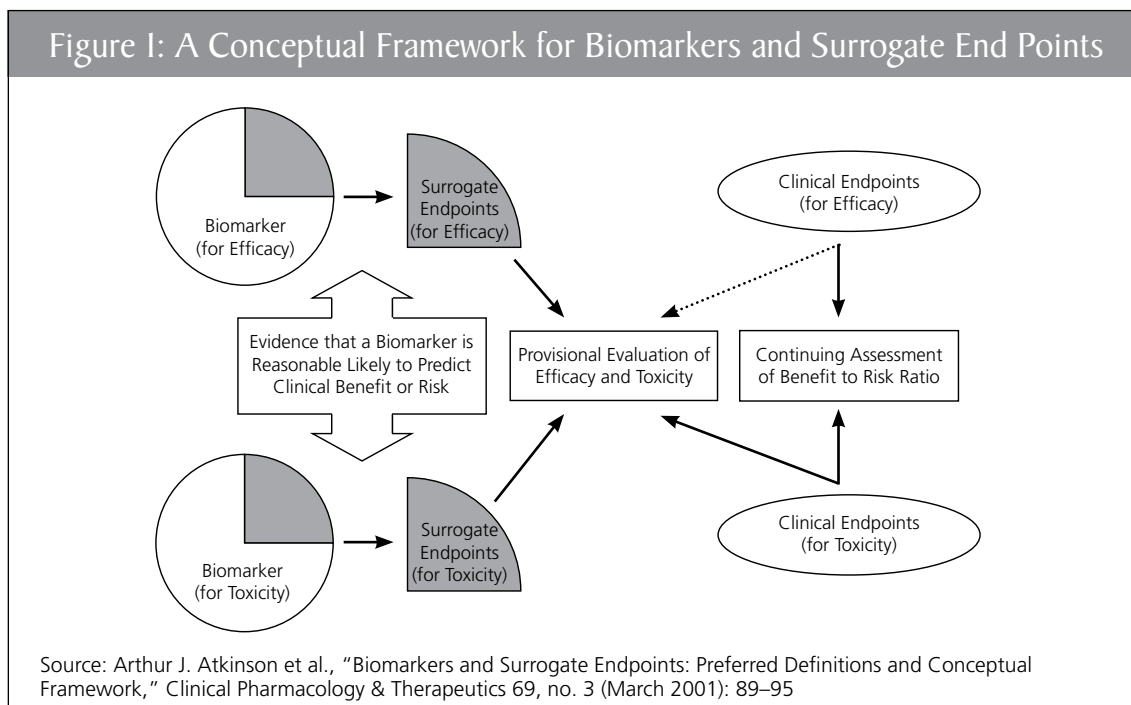
Validating Biomarkers

So we arrive at what the FDA calls “validating” biomarkers. They aren't manufactured by drug

companies, but that detail aside, we are back to 1962. Once again, the agency is struggling to decide how to decide when a molecule is likely to affect clinical health, for better or worse. At issue now are the patient-side molecules that a candidate drug will interact with, directly or indirectly, in good ways or bad.

For over two decades, the FDA has accepted—in principle—the use of biomarkers in drug licensing. The FDA, NIH, and Congress have been issuing general and vaguely encouraging biomarker pronouncements and guidelines, as well as launching related studies, since the late 1980s. In 1997, Congress directed the FDA to establish a program to accelerate the process.

The FDA, however, has unlimited discretion to remain dissatisfied with the quality of biomarker science, and, by and large, it has. The FDA points out, correctly, that linking what happens down there to what then happens up here can be tricky, and if we get the linkage wrong, the FDA may end up licensing drugs that are useless or worse. So the FDA won't accept the use of biomarkers until it is convinced that their use is “reasonably likely” to translate into clinical benefits. What kind of convincing should it take?



Molecular medicine often determines how strongly molecules—cholesterol, for example, or a high-cholesterol gene—are linked to clinical problems by searching for statistical correlations in large databases of patient records that include both molecular and clinical data. Strong links then point to promising drug targets. And they can be found before a clinical trial of, say, a cholesterol-targeting drug begins.

The same statistical tools can then be used to analyze links between drug-biomarker combinations and clinical effects. As it's acquired, this information can be used to refine prescription protocols in ways that improve both efficacy and safety. Such studies have identified genetic biomarkers that can tell you in advance whether you will respond well or badly to a growing number of drugs, among them, anticoagulants, antidepressants, painkillers, and drugs used to treat heart disease, high blood pressure, hepatitis, and various cancers.

But here's the catch: most of the drug-biomarker science can't be developed before human trials begin. FDA protocols allow very little of it, if any, to be developed during the front-end licensing trials. So most of this invaluable predictive molecular science is developed after a drug has been licensed and prescribed to many patients—many of whom, we discover, should never have used it.

A drug designed to target an estrogen receptor, for example, should obviously be tested only in the ER+ breast-cancer patients whose tumors present that target. But if the breast-cancer drug's performance also depends on how it is metabolized in the patient's liver, as tamoxifen's does, the existence of a genetic marker that identifies the patients with the right liver often won't be discovered until doctors begin exploring why different ER+ patients respond differently to the same targeted drug. A drug's selective efficacy can also depend on a wide range of other biomarkers that are hard to identify in advance. Hitting the drug's intended target may not suffice: complex diseases may respond only to multipronged attacks—in which case, the selection criteria for testing today's drug ought to include the selection of other drugs needed to complete the synergistic

cocktail. Which means that it may be impossible to test the drug in the right biochemical environment until complementary drugs are available—and the same may be true for each of those other drugs.

Before a trial begins, it is even more difficult to specify selection criteria for excluding patients in whom the drug will cause serious side effects. The FDA itself helped launch a nonprofit consortium of ten drug companies and academic institutions to compile a global database of genetic links to drug side effects. In 2010, the group released data that help predict when drugs are likely to cause serious harm to a patient's liver or trigger a potentially lethal allergic response. Similar initiatives have exposed genetic variations that make other drugs ineffective and are developing genetic-profile standards to guide more accurate prescriptions. Better late than never; but detecting these links during front-end licensing trials would have been very much better.

Many of these links could, and should, be detected earlier, because another way to develop drug-biomarker science is to study how individual bodies interact with drugs at the molecular level—as was done, for example, in Stanford's iPOP study. The FDA knows that, too, and it recently began approving a range of what are, by Washington's standards, innovative “adaptive” trial protocols that allow that to happen.⁷ But the FDA remains slow and reluctant to approve such trials and unwilling to accept the complex analytical tools that extract reliable scientific knowledge efficiently from extremely complex data sets.

The problem for the FDA is that robust drug-biomarker science can't be fully developed without testing drugs in a broad range of biochemically different patients and carefully studying and comparing their responses. That means removing the FDA's cherished blindfolds and replacing simple trial protocols that analyze comparatively tiny amounts of data with complex protocols that analyze torrents of data—not the kind of change that ever happens quickly in Washington. The FDA, as currently structured and funded, lacks the institutional resources—and perhaps also the expertise—to keep up with the converging,

synergistic power of the biochemical and digital revolutions. Bureaucratic inertia may also be a factor—the indiscriminate testing required by the FDA’s current trial protocols is familiar and much easier to regulate. At stake, unfortunately, is the entire future of molecular medicine.

The New Gold Standard of Drug Science

Three significant loopholes in the existing drug law have already shown us how today’s gold-standard molecular medicine evolves when doctors are given enough latitude to develop much of the drug science from the bottom up. The first two loopholes can bring the FDA fairly close to what might be called “tool-kit licensing”: license drugs as molecular scalpels or sutures in search of the right disease. The third (and, by far, the largest) loophole allows doctors to start using drugs in exactly that way as soon as they are licensed. Ignore the label, and prescribe the drug to patients whose disorder presents the target that the drug was designed to control. Use the available molecular tools simultaneously or sequentially, in a way that makes mechanistic sense, much as surgeons use their tools. Work out the connections between molecular and clinical effects on your own, one patient at a time.

The 1983 Orphan Drug Act directs the FDA to be flexible in evaluating evidence that a drug is effective. The act covers drugs directed at a rare disease, many of which are caused by a single, rare genetic disorder, associated with a single protein that an effective drug can target. This makes it easy to frame trials that fit the drug to the right patients from the get-go and track at least one key aspect of its performance at the molecular level. The act then gives the FDA broad flexibility to license drugs on the strength of individual patient case reports, or even studies conducted in animals or laboratory glassware. Drugs designed and licensed this way are, in effect, recognized and used as molecular tool-kit drugs from the start.

The FDA has designated as orphans about 7,000 rare conditions that collectively affect some 30 million Americans, and it has approved about 350 orphan

drugs.⁸ The orphanage currently fosters about one-third of the FDA’s successful graduates and is now home to “the most rapidly expanding area of drug development.” This is widely viewed as a “roaring success.” Over half of all certified orphans end up as wards of Big Pharma, and quite a few end up treating big crowds, when it turns out that the drug’s molecular target propels other diseases as well.

Then there is the FDA’s own accelerated-approval rule, promulgated in 1993, codified and somewhat expanded by Congress in 1997, and endorsed again in 2012. When the disease is sufficiently serious and available treatments are inadequate, a new drug can get to market by demonstrating that it does indeed produce its intended molecular-scale effect—lowering blood-sugar levels, for example—or, more generally, that it produces favorable changes in what the FDA calls “surrogate end points” without waiting for favorable changes in clinical effects that often take much longer to surface. The front-end trials need not resolve concerns about how the drug’s performance might be affected by many aspects of biochemical diversity or about long-term side effects. The manufacturer must still complete controlled trials after the drug is conditionally licensed; meanwhile, the drug can be prescribed by unblinded doctors who can gather information that clarifies how it can be used well. The license is rescinded if things don’t pan out.

Finally, the 1962 law left doctors free to prescribe licensed drugs “off-label.” Once a drug is licensed for one purpose, however narrow, it may legally be prescribed for any purpose. The doctor and patient will have some assurance that the drug isn’t immediately toxic, but efficacy is entirely up to them. The FDA itself brazenly relied on this aspect of the law to help rush thalidomide into the U.S. market. After desperate HIV patients began smuggling the drug into the U.S., the FDA asked drug companies to consider cashing in on the leprosy epidemic that was not sweeping across America. Celgene accepted the invitation, presented leprosy-related clinical data, and the FDA licensed thalidomide for sale—to leprosy patients—in 1998. Sales boomed, overwhelmingly to HIV-positive patients.

But for these three major licensing loopholes, millions of people alive today would have died in the 1990s. Almost all the early HIV- and AIDS-related drugs—thalidomide among them—were designated as orphans. Most were rushed through the FDA under the accelerated-approval rule. Many were widely prescribed off-label. Oncology is the other field in which the orphanage, accelerated approval, and off-label prescription have already played a large role. Between 1992 and 2010, the rule accelerated patient access to 35 cancer drugs used in 47 new treatments. For the 26 that had completed conventional follow-up trials by the end of that period, the median acceleration time was almost four years.

