Despite high efficacies, precision medicines face the same problems as other drugs: high attrition rates and acquired resistance. Targeting small patient subgroups also makes it hard to offset the high price of bringing a drug to market. Development programs must use all information relevant to bioactive compounds, biological systems and diseases to counter these issues. Frameworks that efficiently support data-driven drug development should integrate sources and reveal connections to ensure better understanding and use of data.
The term “personalized medicine” entered mainstream rhetoric enveloped in the hype around the genomic revolution and the promise of treatment tailored to individual patient needs. The idea of “a drug designed for me” grew out of the prediction that rapidly decreasing DNA sequencing costs would spur mass production of genomic data that could be correlated to information on the demographics, medical histories and drug responses of entire populations (1). Sophisticated computer algorithms would then analyze trends and connections among tens of thousands of patients to deduce relevant treatment options for an individual based on similarities in genomic profiles.

The reality of personalized medicine looks quite a bit different. Instead of being tailored to the individual, personalized medicine hinges on delivering a dedicated drug to a precisely defined subgroup of patients. It extracts a biologically defined subgroup from the overall patient population based on a biomarker that predicts a positive response to therapy, and then focuses the clinical development of a candidate drug for that subgroup. In light of this discrepancy between promise and reality, the U.S. National Research Council introduced a more accurate term: precision medicine (2).

Many precision medicines developed to date have demonstrated astounding efficacies and are encouraging therapy alternatives for medical areas that have limited treatment options, such as non-small cell lung cancer. However, the current business model of precision medicine is likely unsustainable into the future. The average out-of-pocket cost of bringing a drug to market is 1.4 billion USD (3), and targeting the specific needs of patient subgroups, though effective, means a smaller base to recover development costs. Something needs to change.

Not all Patients are Alike

There has been a shift in the pharmaceutical industry from a blockbuster approach to drug development with the assumption that “one drug fits all” to the paradigm of precision medicine that aims to deliver “the right drug to the right group of patients.” Key to this shift was the realization that low efficacies of cancer drugs might be attributed to the heterogeneity of tested patient populations. That is, biologically driven differences among patients diagnosed with one type of cancer were essentially diluting the strong therapeutic effect that a drug might have for a specific subgroup of that population (4). Precision medicine emerged as a more meaningful and potentially more successful way of bringing new drugs on to the market (Figure 1). Information about patient traits that impact response to a drug presented an opportunity to dodge the high attrition rates that have plagued the pharmaceutical industry for decades.

The outcome of this shift is noticeable. Efficacies of newly approved precision medicines like crizotinib and ceritinib are exceptionally high (>50% response rate), which enabled fast-tracked approvals so that patients could benefit from their development as soon as possible (5).
Drug Development Sidesteps Biological Complexity

Most drugs these days are still designed using a reductionist approach that isolates the interaction between a therapeutic compound and a single biological molecule. Drugs are, however, ultimately administered to organisms where molecules operate in a complex orchestration of DNA, RNA, proteins, steroids, metabolites and more. Experience has shown that redundancy and plasticity built into networks of biological molecules can eliminate the impact of a drug on a single target and even render pathogens or tumors resistant to medication (6).

Precision medicines are not exempt from the effects of biological complexity. Unlike most blockbuster drugs, development of precision medicines enhances the narrow characterization of a lead and its impact on a chosen target with data on patient characteristics that influence response to treatment. Thus, input informing development captures knowledge about the drug target and the benefiting patients, but still ignores information about the environment in which the drug exerts its effect, namely, biological systems.

Properties of biological systems that have thwarted the efficacy of promising medications have also rendered precision medicines unusable. In many cases, patients respond to a treatment for only a short time before developing drug resistance. Non-small cell lung cancer patients, for example, respond initially very well to the therapies gefitinib and erlotinib, but almost invariably develop resistance after 10–12 months of treatment (7).

Pharmaceutical companies typically respond by producing second-generation drugs. A surprising number of different mechanisms underlie acquired resistance to cancer therapies, mostly genetic mutations, and these are targeted in development strategies. In this way, lessons learned from first-generation drugs and genomic data from patients inform the development of new therapies.

Second-generation drugs are rapidly entering the clinical setting for patients with resistant tumors, but these new drugs target individual genetic mutations linked to the observed resistance. This means that the resulting drugs benefit individual subgroups of an already stratified patient population (i.e., only those diagnosed with a particular cancer and who have the culprit mutation). With this narrower indication scope, it becomes increasingly difficult for pharmaceutical companies to recoup the expense of developing these targeted therapies. The very essence of precision medicine—efficacious drugs for specific subgroups of the patient population—may lead drug development into a scientific and financial cul-de-sac.