Together, HIV and some cancers have also gone on to demonstrate what must replace the binary, yes/no licensing calls and the preposterously out-of-date Washington-approved label in the realm of complex molecular medicine. The new gold standard of molecular medicine looks nothing like the old.

Engine versus Experts

The first HIV drug to arrive in Washington—AZT—had been developed (as a cancer drug) in the early 1960s but never licensed. Tested against HIV in the lab two decades later, it looked promising. But HIV is typically invisible and seemingly harmless for about five years after the initial infection. What if HIV was able to mutate its way into an AZT-resistant form faster than that? Or caused grave side effects that took four years to surface? AZT couldn't prove that it was good for patients—at least, not to Washington's satisfaction—any faster than HIV killed them.

So the FDA approved a first AZT trial limited to HIV patients who had also been infected with a rare form of fungal pneumonia, one of the most common epitaph killers when the patient develops full-blown AIDS. The trial had to be terminated prematurely, when the dead-patient count reached 19-1 against the placebo—doctors can't ethically keep prescribing a placebo just to run up the score once it becomes clear that the drug works. The FDA immediately licensed AZT for use by HIV-plus-fungus-positive patients. The fungus restriction was, of course, widely ignored.

It took another three years for the FDA to broaden AZT's license to cover early-stage treatment. Soon after, the FDA formalized its accelerated-approval rule. By early 1998, the rule had expedited the licensing of some 27 cancer and HIV drugs, along with 16 drugs for other conditions, several of which most commonly occur in cancer or AIDS patients.⁹

HIV quickly developed resistance to AZT. In the interim, however, biochemists had been designing other HIV drugs. The FDA gunned its licensing engine, and doctors were soon concocting three-drug cocktails that the virus isn't nimble enough to evade. About 20 HIV drugs have since been approved worldwide; they are typically used in about ten fairly standard cocktails. The efficacy of each cocktail depends on which strain launched the infection and how it has evolved inside the patient being treated. Different forms of the disease predominate in different countries and track gender, sexual practices, and other factors.

So, viewed from the treatment perspective, medicine is now dealing here with about ten different diseases, each of which is forever poised to mutate into some new, untreatable form. Treatments work well when the doctor selects just the right trio of molecular scalpels from the available drug tool kit. Selecting them isn't easy because so many different variables can affect how each possible combination of drugs performs in different patients. Until quite recently, trial and error played a large role. The doctor started with one mix, monitored viral loads and other biomarkers in the patient's bloodstream, and adjusted the treatment accordingly.

Today, the process is guided by sophisticated analytical engines fueled by huge collections of patient records that include data on HIV genotypes, treatment histories, and responses, along with patient age, gender, race, and route of infection; patient genotypes will undoubtedly be added sooner or later. As of late 2011, the largest such engine—Europe's EuResist Network, powered by IBM technology—was using data from 49,000 patients involving 130,000 treatment regimens associated with 1.2 million records of viral genetic

sequences, viral loads, and white blood-cell counts.¹⁰ As described by its manager, the EuResist database is “continuously updated with new data in order to improve the accuracy of the prediction system.” When tested against 25 actual case histories that weren’t already in its database, EuResist beat nine out of ten international experts in predicting how well the treatments had performed. The study was dubbed “Engine versus Experts.”

Whatever we may call it up here, there is no single disease down there, and the disease down there tomorrow will be different from today’s. When the FDA licensed the individual drugs or the cocktails, it clearly lacked “substantial evidence” that the drugs or cocktails would perform effectively when directed against any substantial fraction of all the variations in HIV and patient chemistry that they might encounter in the future. That evidence was acquired later and is now translated into complex treatment protocols by experienced doctors or analytical engines like EuResist. The virus continues to evolve, so the cocktails will remain safe and effective, in any meaningful sense of those words, only so long as we continue to prescribe them as directed by continuously updated databases. Whatever they permit or proscribe, the FDA’s licenses and labels will always lag far behind the virus.

Algorithms Replace Labels

Cancer drugs were the other early beneficiaries of the three main licensing loopholes. But cancer cells present a far broader range of biochemical complexity, and the FDA has licensed only a tiny fraction of the drugs that oncologists need to beat them.

Gleevec got the benefit of both the orphanage and the accelerated-approval rule. In the Gleevec-versus-CML trials launched in 2000, doctors assessed the drug’s performance by tracking two types of cell counts. The FDA reviewed the results in three months, conceded that it didn’t yet know whether the drug would keep patients alive longer, and in 2001 licensed Gleevec, anyway.

Almost immediately, oncologists began experimenting with Gleevec in the treatment of other cancers,

and it soon landed a second license to treat a rare gastrointestinal cancer. Other orphan designations followed, and the drug has been widely prescribed off-label. At its peak, little orphan Gleevec was raking in \$5 billion a year. Gleevec and other orphan billionaires epitomize the gulf between the old medicine and the new. The orphanage still defines disease from the top down. Biochemists and doctors fit one drug to multiple diseases by finding a molecular target that they share.

But Gleevec also fails to help about one CML patient in ten. To put it another way, it is the old medical taxonomy that has failed: CML, we now know, is one name for at least two distinct diseases, each of which can spawn others. About two out of every five patients on Gleevec benefit at first but then relapse because their cancer cells mutate into a Gleevec-resistant form. Most cancers exhibit similar behavior—they mutate so frequently that, viewed from a biochemical perspective, “the cancer” is really an engine for spawning a limitless number of different cancers. At major research hospitals, oncologists now sequence the complete genome of different parts of a single tumor, in a search for targets that will be used to guide treatment. The therapies that work often consist of complex drug cocktails that are tailored to—and repeatedly adjusted to track—the disease’s dynamic complexity.

Working with the drugs that they do have, oncologists routinely prescribe cancer drugs and cocktails far outside the boundaries that were tested in blinded licensing trials and are set out in the FDA-approved label. A nonprofit alliance of 21 leading cancer centers evaluates and publishes information on off-label uses. Off-label and cocktail therapies sometimes end up being steered through the rigid, slow, and expensive trials scripted by the FDA. But as a practical matter, the vast majority never will be; there are just too many combinations of drugs, dosages, and patient profiles to explore and calibrate.

With breast cancer, the bottom-up development of the drug science has already traveled a good distance down the same path as HIV. Defined by its clinical symptoms, breast cancer is a single disease that kills about 40,000 Americans a year. For oncologists,

however, the disease now comes with initials—ER, PR, and HER2, for example—with a plus or minus sign attached to each one, depending on whether the malignant cells have receptors for estrogen, progesterone, or a human epidermal growth factor. The medical literature first mentioned the “triple negative” form in late 2005; it has since been the subject of hundreds of research papers. Drugs are prescribed accordingly. Tamoxifen, for example, is used to block estrogen receptors on the ER+ form of breast cancer. But estrogen itself is used to treat other ER- forms, and some studies indicate that estrogen can be used prophylactically to lower the incidence of breast cancer in some postmenopausal women. “The story of estrogen’s role in breast cancer,” an article in the *Journal of the National Cancer Institute* recently observed, “is starting to look like Dr. Jekyll and Mr. Hyde.”¹¹

Over a decade after tamoxifen (an estrogen blocker) was licensed, studies revealed that most of the effective blockers are produced when tamoxifen is metabolized in the liver. But significant numbers of women (the numbers vary significantly across ethnic lines) have two copies of a gene that produces the enzyme in a form that can’t activate the drug, and women who have one copy activate much less of it.

Multidrug breast-cancer regimens have to be tailored to fit all the biochemical variations in tumors, livers, and other parts of the patient’s body that may affect each drug’s performance. The regimens are often adjusted during the course of treatment, as the mutating cancer cells develop resistance to some drugs and susceptibility to others. Prophylactic drugs may well have to address a quite different set of biochemical processes. Some women, for example, are very likely to develop cancer in at least one breast because they carry flawed versions of a gene that produces a protein involved in repairing genetic errors.

So much for the magic-bullet disease struck by a magic-bullet drug. Oncologists now speak of treatment “algorithms”—sets of rules for selecting and combining the array of available drugs in a much broader array of cocktails. A consensus

statement released by breast-cancer specialists in 2009 announced that “the previous attempt to produce a single-risk categorization and a separate therapy recommendation are no longer considered appropriate.” Three years later, a major international study of genes that promote or suppress breast cancer concluded that breast cancer is now an “umbrella term” for “10 quite distinct diseases.” Biochemists and oncologists now have in hand a new “molecular map” to guide both treatment and the development of new drugs.¹² The maps and algorithms will undoubtedly continue to be refined for years to come.

Digital engines will almost certainly end up doing most of the refining. In February 2013, IBM announced the arrival of a new engine—Interactive Care Insights for Oncology, powered by Watson—that apparently aims to do for oncology what EuResist does for HIV.¹³ Developed in partnership with WellPoint and Memorial Sloan-Kettering, and powered by the supercomputer that won the engine-versus-experts challenge on Jeopardy, the engine was initially drawing on “600,000 pieces of medical evidence, two million pages of text from 42 medical journals and clinical trials in the area of oncology research. Watson has the power to sift through 1.5 million patient records representing decades of cancer treatment history, such as medical records and patient outcomes.... Watson continues to learn while on the job, much like a medical resident, while working with the WellPoint nurses who originally conducted its training.”

Disassembling Patients

The magic bullets beat the easy problems. Most diseases that medicine is now struggling with will, in all likelihood, turn out to be much more difficult—more like HIV or breast cancer than cholera. In the grand biological scheme of things, simple, static, one-size design is the path to extinction. Survival lies in complexity: the temporal complexity of viruses such as HIV, which thrive by mutating very fast; or the complexity of cancers, which reveal the human body’s capacity to spawn biochemical complexity at its malignant worst.

Researchers investigating the wild mutability of cancer cells recently discovered that humans share with apes a biochemical quirk that introduces “copy number variations” (CNVs) into our genome.¹⁴ When our cells replicate, whole paragraphs and pages of genetic code are sometimes duplicated, written backward, abridged, or ripped out. CNVs occur in our reproductive cells, too. Their discovery, in the words of one geneticist, has lifted the veil “on a whole new level of genetic diversity.”

Each of us also carries thousands of genetic spelling errors—“single nucleotide polymorphisms,” or “snips.” A recently published study analyzed snips in potential “drug target genes” of 14,000 individuals thought to be particularly susceptible to heart attacks, strokes, obesity, and other health problems.¹⁵ The average subject was found to carry about 14,000 snips, about 12,000 of which were very rare. Each subject carried an estimated 300 genes with rare variants (found in less than 0.5 percent of the population) that would disrupt a protein’s functionality in ways that were likely to undermine health and affect how the individual might respond to drugs. Most of the rare variants, as the *Science News* report on the study put it, are “practically secret family recipes. Others reveal the distinct flavor of geographic regions, much like wines or cheeses.”

Biochemists used to assume that when common disorders ran in families, they were caused by a common variation in a single gene, or perhaps a small cluster of genes. But seemingly common disorders, it now appears, often reflect large numbers of rare, distinct genetic flaws that cause similar clinical symptoms. A neural connection that depends on the interaction of two different proteins, for example, can be disrupted by a flaw in either of the two underlying genes. Evidence in favor of the “common-disease rare-variant” hypothesis is rapidly accumulating. Hundreds of different proteins that control the interfaces between nerve cells, for example, can apparently play roles in choreographing Alzheimer’s, Parkinson’s, epilepsy, and more than 130 other brain disorders.

The endlessly diverse biochemical ecosystems that shape our health also determine how a drug performs

in different bodies. To beat most biochemically complex diseases, we will need a pharmacy stocked with a concomitantly large and diverse array of targeted drugs, together with complex protocols for prescribing complex treatments. We will develop this cornucopia of drugs and treatment regimens only by extracting vast amounts of biochemical information from a very large number of human bodies and working out how the pieces interact.

The gold-standard drug science that gets these drugs licensed will be anchored in mechanistic facts about how specific arrays of other molecules will affect a drug’s clinical effects. Those facts alone aren’t sufficient but are necessary: without them, many drugs that we need can’t perform well, and most diseases won’t be cured. As the EuResist engine and breast-cancer treatment algorithms illustrate, the best predictions of how drugs will perform are provided by a sophisticated and continuously improving mix of the rock-solid biochemical facts and empirical data—with the mix shifting steadily toward the former.

These engines and algorithms still rely on empirical data—but they do so not to pass final judgment on any single drug or drug cocktail but to reveal more complex patterns that can be used to transform the core, patient-specific biochemical facts into a personalized prediction of likely clinical effects, good and bad, that targeted drugs will have in the unique biochemical environment of an individual body. Every advance in the biochemical science diminishes the need for empirical correlations by narrowing the scope of the biochemical uncertainty. As Dr. Janet Woodcock, director of the FDA’s Center for Drug Evaluation and Research, put it in 2004, “biomarkers are the foundation of evidence based medicine—who should be treated, how and with what.... Outcomes happen to people, not populations.”¹⁶

PART 2: PRECISION MEDICINE AND THE FDA

In late 2011, a committee convened by the National Research Council (NRC) at the request of the National Institutes of Health (NIH) released a landmark report addressing the need for “a

new taxonomy of human diseases based on molecular biology” and outlining how that taxonomy might be developed.¹⁷ Redefining diseases on the basis of their biochemistry, it concludes, will require the analysis of “biological and other relevant clinical data derived from large and ethnically diverse populations,” in a dynamic, learn-as-you-go collaboration among biochemists, clinical specialists, patients, and others. As it happens, good *drug* science requires much the same—a drug is just one more molecule added to the molecular ecosystem that constitutes a body. The NRC report assumes as much when it recommends that doctors be allowed to consult the proposed network to find out how other patients have fared when already-licensed drugs are prescribed outside the FDA-approved boundaries. As the NRC report makes clear, the objective is “precision medicine.” A molecular taxonomy of disease is only the starting point that leads to precisely targeted drugs and precise prescription protocols.

The several elements of precision medicine are tightly linked. Every time we prescribe a targeted drug, whether during a licensing trial or thereafter, we simultaneously test and have the opportunity to improve our molecular understanding of the disease that it targets. We confirm that the bacterium is the cause of the disease, for example, by targeting it with an antibiotic and watching the patient recover—or discover that the microbe has mutated into some new form when the previously effective drug fails. Every drug is also a potential cause of other diseases. Tamoxifen suppresses some forms of breast cancer but raises the risk of some forms of uterine cancer.

Precision medicine hinges on systematic patient selection—selection that is based on the drug’s intended target and on unintended targets associated with side effects and on other drugs that may be prescribed at the same time. The include/exclude calls will often have to be repeated on the fly, as the patient’s biochemistry changes (or fails to change) during the course of treatment. A good clinical trial of a good drug will develop the information that future doctors will use to select the patients who have what it takes to make the drug perform well. The

best protocols will be based on molecular markers and effects; the whole point, after all, is not to wait for clinical effects to reveal whether the drug was prescribed well.

But the FDA’s current protocols treat patient selection as a problem that the drug company must solve before the clinical trial begins or, to a limited extent, when it is in its very early phases. At best, this means that the drug is prescribed to many patients whom it fails to help or even harms during the trials, and to still more of the wrong patients after it’s licensed, until enough post-licensing data accumulates and reveals how to prescribe the drug more precisely. At worst, drugs that some patients desperately need don’t get licensed because the trials include too many of the wrong patients. Either way, testing a drug in many of the wrong patients wastes a great deal of time and money. At some point, the cost of relying on this very inefficient process to try to solidify the science up front surpasses how much the drug is likely to earn years later in the market. We then have an economically incurable disease.

The PCAST Proposal

Nine months after the NRC issued its report, President Obama’s Council of Advisors on Science and Technology (PCAST) released a report, “Propelling Innovation in Drug Discovery, Development, and Evaluation,” which picks up roughly where the NRC report leaves off. The FDA’s standard trial protocols, the PCAST report notes, “have only a very limited ability to explore multiple factors ... [among them] individual patient responses to a drug, the effects of simultaneous multiple treatment interventions, and the diversity of biomarkers and disease subtypes.” These protocols lead to clinical trials that are “expensive because they often must be extremely large or long to provide sufficient evidence about efficacy.” The report goes on to outline a proposal for ushering the FDA into the future of molecular medicine. It has five main elements:

- The FDA should use its existing accelerated-approval rule, which “allowed for the development of pioneering and lifesaving HIV/AIDS and

cancer drugs over the past two decades,” as the foundation for reforming the trial protocols used for all drugs that address an unmet medical need for a serious or life-threatening illness.

- The molecular science used to select targets and patients should be anchored in human rather than cell or animal data, and it can be developed, in part, during the clinical trials. Before the trials begin, statistical studies of naturally occurring genetic variations can provide valuable guidance on biomarkers to target and track and will grow increasingly useful as databases that combine genomic and clinical data grow larger. “Clinical investigational studies with small numbers of patients but extensive data gathering” are an “extremely valuable” alternative.
- The FDA should adopt “modern statistical designs” to handle the data-intensive trials and explore multiple causal factors simultaneously—among them, “individual patient responses to a drug, the effects of simultaneous multiple treatment interventions, and the diversity of biomarkers and disease subtypes.” These designs are much more efficient than the FDA’s conventional protocols, and the patients involved receive, on average, better treatments.
- The FDA should also “expand the scope of acceptable endpoints” used to grant accelerated approval. Specifically, the FDA should make wider use of “intermediate” end points—indications that a drug provides “some degree of clinical benefit to patients” though the benefits “fall short of the desired, longterm meaningful clinical outcome from a treatment.” The FDA has granted only 11 such approvals in the past 20 years. It should “signal to industry that this path for approval could be used for more types of drugs” and “specify what kinds of candidates and diseases would qualify.”
- These initiatives should be complemented by greater rigor in enforcing and fulfilling requirements that follow up confirmatory studies that demonstrate the actual efficacy of drugs on

clinical outcome, and the FDA should continue and possibly expand its use of reporting systems that track both efficacy and side effects in the marketplace. The FDA should also consider a process of incremental licensing that begins with accelerated approval for use of the drug only in treating “a specific subpopulation at high risk from the disease” when larger trials would take much longer or wouldn’t be feasible. The license could then be broadened to authorize broader use upon the successful completion of broader trials. The FDA would “strongly discourage”—but not forbid—off-label use in the interim.

Vigorously implemented, these proposals could go a long way toward aligning FDA regulation with the drug development tools and practice of modern molecular medicine. The accelerated-approval rule puts the focus on molecular-scale or other low-level effects from the start. Protocols that allow the efficient, integrated development of drug-biomarker science lead to smaller, less expensive trials because they lead simultaneously to narrower, safer, and more effective prescription protocols—or to the conclusion that the drug has no useful role to play.

Broadening the standard for accelerated approval to include successful achievement of “intermediate” end points is a good starting point in addressing the most fundamental issue of all: What should it take to meet the federal drug law’s demand for “substantial evidence” in the age of molecular medicine backed by the pattern-recognition power of digital technology? The PCAST report addresses that question only indirectly; it needs to be addressed head-on.

Adaptive Trials

To acquire the tools that medicine needs to deal successfully with complex diseases that require complex treatments, we will have to develop treatment regimens piece by piece, each piece consisting of a drug and a solid understanding of how a cluster of biomarkers can affect that drug’s performance. Demanding a front-end demonstration that each piece will deliver clinical benefits on its own will only ensure that no treatment for the disease is

ever developed. An intermediate end point—“some degree of clinical benefit”—suggests that the drug is interacting in a promising way with a molecular factor that plays a role in propelling the disease; that is the best we can expect from any single piece. Even that requirement may be too demanding—used on their own, the individual constituents of some multidrug treatments that we need may never be able to deliver any clinical benefit at all.

When the first HIV protease inhibitor, for example, showed that it did its job and thus lowered viral loads, it was a drug that medicine clearly wanted to have on the shelf—even though it would take several more years to develop additional drugs and assemble cocktails that could suppress the virus almost completely, and for a long time; likewise with the first estrogen inhibitor for breast-cancer patients. A drug may have some modest, short-term effect on a patient’s clinical health but have no lasting effect on the progress of the disease because the virions and cancer cells are quick to mutate their way past any single-pronged attack. But if the drug offers a biochemically new approach to attacking the disease, it should be licensed, anyway. A successful attack on a biochemically nimble virus or cancer has to begin somewhere, and the place to begin is with a targeted drug that has demonstrated its ability to disrupt some molecular aspect of the disease’s chemistry in a way that had some promising effect, in some patients, at some point further along in the biochemical process that propels the disease.

By allowing broader use of the drug by unblinded doctors, accelerated approval based on molecular or modest—and perhaps only temporary—clinical benefits launches the process that allows more doctors to work out the rest of the biomarker science and spurs the development of additional drugs. The FDA’s focus shifts from licensing drugs, one by one, to regulating a process that develops the integrated drug-patient science to arrive at complex, often multidrug, prescription protocols that can beat biochemically complex diseases. The FDA already has the authority to monitor and regulate that follow-up process and to modify or rescind the initial license if the clinical benefits don’t materialize.

Adaptive trials can be structured in many different ways; the details are beyond the scope of this paper. The PCAST report includes a description of the I-SPY 1 (2002–06) and I-SPY 2 (ongoing as of early 2013) trials of breast-cancer drugs.