Accumulating records of acquired resistance and continuous discovery of yet another underlying mechanism, raises the question whether such an “arms race” against a never-ending series of acquired resistances is defensible. Furthermore, with a susceptibility window of only a few months for each drug, is it even possible to develop enough drugs to address all forms of resistance and meet the needs of all patient subgroups?

Figure 1. The evolution of drug development models, from “blockbuster” to precision medicine and into the future. Each progressive step is informed by increasing amounts of data, first from a datastream with information on disease and drug mechanisms, then from a second datastream with information on patient characterization. Looking into the future, complexity of data informing drug design and development will increase as focus shifts from single targets to complete biological networks.
Properties of biological systems that have thwarted the efficacy of promising medications have also rendered precision medicines unusable.

Process-Wide Repercussions of a Narrowing Scope

Creating second-generation drugs that address each mechanism of resistance in isolation means that each new drug is relevant for only a subset of patients treated with the first-generation drug. Drawing this logic into subsequent rounds of acquired resistances leads to the conclusion that every new defensive line of drugs will be indicated for smaller and smaller patient subgroups. This trend is not restricted to acquired resistance. Different molecular aberrations implicated in lung and breast cancer that are targeted by approved and candidate drugs also show decreasing frequency in patient populations (Figure 2).

The repercussions of this narrowing in patient subgroups play out at multiple stages in the drug development pipeline. For one, the number of candidates entering and moving through the pipeline will need to grow substantially. This puts pressure on discovery programs to identify more, increasingly specific leads. Then, with attrition rates of 80–90% common throughout industry, numerous candidates will need to be processed so that sufficient therapies successfully make it to market. Given the predominance of late-stage failures, pushing more candidates through the pipeline could mean a wildly disproportionate R&D investment.

Clinical trials are also impacted. If the trend of shrinking patient subgroups holds, limiting test populations in clinical trials to only subjects predicted to benefit from a drug will make it increasingly difficult to find the number of patients needed to demonstrate drug efficacy and safety. So far, this has not been an issue because precision medicines have had tremendous efficacy rates. But in the future, clinical utility of precision medicines with lower efficacies will be less clear cut and demand larger sample sizes to demonstrate statistical significance. As a result, trials will take longer and cost more.

Raising the price tag on medications to recover development costs, however efficacious the drugs may be, is simply not realistic. The only option to escape the cul-de-sac created by current development approaches for precision medicines is to broaden the utility scope of existing and new drugs and to cut back development costs. Only then will continued investment in precision medicine be warranted.
The only option to escape the cul-de-sac created by current development approaches for precision medicines is to broaden the utility scope of existing and new drugs and to cut back development costs.

Networks Open New Paths
Sifting through the hype around precision medicine, we must acknowledge that successes rest on the precise selection of patients to receive treatment, and little has changed about the way drugs are designed. These are single-target drugs matched to specific patient subgroups. That is, precision medicines are an improvement over other medications only because they are precisely indicated for the right patients, and not because they are created to tackle diseases with innovative modes of action.

The resistance mechanisms observed with various precision cancer treatments highlight the fallacy of assuming that cancer can be treated by knocking out a single aberrant molecule. First, cell function is based on the activity of signaling pathways that are highly redundant, interconnected networks of multiple relevant molecules. Shutting down one aberrant pathway causing tumor growth may lead to compensatory activity in another. Second, the genetic profile of a tumor changes over time. New aberrations in tumor cells may render a target molecule unsusceptible to a drug.

Opening a path out of the scientific and financial dead end predicted for precision medicine will require placing the complexity of drug action and the dynamic nature of disease at the heart of development strategies. Drug discovery, design, optimization and clinical evaluation must be guided by as much information as possible to ensure that a drug candidate is vetted to work in a full-blown organism. Drug development must be data-driven, not target-driven.

Precision medicines that deliver a powerful therapeutic effect while simultaneously circumventing acquired resistance or off-target effects will be drugs designed to modulate complete networks of interacting biological molecules, rather than inhibiting or promoting the activity of a single target. Before any assay to discover active compounds, a development team must piece together as comprehensive a picture of relevant biological pathways based on a wide range of existing data and data to be produced. They could then leverage the promiscuity of compounds to “hit” several critical components of a network. Alternatively, they could combine several.