In brief, adaptive trials gather a great deal of data, focusing at first on effects down at the bottom—tracking genes, proteins, microbes, and other biomarkers that control the trajectory of the disease and cause different patients to respond differently to the same treatment regimens. As Stanford’s iPOP study demonstrated, and the PCAST report notes, the data-intensive study of quite small numbers of patients can substitute for the statistical analyses of crowds. The protocols evolve as the trial progresses and the collective understanding of the drug-patient molecular science improves.

Data-pooling networks and pattern-recognition computers should be used to systematize the process from the outset. Informed by a constantly expanding database of patient experience, the computers will be engaged in the rigorous process of learning incrementally from uncertain observations of complex phenomena—a process, as discussed shortly, that relies on Bayesian (or similar) statistical methods.

The selection of additional biomarkers for use in refining the selection of additional patients to include in the trials can be guided by a mechanistic biochemical understanding of why the biomarkers are relevant, along with the types of data already used by the FDA when licensing orphan drugs. Laboratory tests, such as those already developed to mimic various aspects of the human liver or heart cells, can be used to confirm that a drug can indeed interact with a biomarker in a way likely to affect the drug’s performance. The in-depth investigation of the response of individual patients, coupled with today’s sophisticated laboratory tests, can do much to ensure that the biomarkers that are used to stack the patient deck in a drug’s favor are based on objective criteria rather than on wishful thinking. The analytical engines that quantify the strength of links between drug-biomarker combinations and clinical effects need not even know whether the effects are

medically good or bad; regulators can see to it that the computers wear the blindfolds.¹⁸

If the analytical engine is doing its job well, the adaptive trial will progressively hone in on the taxonomic aspects of the disease—if any—that determine when a drug can perform well down at the molecular and cellular level, along with biomarkers that determine when the drug causes unacceptable side effects. The drug's clinical performance should steadily improve as treating doctors gain access to the information that they need to predict when the drug will fit the patient. If performance does not improve, either the drug or the engine is failing; either way, the trial should stop. If performance does keep improving, the trial can start expanding again—more clinicians can enlist and treat more of the right patients.

If the drug's numbers continue to improve, what next? One possibility is to revert to conventional blinded trials that use patient-selection criteria supplied by an engine powered by data collected up to that point. But with comprehensive tracking and reporting systems in place, a better alternative is to allow biochemists, unblinded clinicians, and Bayesian engines to continue to develop the patient-selecting biomarkers as long as the drug is used. The FDA already relies on this process to expose rare, long-term side effects that don't surface during front-end trials.

Adaptive licensing is a necessary corollary to the adaptive and open-ended development of the drug-patient science, and formalizing it would also force Washington to be more candid about the scientific realities of the drug-licensing process. When, if ever, a drug company should be able to start selling a drug for profit, and for what medical purposes, can be guided and limited by the accuracy of the constantly evolving databases and analytical engines that link known molecular effects to desired clinical effects. But who decides—and how they decide—that the engines are accurate enough to justify using a drug to treat a particular patient or disorder are not strictly scientific questions, and Washington should stop pretending that they are. As the databases grow and the analytical engines improve, the authority to

make the final calls should shift progressively from Washington to professional medical associations whose members are engaged in the battle against a disease, on down to front-line doctors and patients.

As others take charge of judging when it is in a patient's best interest to start tinkering with his own molecular chemistry, the FDA will be left with a narrower task—one much more firmly grounded in solid science. So far as efficacy is concerned, the FDA will verify the drug's ability to perform a specific biochemical task in various precisely defined molecular environments. It will evaluate drugs not as cures but as potential tools to be picked off the shelf and used carefully but flexibly, down at the molecular level, where the surgeon's scalpels and sutures can't reach. The FDA will retain the power to require that the drug be prescribed only by certain specialists and only to patients who are tested and tracked to ensure that the drug is prescribed in ways consistent with what is known about its effects. The data gathering and analytical engines used in adaptive trials can also be used to systematize the essential and rapidly expanding sphere of off-label drug prescription.

Safety is (and will forever remain) a trickier issue than efficacy. All drugs will continue to be screened at the threshold for toxicity before adaptive trials begin. Genetic factors that are linked to some fairly common side effects, such as those linked to the body's ability to metabolize a drug, have already been identified, and unblinded trials can search systematically for others. But other side effects may always be lurking just over the horizon. Some balancing between known benefits and unknown risks will always be required, and the balancing should itself be an ongoing process, as clinical experience accumulates. The best that science and regulation can do for the individual patient is provide the best possible estimates of how much confidence can be placed in the personalized prediction made by a well-designed analytical engine. If the drug is effective for some purposes and the engines are doing their job, the drug's overall performance should steadily improve as we steadily improve our ability to link both good effects and bad to patient-specific biomarkers.

Though much of Washington will recoil at the idea, we should conduct systematic comparative effectiveness studies of the regulatory process itself. However the front-end trial is scripted, one of its purposes is to establish a reliable basis for prescribing a drug safely and effectively to future patients. Conventional FDA trials provide one familiar path to that end, centered on human expertise and one specific type of statistical investigation. Adaptive trials and Bayesian analyses of large patient databases offer a different path to the same end. Those two alternatives can be tested against each other. As was done in the “Engine versus Experts” study of EuResist, there are systematic ways to find out if adaptive trials used to educate a Bayesian computer can provide better predictive guidance and provide it sooner than trials scripted by the FDA, with the details of what was learned collapsed into FDA-approved labels. Clinical experience with a drug that is widely prescribed off-label in ways later vindicated in FDA-approved clinical trials offers further opportunities to test how the Bayesian computers measure up against the empirical and analytical methods of the past.

Enlisting the Right People

What it will take to get drug companies, doctors, and patients engaged in adaptive trials is a separate question. Experience with HIV and AIDS drugs and an early adaptive trial of a Pfizer drug for acute stroke therapy indicates that patients are considerably more willing to volunteer for trials in which they are guaranteed some kind of treatment than for trials in which they take their chances on the flip of a coin.¹⁹ Drug companies and doctors, however, may hesitate to start prescribing new drugs under less tightly controlled conditions until they are confident that the data acquired will be analyzed using rigorous statistical methods, not cherry-picked in an unscientific search for anecdotes that can be used to condemn a drug at the FDA or launch lawsuits. The FDA side is easily addressed. The vaccine compensation law already provides one reasonably fair and accurate alternative to the wildly unpredictable tort system.

We also need to find reasonable ways to integrate the clinical development of drug science with the

sale of drugs for treatment. Manufacturing drugs (particularly monoclonal antibodies) in small quantities can be very expensive, and small biotechs do much of the pioneering work. Developing drugs to treat complex, slow-moving diseases will require many years of involvement by many patients. We should revive rules drafted in the HIV-driven 1980s (and still on the books) that, in appropriate circumstances, allowed manufacturers to charge patients for the cost of manufacturing drugs distributed under investigational licenses. Pay-for-performance schemes, already used in Europe, should be considered in the United States, too.

To keep private capital engaged in the long-term pursuit of ever more complex diseases, we will also need to address intellectual property rights. Much of the development cost and value of new drugs is now anchored in the development of databases that link molecular scale to clinical effects. Current patent and data-exclusivity rules address the right issues—but not broadly enough to span the continuous, dynamic process of developing the drug-patient science that we need.

That a drug trial must often begin with an imperfect molecular understanding of a disease’s biochemistry also raises a question of institutional competence. At present, the FDA passes judgment, implicitly or explicitly, on two scientifically distinct issues: a drug’s ability to control a molecule down there; and the role that the same molecule plays in causing clinical effects up here. The first obviously involves the drug. But the molecules that precisely define a disease and control its progress are matters of biological science. The FDA has quietly emerged as America’s chief taxonomist of health and disease, policing not just drug-disease interactions but also the disease-defining science and all the diagnostic and prognostic measurements used to judge whether a disease is headed north or south inside the individual patient.

But the NIH, not the FDA, is the agency with the deep expertise in diseases, and it is therefore the agency best qualified to decide when specific, measurable, molecular-scale changes in a patient’s body have some reasonable prospect of playing a

role in changing the trajectory of a disease for the better. The NIH should, at the very least, have independent authority to identify the biomarkers that can play such an important role in improving the quality of drug science and the speed at which drugs are licensed. NRC report cochair Dr. Susan Desmond-Hellmann has suggested that biomarker validation might also come from “other regulators or the American Heart Association or the American Cancer Society.”²⁰

Sooner or later, the individual doctor and patient should be added to that list. The accumulation of molecular and clinical data in public and private databases will steadily improve medicine’s ability to make an accurate, personal, biomarker-based prognosis of how the untreated disease is likely to progress inside the patient. The doctor and patient will thus gain access to concomitantly accurate estimates for how much benefit the individual patient is likely to derive from drugs that modulate molecules involved in propelling the disease. Together, the patient and doctor will then be better qualified than anyone else to decide when it makes sense to start fighting the clinical future of the disease by using one or more drugs to address molecular problems here and now.

Bayesian Statistics and Alternative States of Nature

In their basic conception, Bayesian and other “adaptive” clinical trial protocols aren’t radical or new. It was the advent of digital technology, however, that made them powerful enough to deal with the complexity of molecular medicine.

In 1948, a century after John Snow tracked cholera to the Broad Street pump and removed the handle, his successors at the NIH began searching for handles that might be removed to quell America’s rising epidemic of heart disease. They signed up 5,209 residents of the small town of Framingham, Massachusetts, to participate in a long-term study that would track their cardiovascular health and an array of possible risk factors.²¹ But the researchers faced an immediate practical problem: using

conventional statistical methods to analyze every possible combination of ten high-medium-low risk factors would have required tracking hundreds of thousands of people to get a sufficient number of representatives of each possible combination.²²

At about the same time, Jerome Cornfield, one of the NIH’s own statisticians, set about rediscovering the genius of Thomas Bayes and Pierre-Simon Laplace, the two eighteenth-century fathers of Bayes’ theorem. The one-line Bayes formula provides a systematic way to calculate “reverse probability”: how confidently we can attribute an observed effect (lung cancer or a heart attack, for example) to a suspected cause (cigarettes or high cholesterol). Cornfield’s landmark 1951 paper demonstrated how statistical methods based on that theorem could be used to establish with high confidence that most lung cancers had been caused by cigarettes. As Sharon Bertsch McGrayne recounts in her 2011 book, *The Theory That Would Not Die*, Bayes has since emerged as “arguably the most powerful mechanism ever created for processing data and knowledge.”²³

Pinning down reverse probabilities with high confidence is extremely difficult when a single effect might be the product of many causes that occur in different combinations or interact in complex ways. When conventional statistical tools are used, getting robust answers for all possible combinations of all relevant factors requires massive amounts of data. Bayesian statisticians converge on correct answers much more quickly and efficiently by adding science to the analysis in a way that progressively narrows the range of uncertainty that must be addressed by purely statistical correlations.