Figure 2. Frequency of mutations targeted by approved and candidate drugs to treat breast and lung cancer. Over time, the size of patient groups that benefit from approved and future cancer drugs decreases, as the frequency of target aberrations is lower. Percentages of mutations in genes for EGFR, KRAS, MET, BRAF and PIK3 represent the average of the highest reported frequency in different ethnic groups. Data extracted from Jørgensen 2011 (4) and El-Telbany & Ma 2012 (8).
single-target drugs into an effective therapy. For that purpose, they need a clear-cut understanding of how the drugs interact with one another (9).

Precision medicines also need a maximized indication scope. This may require abandoning traditional definitions of disease and focusing instead on the biological pathways involved in generating a disease phenotype. So, for example, cancer may no longer be defined by the affected organ, but by the biological pathways that are dysfunctional in tumor tissue. Again here, large amounts of data must feed into constructing a far deeper and expanded understanding of the biological activity responsible for a diseased phenotype, how mechanisms may be interconnected, and how microenvironmental conditions can modulate disease and drug action.

Network biology and network chemistry aim to capture the complexity of biological systems, how they responding to bioactive compounds, and how compounds interact with one another. Applied to early stages of drug development, network-based approaches can revolutionize the definition of indications requiring attention, and the discovery and vetting of leads that are suitable to tackle therapeutic needs. Applied to clinical stages, network-based approaches can unify trial objectives so that multiple drugs can be tested within the same infrastructure and immediate feedback can inform development priorities.

Efforts to create and implement network-based approaches are underway. Initiatives in bioinformatics are creating ways to visualize, evaluate and interpret large amounts of data, and feed resulting insights into drug design and optimization. Isolated drug-target interactions are being collected into larger models of complete signaling pathways, and researchers are working out methodologies to analyze and interpret the vast bodies of genomic, epigenomic, transcriptomic, proteomic, phosphoproteomic, and metabolic data available in multiple scientific repositories (10).

In situations where knowledge is lacking about how components in a biological system are connected and interact, network-based mathematical strategies enable data-driven extraction of patterns and properties from multimodal datasets that can serve as basis for modeling. For example, weighted correlation network analysis, can assist in summarizing clusters of molecules with linked functions and reveal new signaling pathways that may be relevant to a disease but would have otherwise gone unnoticed because connections were not obvious from looking at individual system components (11).

Clinical trial experts are also exploring network-based approaches that promote the use of data to inform decisions. The I-SPY-2 TRIAL (Investigation of Serial Studies to Predict Your Therapeutic Response with Imaging and Molecular Analysis 2) is an innovative clinical trial that can be viewed as a collection of several trials unified by a shared infrastructure and the integration of information derived from all treatments. A profile of a patient’s breast tumor is made upon entering the trial. This profile is used to match the patient to one of several therapeutic agents being tested. The selected agent is the one that has demonstrated best efficacy for the patient’s tumor based on results from patients already enrolled in the trial. That is, the trial allows a continuously optimized scheme for assigning new patients to treatment groups.
Concurrently, results from the different treatments are compared to remove tested drugs that prove to be ineffective or toxic, and new drugs are added to the trial as new treatment options.

The shared infrastructure has advantages over conventional trial design. First, comparing the multiple treatments informs decisions earlier about whether a particular therapeutic agent should continue to be pursued. Second, the extensive profiles of patient tumors provide data to support new indications for a drug if it demonstrates efficacy associated with a different subgroup of patients than originally planned. Third, data shared across several treatment groups can help overcome limitations due to small sample sizes arising from patient selection. Fourth, collected data are used to improve the disease models underlying the design and optimization of each therapeutic drug tested. The latter is a critical feedback loop to inform drug development with real and relevant clinical data, and to generate a more comprehensive and accurate picture of overall disease mechanisms.

**Building Information Frameworks for Data-Driven Drug Development**

Dr. Herbert Köppen, a veteran of the pharmaceutical industry who advocates integrating network biology approaches into drug development processes, points out that “implementing those tools and ensuring that they are a good proxy of the human body will require intensive work. The technologies are already available, but the workflows need to be developed and then integrated into the framework of pharmaceutical research.”

Along the same lines, Dr. Scott Lusher, an expert in IT solutions for data-driven scientific research, explains that integrating these tools will require a change in mentality. “Medicinal chemists, biologists, pharmacologists and all other team members will need to work with, share, trust and interpret data differently.”