Richard Wilson, a Harvard professor of physics, provides a simple illustration: How believable is a child’s report that “I saw a dog running down Fifth Avenue”? To answer the question using conventional “frequentist” tools, one might conduct a study of children randomly assigned to walk a path with Fifth Avenue–like pedestrian traffic and distractions for all, but a dog briefly included only half the time. Statisticians can tell us how many children would have to be tested to arrive at a reliable measure of

how much we can trust such reports, assuming that all the factors that they can't control for—the child's eyesight, veracity, yearning for a puppy of his own, and so on—are randomly distributed among children. Dog size may be a factor, too; so if FDA statisticians were in charge, they would want a representative mix of breeds, from Great Danes to Chihuahuas. Reports of a lion sighted on Fifth Avenue would require new trials with the right mix of lions, and likewise with stegosaurus reports. The FDA could handle them all, so long as someone was willing to pay for each trial.

A Bayesian, however, would start at a different point, and arrive at reliable answers much faster. We are dealing here with a typical reverse probability problem: we have an observed effect—child chatter—and we are wondering how confidently we can attribute it to the suspect cause. But we are talking Fifth Avenue, where dogs are quite common. Accepting an “I-saw-a-lion” report requires additional information: Were the Ringling Brothers in town, and did their truck crash? “I saw a stegosaurus” is never believable, not even if Steven Spielberg is in town. The reliability of each report depends not only on the child but also on knowledge that has nothing to do with the child—knowledge about where lions roam and dinosaurs don't.

Bayes provides a systematic, rigorous way to insert that kind of external knowledge into the analysis when calculating reverse probability. Indeed, the rise of modern Bayesian analysis began with the recognition that (in McGrayne's words) “statistics should be more closely entwined with science than with mathematics.” As one Bayesian analyst put it: “The limit of [frequentist] approaches just isn't obvious until you actually have to make some decisions. You have to be able to ask, ‘What are the alternative states of nature, and how much do I believe they're true?’ [Frequentists] can't ask that question. Bayesians, on the other hand, can compare hypotheses.”

We already have good numbers for many of the alternative states of nature on Fifth Avenue, and if we didn't, we could acquire them without conducting a long series of double-blind trials. The example often used in medical textbooks addresses the use of

mammograms in the routine screening of 40-year-old women—the results have an 80-20-10 accuracy rate, with the ten being “false positive” reports of a tumor that isn't there. So these mammogram reports are, of course, wrong 97 percent of the time—29 out of every 30 frightened patients whom they send scurrying for a biopsy or some other test don't need it. If you have no idea where that “of course” came from, and don't believe it, you're in good company: surveys indicate that many American doctors don't either. But for Bayesians, this is a simple calculation.¹ Mammograms are usually wrong not because radiologists are incompetent but because breast cancer is rare—more lion than dog. When used in routine screening for rare diseases, any test that is even a bit less than perfect will report many more false positives than true positives because the number of healthy individuals screened will dwarf the number who are sick.

Cornfield helped design the Framingham heart-disease study in 1948; a decade later, it still hadn't lasted long enough, nor was it large enough, to pin down any risk factor with high confidence. But using Bayesian analysis, statisticians can refine probabilities as fast as new evidence is acquired, and in the search for rare causes, we often acquire information about suspects that do *not* matter much more quickly than information about those that do. In the first decade of the study, 92 of 1,329 adult males had experienced a heart attack or serious chest pain. Based on a Bayesian analysis of the various combinations of risk factors presented by those who had and hadn't, Cornfield was able to reframe the study around just four risk factors: cholesterol, smoking, heart abnormalities, and blood pressure. The “multiple logistic risk function” that he developed has been called “one of epidemiology's greatest methodologies.”²⁴

Using the limited amounts of data obtained during the early phases of a trial to narrow the trial's focus

¹ Among 40-year-old women picked at random, about 40 out of 10,000 have breast cancer, and 9,960 don't. Radiologists with an 80-20-10 skill for reading mammograms will detect 32 of the 40 actual cancers, and miss 8. They will also suspect cancer in 996 women who don't have it. So, of the 1028 positive calls made, 32 (about 3 percent) will be correct.

is *not* the same as concluding that we are now highly confident that we know which risk factors matter; we are just more confident than we were a while ago, we can calculate how much more, and, when the numbers look encouraging enough, we can focus more of our attention on some factors and less on others. Later results may either reinforce that early confidence boost or undermine it—the process is self-correcting. And it worked well in Framingham. Using data acquired in the early years of the study to narrow the range of what remained to be explored statistically, Cornfield hastened the arrival of statistically robust correlations that have since helped save millions of lives.

Statistics Bounded by Molecular Reality

Biologists now rely almost entirely on Bayes or closely related analytical tools to track complex diseases, or their absence, back to genetic and other molecular factors. Figures 2, 3, and 4 (below) illustrate the probabilistic causal networks that emerge when modern Bayesians use powerful computers to analyze large amounts of detailed molecular and clinical data. As those analyses demonstrate, Bayesians can begin with many suspect biomarkers, each one linked to all the others—the strength of each link initially based on such things as biochemical logic, laboratory experiments, and experience with other diseases and drugs, and therefore not much better than a guess—and then systematically adjust all the numbers until they align with all the available data on combinations of biomarkers that were present or absent in patients who did or didn't develop the disease. With enough data to analyze, the biomarkers that play no role will drop out of the picture. Those most strongly associated with the disease will, in one typical graphical representation, migrate toward the center of the graphic, closest to the point that represents the disease itself.

The same Bayesian statistical methods can be used to analyze links between unusually good health and the underlying causes—genes, for example—that keep some heavy smokers cancer-free or that allow some patients to control HIV on their own. And they

can add lung-cancer drugs and various measures of a cancer's advance or retreat to their analyses as well.

The FDA itself is often a closet—albeit an ad hoc and therefore inept—Bayesian. Sometimes there is no other ethical or practical alternative. A separate team of unblinded doctors typically monitors the results of clinical trials from a distance and can halt the trial if the results seem so clearly good or bad that continuing the trial would be unethical; the trial of AZT, the first HIV drug, ended that way. And the investigation of drug side effects that aren't bad enough to halt a trial invariably involves an ad hoc Bayesian process to identify—though not to pin down with high statistical confidence—possible side effects. The FDA's trial protocols make do, for the most part, with careful monitoring for adverse responses during the course of the trial and ad hoc searches for patterns that suggest that the drug is to blame. If the drug gets licensed, the label will warn doctors to look out for such effects, and the FDA has in place various processes for collecting reports of other side effects that the drug may have caused thereafter. But ad hoc Bayes is a far cry from the real thing, and most of the statistically robust molecular-based safety science emerges, if ever, after the drug is licensed—at which point the FDA itself uses Bayesian statistical methods to analyze the data.

Much of the efficacy side of drug science is now developed in the same way because the only practical way to learn how to treat biochemically complex diseases is to get the molecular tools into the hands of front-line doctors and let them learn about efficacy in the same way that they learn about side effects: by learning as they treat, with eyes wide open. As we have seen, the flexibility of the Orphan Drug Act and the accelerated-approval rule allows the FDA to accept limited or uneven evidence of efficacy and allows doctors to work out the algorithms and details in an ad hoc, adaptive process later on; the off-label loophole allows doctors to launch the same process without any relevant FDA-approved evidence of efficacy at all. But here, too, the ad hoc Bayesian analysis gets started late, if at all, and isn't statistically rigorous.

The NRC report includes an illustration.²⁵ Until recently, clinicians divided lung cancers into two main types: small-cell and non-small-cell. In 2003 and 2004, the FDA granted accelerated approval to two drugs (Iressa and Tarceva) on the strength of their dramatic effects in about one in ten non-small-cell patients. During the next two years, the drugs were prescribed to many patients whom they didn't help, and several follow-up clinical trials seemed to indicate that the drugs didn't work, after all—probably, we now know, “because the actual responders represented too small a proportion of the patients.”

Meanwhile, the NRC report continues, the molecular disassembly of lung cancer had begun its explosive advance. In 2004, researchers had identified the specific genetic mutation that activates the epidermal growth factor (EGF) receptor for the enzyme that these two drugs inhibit. “This led to the design of much more effective clinical trials as well as reduced treatment costs and increased treatment effectiveness.” By conditionally licensing a pair of one-in-ten drugs, the FDA had launched an adaptive process that finished the job.

Under current FDA trial protocols, however, such launches often depend on luck and circular science. The original clinical trial happens to include enough of the right patients to persuade the FDA to license the one-in-ten drug. The fortuitously and just-barely-successful completion of the FDA-approved trial starts the process that may ultimately supply the information that ideally would have been used to select the patients to include in that first trial.

Countless other valuable drugs have almost certainly been abandoned not because they didn't work but because medicine hadn't yet found out how to contract the clinical trial, while Washington's statisticians insisted on expanding it willy-nilly. According to a recent consensus report issued by a coalition of cancer experts drawn from the industry, academia, and the FDA itself, the agency still usually relies on “traditional population-based models of clinical trials ... designed to guard against bias of selection.” Such trials “may form the antithesis of personalized medicine, and accordingly, these trials

expose large numbers of patients to drugs from which they may not benefit.”²⁶

Tarceva remains on the U.S. market, but not Iressa.²⁷ In early 2005, Iressa became the first cancer drug to be withdrawn after the required follow-up trials failed to confirm its worth to the FDA's satisfaction.²⁸ In 2011, after further trials failed to establish that Iressa extends average patient survival and serious side effects surfaced, the manufacturer halted further testing in the U.S.²⁹ The drug had been licensed in Europe and other countries, subject to further study on how to identify patients whom it can help. So Iressa may yet return to the U.S., after doctors and patients in Europe and elsewhere finish developing the biomarker science that medicine needs to prescribe Iressa more precisely.

Iressa survival times and side effects vary widely among patients. As Bruce Johnson, a researcher at Boston's Dana-Farber Cancer Institute and a doctor involved in the original Iressa trials, remarked in 2005: “For us as investigators, at this point, there are at least 20 different mutations in the EGF receptors in human lung cancers, and we don't know if the same drug works as well for every mutation ... which is why we want as many EGFR inhibitor drugs available as possible for testing.”³⁰ When the FDA rescinded Iressa's license, it allowed U.S. patients already benefiting from its use to continue using it. One such patient who started on Iressa in 2004, when he had been given two to three months to live, was still alive eight years later, and walking his dogs several miles daily.

* * *

A series of frequentist drug trials can eventually yield the same answers as a single adaptive Bayesian trial: each separate trial will test a different combination of suspect causes in a suitably large number of patients, and when every combination of biomarkers has been tested, we will be statistically confident that we know how likely it is that the drug's good or bad effects can be attributed to each combination. But if the disease is biochemically complex, a great deal of time and money will be spent testing suspect causes that don't play any role.