Drug developers will need to explore and interpret a highly complex picture of correlated molecular targets interacting with multiple therapeutic compounds in a heterogeneous and changing physiological environment to extract key insights relevant to a drug’s design. They will need to operate at the intercept of a growing number of data sources, experimental methods and disciplines (12).

Coupled to this change in mentality and the implementation of new methodologies is the need to build information frameworks that support data-driven approaches. Such frameworks must allow integrating data from multiple sources and accurately informing drug development with knowledge emerging from the interplay of multiple scientific disciplines. Just like network biology and network chemistry focus on understanding properties that emerge from networks and not individual network components, these frameworks should reveal insights from the complete information landscape that are not visible when restricted to the perspective of a single knowledge domain.

To be effective, three elements must be foundational pillars of any framework designed for data-driven drug development. First, the data used must be high-quality. Data must be an accurate reflection of what is reported in literature, internal information systems, and other sources. Data should be qualified with information about how and why they were collected so that they are used only in ways that match the context in which they were generated. Equally important are integrated mechanisms to assess the quality of data and exclude content that does not meet quality criteria. The use of improperly curated data can mislead drug development, lengthen research projects, cause costly mistakes and, in worst cases, undermine patient safety.

Integrating different types of information from different sources requires processing channels that bring disparate data together and highlight synergies where connections and comparisons are possible. This brings us to the second pillar: the frameworks should build on organizational constructs such as taxonomies, ontologies and dictionaries. Such constructs enable translating unstructured or semi-structured data into structured data that are findable and
A data-driven approach, built on these cross-disciplinary information frameworks, will loosen the restrictive path of precision medicine and highlight ways to make medicines effective but not limited.

Usable in meaningful ways. Practically, they are the source of indexing used to tag and retrieve data, and they inform production steps that normalize and standardize data entering a framework. Functionally, they serve as the basis to organize data and thus, underlie connections and synergies that emerge from the overall body of information.

The third pillar is that the organizational constructs of these frameworks must reflect an accurate and comprehensive perspective of all disciplines relevant to data-driven drug discovery—disciplines that enable a precise understanding of disease biology, that elucidate patient heterogeneity and that explain genomic, proteomic, metagenomic, histological, disease history, patient environment and other data. To remain accurate and comprehensive, the organizational constructs will need to be flexible, accommodating not only expansion of knowledge but also shifts in paradigms, debunking of previous understanding, and a very large vocabulary that continues to grow and evolve. This cross-disciplinary view of a highly dynamic knowledge landscape serves as the context of data-driven drug development and thus, is an imperative to make sense of data. Only data that are interpreted correctly can accurately inform daily decisions about a compound, the biological theater in which it exerts its effect, and the disease it aims to address. And only making the best decisions at every development stage can improve the way that drugs are made.

For decades, experts in the organization and dissemination of scientific knowledge have been creating information frameworks that facilitate merging knowledge and uncovering the insights that result from the whole. These frameworks support the exploration of new ideas and grant access to the data needed to pursue a new investigative path. A data-driven approach, built on these cross-disciplinary information frameworks, will loosen the restrictive path of precision medicine and highlight ways to make medicines effective but not limited. Ultimately, development teams leveraging such information frameworks will be able to translate large amounts of correlated, diverse and often disparate data into insights that will drive the creation of not just more precise, but better drugs.
REFERENCES


R&D Solutions for Pharma & Life Sciences

Elsevier’s R&D Solutions is a portfolio of tools that integrate data, analytics and technology capabilities to help pharmaceutical & life sciences customers realize the potential of data-driven drug discovery and development.

LEARN MORE

To learn more about Elsevier’s R&D Solutions, visit elsevier.com/rd-solutions/pharma-and-life-sciences or contact your nearest Elsevier office.

ASIA AND AUSTRALIA
Tel: +65 6349 0222
Email: sginfo@elsevier.com

JAPAN
Tel: +81 3 5561 5034
Email: jpinfo@elsevier.com

KOREA AND TAIWAN
Tel: +82 2 6714 3000
Email: krinfo.corp@elsevier.com

EUROPE, MIDDLE EAST AND AFRICA
Tel: +31 20 485 3767
Email: nlinfo@elsevier.com

NORTH AMERICA, CENTRAL AMERICA AND CANADA
Tel: +1 888 615 4500
Email: usinfo@elsevier.com

SOUTH AMERICA
Tel: +55 21 3970 9300
Email: brinfo@elsevier.com