The doctors conducting FDA-approved blinded trials have no choice. The patient-selection criteria must be specified and approved at the outset of the trial. The FDA's "controlled" trials deliberately exclude controls that unblinded doctors guided by Bayesian statisticians might otherwise develop and use to guide the inclusion of new patients and the exclusion of older ones as the trial progresses. FDA protocols do allow "subgroup analysis" of the results at the conclusion of some trials but only using statistical analyses that are heavily stacked against approving the drug.³¹

Instead, the FDA's frequentist statistical methods consign to chance everything that isn't understood and addressed at the outset and let statistical analysis take it from there. These methods assume that when a drug lands inside a human body, anything is possible, but some things are just less likely than others; assuming a specific probability-distribution for the limitless number of unknown drug-patient interactions keeps the statistics and the trials manageable. But while new drugs can surprise us in many unanticipated ways, biochemistry is not a realm in which anything is possible. How drugs and human bodies interact is constrained by solid rules of biochemical science, and we now have the power to identify those constraints, molecule by molecule, and thus narrow how much we need to rely on blind statistics.

The alternative states of nature that can affect a drug's performance are largely defined by all the biomarkers that can interact with the drug in all the different combinations that occur in patients who use it. A clinical trial of a drug that targets a biochemically complex disease will always begin with an uncertain and incomplete understanding of the drug-biomarker science. Bayesian choreographers of clinical trials can deal with many suspect biomarkers and recalculate the strength of the links among drug-biomarker clusters and various measures of the patient's health as fast as they acquire data about how different patients respond well or badly. Bayesians can likewise deal with complex, multidrug regimens from the start and continue refining them forever.

They can, for example, incorporate what science has long known—or just found out—about how

different breast-cancer or HIV molecular receptors affect a drug's performance, or about how fast cancer cells or HIV mutate at different stages of their assault on our bodies. The EuResist analytical engine takes into account the fact that three important classes of HIV drugs are used to target three different aspects of HIV's chemistry. Bayesians can start quantifying the likelihood that a new drug will perform well as soon as *any* possibly relevant biochemical information is acquired. They can begin with evidence acquired in glassware and test animals. As we shall shortly see, they can start quantifying the likelihood that a drug will successfully reach its intended target without causing side effects by considering the experience gained and biomarkers validated with other chemically similar drugs.

None of these sources of data can finish the job. But they can help launch an efficient, robust, self-correcting process that can, as it tracks a drug's effects across biochemical space and time, steadily improve our confidence in our ability to select the patients in whom the drug will perform well. Unlike the FDA, Bayesians need not select some arbitrary number of patients to be tested in a trial that will end at some arbitrary point in time. However simple or complex the disorder, the accumulation of valuable data can—and should—continue for as long as the drug is prescribed.

In the early stages of a drug trial, the negative information will be more valuable than the positive. The negative data points are the ones that allow the trial to hone in on the molecules that do matter and then stack the patient deck to increase the likelihood of a positive outcome in the next patient tested. As data accumulate, multi-patient analyses expose the patterns that can be used to understand the implications of the torrents of data extracted from a single patient, spot molecular changes that foreshadow clinical benefits or problems, and guide customized treatment.

This process will systematically converge on the science that ultimately matters: the complex, data-rich, integrated drug-patient science. As the FDA's own Dr. Janet Woodcock put it in 2004, drug science

is, at best, a “progressive reduction of uncertainty” about effects—or an “increasing level of confidence” about outcomes that comes with the development of “multidimensional” databases and “composite” measurements of outcomes.³²

* * *

The calculations required to extract cause-and-effect patterns from large volumes of complex data are so difficult that an appreciation of the full power of Bayesian analysis had to await the digital revolution. The digital wizards, as it happened, needed the power themselves; their devices and networks are constantly racing to link what matters to you right now with just the right puff of data stored somewhere in the vast digital cloud that surrounds you. Doing that efficiently is essential, which means anticipating what you want before you ask for it, which digital Bayesians do by learning from experience about the alternative states of nature commonly found in your microprocessor or brain.

Digital Bayesians can handle the rest of your body, too. Andy Grove, a founder and, for many years, pioneering CEO of Intel, has urged the FDA to catch up with the advent of computing power that “now makes possible a process that, in its early phases, enlists patients much more flexibly, to “provide insights into the factors that determine ... how individuals or subgroups respond to the drug, ... facilitate such comparisons at incredible speeds, ... quickly highlight negative results, ... [and] liberate drugs from the tyranny of the averages that characterize [FDA-scripted] trial information today.”³³ 23andMe, a provider (with Google and Genentech connections) of genetic sequencing services, recently announced that it would allow other providers and software services to develop applications that would interact with the data entrusted to 23andMe by its customers. Hundreds soon did. Their interests, *Wired* reported, include “integrating genetic data with electronic health records for studies at major research centers and ... building consumer-health applications focused on diet, nutrition and sleep.”³⁴ For individuals, 23andMe’s platform will, in the words of the company’s director of engineering, serve as “an operating system for your genome, a way that you can authorize what happens with your genome online.”

Meanwhile, Washington remains focused on why ordinary citizens should *not* be permitted to read their own biochemical scripts. The FDA is determined to protect us from reports provided by diagnostic sniffers or companies like 23andMe that, however biochemically accurate they may be, might lead to “unsupported clinical interpretations.”³⁵ But the fastest way to develop support for clinical interpretations is to do exactly what 23andMe wants to help lots of people start doing today: feed a steady stream of biochemical data into the rapidly expanding digital cloud of biochemical-clinical data, to be continuously probed by Bayesian engines, to progressively refine our understanding of all the biochemical factors that do or don’t affect clinical health.

When 23andMe and others let the rest of us catch up with the iPOPing professors at Stanford and gain easy access to the digital engines that can discern causal patterns in torrents of data, the first thing each of us should do is establish a baseline profile of our excellent health and keep it up to date thereafter. With that information securely stored and pooled with enough data from other patients, the Bayesian engines will take it from there. When we suddenly find ourselves diabetic, they will probably be able to tell us whether a viral infection, a bad diet, or some other factor was to blame. When we try a cure, we will be able to track and at least tentatively evaluate its efficacy almost immediately, down at the molecular level. To establish a control baseline for its crowd science, the FDA directs doctors to prescribe placebos. But as Stanford professor Snyder noted, the patient’s own healthy, unmedicated history can provide the best possible control for tracking a disease to its root cause, starting treatment earlier, and tracking the performance of a drug prescribed to cure it.

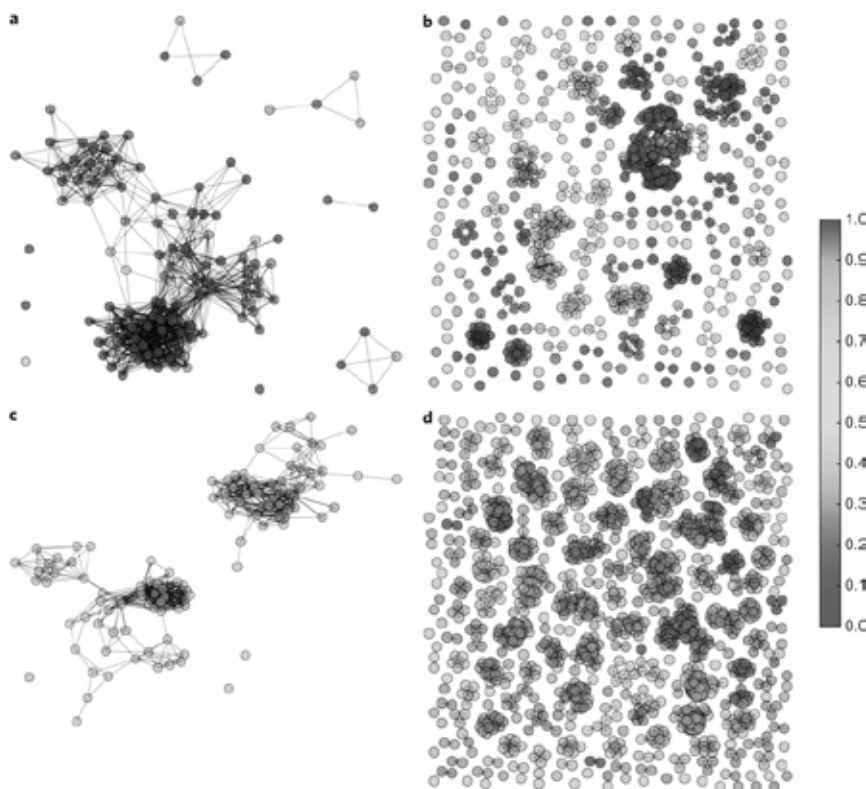
Why isn’t the FDA already on board? It accepts Bayesian methods when licensing devices—such things as lenses, implants, artificial hips, and diagnostic sniffers. In February 2010, it did finally issue a “Draft Guidance” for adaptive drug trials, and, as noted earlier, the FDA has taken a few small, hesitant steps that point to the possibility of a fundamental shift in the way it will script clinical

trials and pass judgment on drug science. But it has clearly failed to proceed at the pace that many outside experts have been advocating for years.

One of the FDA's legitimate technical concerns is, apparently, the Bayesian "prior." In deciding what to make of reports from children or radiologists, or from doctors engaged in a drug trial, Bayesian analysts require estimates of how often lions or women with breast cancer stroll down Fifth Avenue, or how strongly a suspect biomarker affects the drug's performance. These estimates can affect how quickly a Bayesian analysis will converge on a reliable answer, and drug trials must often begin with speculative estimates—too speculative, the FDA

worries—of how various biomarkers might affect a drug's performance. The FDA, however, begins with initial guesses, too—about how many patients must be tested for how long to expose enough detail about our complex biochemical diversity. The main difference is that the FDA buries its estimates in trial protocols and reductionist, unscientific pronouncements about "safe" and "effective" for the crowd. There is, of course, only one reality out there, and if the drug is prescribed to enough patients, the Bayesian and frequentist analyses of the results will invariably converge on the same understanding of how a drug's clinical effects are shaped by the various biomarker combinations presented by different patients. Without enough data, they can both make

Figure 2: Charting the "Drug-Likeness" of Different Compounds



Each dot represents a drug known to interact with one of four different targets. Tight, connected clusters indicate chemical similarity, and the darkest dots represent the most promising "drug-like" compounds. Biochemical properties determine whether a drug is likely to reach its intended target (bioavailability) or to cause unwanted side effects. By comparing molecular structures of different drugs, drug designers can use clinical experience with the old drugs to predict how the new drug will perform.

Source: G. Richard Bickerton et al., "Quantifying the Chemical Beauty of Drugs," *Nature Chemistry* 4 (January 24, 2012): 90–98

mistakes. Because they are willing and able to deal with much more data, the Bayesians will correct their mistakes faster.

Bayesian analyses look messy, mainly because they dare to deal forthrightly with complexity. But Bayesians don't choke when biochemical reality gets complex, either. "Far better an approximate answer to the right question," as one Bayesian put it, "than an exact answer to the wrong question."³⁶

The New Science of Molecular Crowds

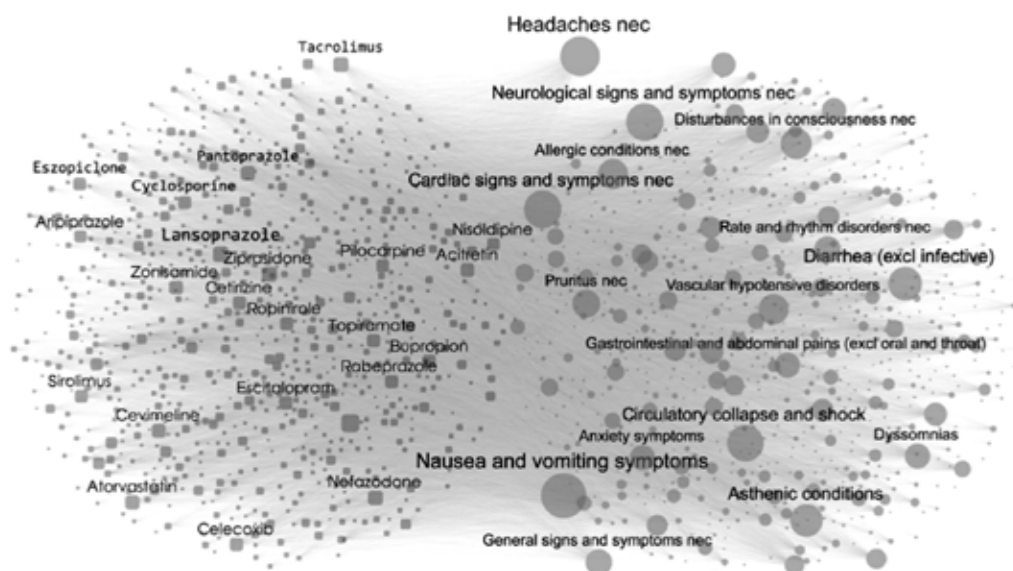
By combining what we know about drugs with what we know about bodies, researchers are already beginning to systematize pharmacological science from end to end. The predictive power of integrated drug-patient science is rapidly moving far beyond anything that pharmacology has previously seen.

In their 2012 paper "Quantifying the Chemical Beauty of Drugs," one research team describes how it pooled information about multiple aspects of the

molecular structures of drugs successfully licensed in the past to arrive at a general algorithm for predicting the likelihood that a candidate drug will be successfully absorbed by the human body and won't have toxic side effects.³⁷ The team used similar tools to quantify the beauty of potential binding sites that a new drug might attempt to target. To the eyes of a biochemist, the measure of a drug's beauty is how likely it is to hook up smoothly with its target.

Another research group combined a catalog of 809 drugs and the 852 side effects known in 2005 with information about each drug's chemical properties and molecular targets in the human body. Network analysis software was then able to predict almost half of the additional side effects that have emerged since then. "We were pleasantly surprised," said Ben Reis, director of the predictive medicine group at Boston Children's Hospital. Part of the network's power comes from the inclusion of information not previously considered in attempts to assess side-effect risks—the drug's molecular weight and melting point, for example, and what specific part of the body

Figure 3: Predicting Adverse Drug Events Using Pharmacological Network Models



Clinical experience links drugs to known side effects (on the right); biochemical properties link drugs to other drugs (on left). The network then reveals new links between drugs and side effects.

Source: A. Cami et al., "Predicting Adverse Drug Events Using Pharmacological Network Models," *Science Translational Medicine* 3, no. 114 (December 21, 2011): 114–27, <http://stm.sciencemag.org/content/3/114/114ra127.abstract>

the drug targets. As Reis notes: “The network encodes a lot of information from other worlds.” The team is now investigating what types of biochemical data have the most predictive value and is studying drug-drug interactions. “We’re moving from a paradigm of detection—where it takes sick people to know something is wrong—to prediction.”³⁸

By mining ten years’ of clinicians’ notes on the treatment of 4,700 patients at a large psychiatric hospital, another team uncovered some 800 unexpected pairings of health problems.³⁹ Adding gene and protein data relevant to about 100 of these pairs revealed previously unknown molecular connections between such conditions as migraines, hair loss, gluten allergy, and schizophrenia. Yet

another team developed what one member describes as an opposites-attract dating service for drugs and diseases.⁴⁰ Using public databases that contain thousands of genomic studies, the digital matchmaker searches for diseases that push a specific human biochemical north and drugs that push it south. Early results suggest that an epilepsy drug might also be useful in treating certain inflammatory bowel disorders, while an ulcer drug might also help treat some forms of lung cancer.

Biochemists call this “repurposing.” Many more such odd couples are certainly out there waiting for us. Life has been repurposing molecules from the beginning, so we now find identical or very similar ones scattered all over the place. And when we find them, we find

Figure 4: Using Electronic Patient Records to Discover Disease Correlations



Overlapping or adjacent conditions such as diabetes (number 26 at top) and hypertension (number 72) suggest a likely molecular kinship. Comparing patient biochemical profiles can then reveal the specific biochemical processes involved.

Source: Francisco S. Roque et al., “Using Electronic Patient Records to Discover Disease Correlations and Stratify Patient Cohorts,” *PLOS Computational Biology* 7, no. 8 (2011), <http://www.ploscompbiol.org/article/info%3Adoi%2F10.1371%2Fjournal.pcbi.1002141>

new commonalities among diseases that point to new uses for old drugs.

This is the new science of crowds: crowds defined not by shared clinical symptoms but by shared clusters of molecules that propel our diseases and interact with our drugs. Life is intrinsically social. Nucleic acids and the proteins that they define, along with all the rest of the chemistry that proteins assemble and control, flow through the river of life as surely as cholera once flowed through London's water supply. Mapping human chemistry exposes the differences between many forms of breast cancer—but it also exposes the shared biochemistry of leukemia and gastrointestinal cancer; and, by way of a drug called pentamidine, the molecular kinship of sleeping sickness in Gambia and fungal pneumonia in AIDS patients; and, by way of thalidomide, insomnia, leprosy, Kaposi's sarcoma (a rare form of skin cancer) and "wasting syndrome" in AIDS patients, and at least two bone-marrow and blood cancers in other patients; and unsightly facial hair and sleeping sickness again (eflornithine). There are differences everywhere, but there are also matches, overlaps, and widely shared forms of molecular strength and weakness.

Going forward, molecular science will link the performance of more drugs to more genes, proteins,

and other molecular constituents of the sick and healthy parts of our bodies. The more we learn, the easier it will get to learn still more, and the cheaper it will get to translate what we know into powerful medicine. The accumulation of increasingly detailed descriptions of how biochemical ecosystems work will progressively lower the cost of designing new drugs and determining, quickly and cheaply, how they can be prescribed safely and effectively.

Databases will expand to include the results of laboratory experiments on microbes and test animals and thus expose biochemical webs and processes that operate in the same way in different species. No bacterium, rat, monkey, or other animal is a good model for an entire person, but some are enough like us in the ways that are relevant to beating a particular disease, and they are now bioengineered to incorporate human—and sometimes the individual patient's—immune-system or cancerous-tumor genes. A good animal model for a human disease is often what launches the ultimately successful search for a drug to beat the disease in people. Extended across species, molecular cartography can thus steadily improve our ability to develop solid human-drug science outside human bodies.

ENDNOTES

1. For a recent review, see Raymond L. Woosley, "One Hundred Years of Drug Regulation: Where Do We Go From Here?," *Annual Review of Pharmacology and Toxicology* 53 (2013): 255–73.
2. Jim Manzi, *Uncontrolled: The Surprising Payoff of Trial-and-Error for Business, Politics, and Society* (New York: Basic Books, 2012), 77; and Curtis L. Meinert, *Clinical Trials: Design, Conduct, and Analysis* (New York: Oxford University Press, 1986), 93.
3. Rock Brynner and Trent Stephens, *Dark Remedy: The Impact of Thalidomide and Its Revival as a Vital Medicine* (New York: Basic Books, 2001), 122–23 (citing interviews conducted in 2000 with Dr. Gilla Kaplan and Dr. Robert Hastings). Brynner and Stephens describe the patient as a man; an earlier source, however, expressly notes that the patient was female: G. P. Ellis and G. B. West, eds., *Progress in Medicinal Chemistry* (Elsevier, Amsterdam 1985), 22:192.
4. "FDA Approves Important New Leukemia Drug," National Cancer Institute at the National Institutes of Health (May 10, 2001), <http://www.cancer.gov/newscenter/pressreleases/2001/gleevecpressrelease>.
5. National Academy of Science, "Disarming a Deadly Virus: Proteases and Their Inhibitors," *Beyond Discovery* (February 2000), <http://www.beyonddiscovery.org/content/view.page.asp?l=111>.
6. Rui Chen et al., "Personal Omics Profiling Reveals Dynamic Molecular and Medical Phenotypes," *Cell* 148, no. 6 (March 16, 2012): 1293–1307; Allison Blass, "Genetics in Action: Researcher Watches His Own Body Develop Diabetes," *Diabetes Mine* (April 26, 2012), <http://www.diabetesmine.com/2012/04/genetics-in-action-researcher-watches-his-own-body-develop-diabetes.html>; and Krista Conger, "Revolution in Personal Medicine: First-Ever Integrative 'Omics' Profile Lets Scientist Discover, Track His Diabetes Onset," Stanford School of Medicine (March 15, 2012), <http://med.stanford.edu/ism/2012/march/snyder.html>.
7. For reviews of many of the major issues, see Malorye Allison, "Can Cancer Clinical Trials Be Fixed?," *Nature Biotechnology* 29 (2011): 13–15; and idem, "Reinventing Clinical Trials," *Nature Biotechnology* 30, no. 1 (January 9, 2012): 41–49.
8. Walter Armstrong, "Pharma's Orphans," *Pharmaceutical Executive* (May 1, 2010), <http://www.pharmexec.com/pharmexec/article/articleDetail.jsp?id=670568&sk=&date=&pageID=2>; President's Council of Advisors on Science and Technology, "Report to the President on Propelling Innovation in Drug Discovery, Development, and Evaluation," iv.
9. Michael A. Friedman (acting commissioner of the FDA), "Statement on Implementation of the FDA Modernization Act of 1997," testimony to the Committee on Commerce, House of Representatives (October 7, 1998), <http://www.fda.gov/NewsEvents/Testimony/ucm115096.htm>; and Ken Flieger, "FDA Finds New Ways to Speed Treatment to Patients," FDA Consumer Special Report on New Drug Development in the United States (January 1995), <https://sites.google.com/site/pharmafuture/fdaconsumerspecialreport>.

10. EuResist, "EuResist Prediction System," <http://engine.euresist.org>.
11. Renee Twombly, "Estrogen's Dual Nature? Studies Highlight Effects on Breast Cancer," *Journal of the National Cancer Institute* 103, no. 12 (June 22, 2001): 920–21.
12. Kate Kelland, "Study Suggests Breast Cancer Is Clutch of 10 Diseases," Reuters (April 18, 2012), <http://www.reuters.com/article/2012/04/18/us-cancer-breast-subtypes-idUSBRE83H10E20120418>; and Christina Curtis et al., "The Genomic and Transcriptomic Architecture of 2,000 Breast Tumours Reveals Novel Subgroups," *Nature* 486 (June 21, 2012): 346–52.
13. "IBM Watson Hard at Work: New Breakthroughs Transform Quality Care for Patients," IBM press release (February 8, 2013), <http://www-03.ibm.com/press/us/en/pressrelease/40335.wss>.
14. A. C. Lionel et al., "Rare Copy Number Variation Discovery and Cross-Disorder Comparisons Identify Risk Genes for ADHD," *Science Translational Medicine* 3, no. 95 (August 10, 2011); and Tina Hesman Saey, "Shared Differences: The Architecture of Our Genomes Is Anything but Basic," *ScienceNews* 175, no. 9 (April 25, 2009): 16–20.
15. Tina Hesman Saey, "Uncommon Carriers: People Have a Surprising Number of Rare Genetic Variants," *ScienceNews* 182, no. 4 (August 25, 2012) : 28–29; Jacob A. Tennessen et al., "Evolution and Functional Impact of Rare Coding Variation from Deep Sequencing of Human Exomes," *Science* 337, no. 6090 (July 6, 2012): 64–69; and Matthew R. Nelson et al., "An Abundance of Rare Functional Variants in 202 Drug Target Genes Sequenced in 14,002 People," *Science* 337, no. 6090 (July 6, 2012): 100–104.
16. Janet Woodcock (acting deputy commissioner for operations), "A Framework for Biomarker and Surrogate Endpoint Use in Drug Development," PowerPoint presentation (November 4, 2004, slides 8, 36, http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CDQQFjAA&url=http%3A%2F%2Fwww.fda.gov%2Fohrt%2Fdoctors%2Ffac%2F04%2Fslides%2F2004-4079S2_03_Woodcock.ppt&ei=V7HRULD1BYWs0AGKyoGIBw&usg=AFQjCNFHWKjhcWus_aEazAH1ME9mNBRhBg&sig2=rBM1OaKD4K4eHypRyR-zcg&bvm=bv.1355534169,d.dmQ).
17. National Research Council, Committee on a Framework for Developing a New Taxonomy of Disease, *Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease* (Washington, D.C.: National Academies Press, 2011), 1.
18. For one proposal along those lines, see Anup Malani, Oliver Bembom, and Mark van der Laan, "Accounting for Differences among Patients in the FDA Approval Process," John M. Olin Law & Economics Working Paper No. 488, Public Law and Legal Theory Working Paper No. 281 (October 2009), 3, 8, <http://www.law.uchicago.edu/files/file/488-281-am-fda.pdf>.
19. Derek Lowe, "What You Need to Know about Adaptive Trials," *Pharmaceutical Executive* (July 1, 2006), <http://www.pharmexec.com/pharmexec/article/articleDetail.jsp?id=352793&sk=&date=&pageID=4>.
20. Quoted in Amy Dockser Marcus, "Working Toward a New Social Contract for Health Care," *Wall Street Journal* (April 17, 2012), <http://blogs.wsj.com/health/2012/04/17/working-toward-a-new-social-contract-for-health-care>.

21. William B. Kannel et al., "An Investigation of Coronary Heart Disease in Families," *American Journal of Epidemiology* 110, no. 3 (January 31, 1979): 281–90.
22. Sharon Bertsch McGrayne, *The Theory That Would Not Die: How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines, and Emerged Triumphant from Two Centuries of Controversy* (New Haven, Conn.: Yale University Press, 2011), 115.
23. *Ibid.*, 224.
24. McGrayne, *The Theory That Would Not Die*, 115–16, 238.
25. National Research Council, *Toward Precision Medicine*, 26.
26. Samir N. Khleif et al., "AACR-FDA-NCI Cancer Biomarkers Collaborative Consensus Report: Advancing the Use of Biomarkers in Drug Development," *Clinical Cancer Research* 16, no. 13 (July 1, 2010): 3299
<http://clincancerres.aacrjournals.org/content/16/13/3299.full>
27. Renee Twombly, "Failing Survival Advantage in Crucial Trial, Future of Iressa Is in Jeopardy," *Journal of the National Cancer Institute* 97, no. 4 (February 16, 2005): 249–50.
28. Simon Frantz, "Iressa Failure Raises Fears about Accelerated Approvals," *Nature Reviews Drug Discovery* 4, no. 94 (February 2005); and "FDA Alert for Healthcare Professionals: Gefitinib (marketed as Iressa)," U.S. FDA website, <http://www.fda.gov/downloads/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/ucm126182.pdf>.
29. Ed Silverman, "AstraZeneca Finally Gives Up on Iressa Approval," *Pharmalot*, February 9, 2011, <http://www.pharmalive.com/astrazeneca-finally-gives-iressa-approval>.
30. Quoted in Twombly, "Failing Survival," 250.
31. For a statistician's view of the arcane details, see Malani et al., "Accounting for Differences," 3, 8.
32. Woodcock, "A Framework for Biomarker and Surrogate Endpoint Use in Drug Development."
33. Andrew Grove, "Rethinking Clinical Trials," *Science* 333, no. 6050 (September 23, 2011): 1679.
34. Daniela Hernandez, "Social Codes: Sharing Your Genes Online," *Wired: Science* (November 9, 2012).
35. Letter from Alberto Gutierrez, Office of *In Vitro* Diagnostic Device Evaluation and Safety, Center for Device and Radiological Health, to Anne Wojcicki, president and cofounder, 23andMe Personal Genome Service, "RE: 23andMe Personal Genome Service" (June 10, 2010), <http://www.fda.gov/downloads/MedicalDevices/ResourcesforYou/Industry/UCM215240.pdf>; and David Gelles, "FDA Calls for Genetic Tests Crackdown," *Financial Times* (June 13, 2010), <http://www.ft.com/cms/s/0/ed476afc-7715-11df-ba79-00144feabdc0.html#axzz2IHhmcNa8>.

36. Quoted in McGrayne, *The Theory That Would Not Die*, 233.
37. G. Richard Bickerton et al., "Quantifying the Chemical Beauty of Drugs," *Nature Chemistry* 4 (January 24, 2012): 90–98; and Rachel Ehrenberg, "Measuring What Makes a Medicine," *ScienceNews* 181, no. 5 (March 10, 2012): 18.
38. Rachel Ehrenberg, "Network Analysis Predicts Drug Side Effects," *ScienceNews* 181, no. 2 (January 28, 2012): 12; and Aurel Cami et al., "Predicting Adverse Drug Events Using Pharmacological Network Models," *Science Translational Medicine* 3, no. 114 (December 21, 2011): 114ra127.
39. Francisco S. Roque et al., "Using Electronic Patient Records to Discover Disease Correlations and Stratify Patient Cohorts," *PLOS Computational Biology* 7, no. 8 (2011): e1002141; and Rachel Ehrenberg, "Mining Electronic Records Yields Connections between Diseases," *ScienceNews* 180, no. 8 (October 8, 2011): 16.
40. Joel T. Dudley et al., "Computational Repositioning of the Anticonvulsant Topiramate for Inflammatory Bowel Disease," *Science Translational Medicine* 3, no. 96 (August 17, 2011): 96ra76; Marina Sirota et al., "Discovery and Preclinical Validation of Drug Indications Using Compendia of Public Gene Expression Data," *Science Translational Medicine* 3, no. 96 (August 17, 2011): 96ra77; Amy Dockser Marcus, "Researchers Show Gains in Finding Reusable Drugs," *Wall Street Journal* (August 18, 2011).

CHAIRMAN

Andrew von Eschenbach, MD

COMMITTEE MEMBERS

Arthur Daemmrach, PhD
Joseph DiMasi, PhD
Henry G. Grabowski, PhD
Paul Howard, PhD
Peter Huber, PhD
Daniel P. Petrylak, MD
Tomas J. Philipson, PhD
Lance K. Stell, PhD
E. Fuller Torrey, MD

PROJECT FDA

Sponsored by the Manhattan Institute

Utilizing twenty-first century technologies to help develop better FDA regulations and a faster and safer drug and medical-device pipeline

Project FDA is a Manhattan Institute initiative which aims to reform the FDA to meet twenty-first century challenges. Under the leadership of former FDA commissioner Dr. Andrew von Eschenbach, Project FDA promotes reforms that can enable the FDA to offer a more predictable, transparent, and efficient pathway for bringing safe and effective new products to patients.

Medicine is on the cusp of a radical transformation. New sciences and technologies are poised to allow physicians to personalize treatment for every cancer patient; arrest or prevent the development of Alzheimer's disease; and radically lower health care costs by reducing the prevalence of expensive chronic diseases. Unfortunately, today's FDA—simultaneously overtasked and underfunded by Congress—has struggled to adapt its regulations to new scientific advances.

Project FDA believes the FDA can become a bridge for innovation, rather than a barrier to it, and that this can be achieved without sacrificing patient safety. For instance, advances in molecular medicine that allow companies to target specific sub-groups of patients, combined with electronic health records, should allow the FDA to streamline and improve time-consuming and expensive pre-market product testing that can take a decade or more, and implement vigorous post-market surveillance of “real world” patients after drugs or devices demonstrate safety and efficacy in early testing. This approach will not only accelerate access to innovative products, it should enhance efforts to safeguard public health.

Project FDA will educate the public on the FDA's vital role in advancing medical innovation; highlight the potential for new sciences to improve health while also lowering costs; and collaborate with patients groups, industry stakeholders, and policymakers to modernize the FDA's policies and procedures.

As part of our mission to advance public dialogue and debate about the importance of supporting medical innovation and personalized medicine, the Center for Medical Progress also hosts Medical Progress Today, a blog that provides a forum for economists, scientists, and policy experts to explore the scientific, regulatory, and market frameworks that will best support twenty-first century medical innovation.

The Manhattan Institute is a 501(C)(3) nonprofit organization. Contributions are tax-deductible to the fullest extent of the law. EIN #13-2912529

<http://www.manhattan-institute.org/projectfda>