Title
Leveraging innovative technology to generate drug response phenotypes for the advancement of biomarker-driven precision dosing.

Permalink
https://escholarship.org/uc/item/0rw277kj

Journal
Clinical and translational science, 14(3)

ISSN
1752-8054

Authors
Oni-Orisan, Akinyemi
Srinivas, Nithya
Mehta, Krina
et al.

Publication Date
2021-05-01

DOI
10.1111/cts.12973

Peer reviewed
POSITION PAPER

Leveraging innovative technology to generate drug response phenotypes for the advancement of biomarker-driven precision dosing

Akinyemi Oni-Orisan1,2,3 | Nithya Srinivas4 | Krina Mehta5 | Jesmin Lohy Das6 | Thu T. Nguyen7 | Geoffrey H. Tison8 | Scott R. Bauer9,10,11 | Maria Burian12 | Ryan S. Funk13 | Richard A. Graham14 | Biomarkers and Translational Tools Community Working Group of the American Society for Clinical Pharmacology and Therapeutics

Abstract

Although traditional approaches to biomarker discovery have elucidated key molecular markers that have improved drug selection (precision medicine), the discovery of biomarkers that inform optimal dose selection (precision dosing) continues to be a challenge in many therapeutic areas. Larger and more diverse study populations are necessary to discover additional biomarkers that provide the resolution needed for a more tailored dose. To generate and accommodate large datasets of drug response phenotypes, time- and cost-efficient strategies are necessary. In particular, a multitude of technological advances that originated for purposes outside of biomedical research (electronic health records, direct-to-consumer genetic testing, social media, mobile devices, and machine learning) have made it easier to communicate, connect, and gather information from consumers. Although these technologies have been used...
INTRODUCTION

Precision medicine is the practice of tailoring a treatment that is optimal for an individual patient. The objective of precision medicine is to maximize efficacy and minimize toxicity without the need for trial-and-error. Precision dosing, which is the selection of the right dose for a patient, is a critical component of precision medicine with the same objective. The successful implementation of precision dosing requires robust biomarkers that can predict response to various pharmacologic doses. Biomarkers have facilitated precision medicine for decades. Well before the Human Genome Project was completed, clinicians were using individual clinical, demographic, and laboratory biomarkers to determine an appropriate pharmacological therapy. Traditional sources of drug information that facilitated the discovery of these early biomarkers were randomized controlled trials (RCTs) and observational studies. The

landmark Framingham Heart Study, for example, produced the Framingham Risk Score, a composite biomarker tool for coronary heart disease risk prediction used to inform the decision for and intensity of treatment with lipid-lowering therapies. Importantly, the development of biomarker tools for precision medicine have been driven by technology (see Figure 1 for an example in cardiovascular therapeutics). With the advancement of biomolecular technology, our enhanced ability to characterize a patient’s molecular profile (genome, proteome, metabolome, etc.) has allowed for improved discovery of biomarkers beyond clinical, demographic, and laboratory predictors. Based on subsequent data supporting clinical validity (strength of association between biomarker and drug response phenotype), many of these molecular markers are progressing toward clinical implementation. Although the majority of these “low-hanging fruits” (biomarker-phenotype associations with large effect sizes) can inform the “right drug for the right patient,” many do not have

with success in the health sciences for an array of purposes, these resources have not been fully capitalized on for precision dosing. This perspective will touch on how these innovations can be used as data sources, data collection tools, and data processing tools for drug-response phenotypes with a unique focus on advancing biomarker-driven precision dosing.

FIGURE 1 Timeline for technological innovation and the identification of select clinical biomarkers in cardiovascular therapeutics. Shown are select biomarker developments (top half) and accompanying technological innovations (bottom half) for cardiovascular therapeutics. Similar progress has been made in other disease states. With the advancement of biomolecular technology, our enhanced ability to characterize a patient’s molecular profile (genome, proteome, metabolome, etc.) has allowed for improved discovery of biomarkers beyond clinical, demographic, and laboratory predictors. In order to best leverage these advances for precision dosing, we need creative strategies for gathering drug response phenotypes from data with large, diverse study populations in a time- and cost-efficient manner. AST, aspartate transaminase; CK, creatine kinase; DNA, deoxyribonucleic acid; LDH, lactate dehydrogenase; PCR, polymerase chain reaction.
sufficient resolution to inform the “right dose.” This is, in part, due to limitations in current drug response phenotypes from the aforementioned traditional sources of drug information. In order to maximize the utility of molecular biomarkers for precision dosing, we need drug response phenotypes from larger and more diverse study populations (e.g., from real-world data sources). Per the US Food and Drug Administration (FDA), real-world data are “the data relating to patient health status and/or the delivery of health care routinely collected” (www.fda.gov/ce-research/science-and-research-special-topics/real-world-evidence). This multidisciplinary perspective on behalf of the American Society for Clinical Pharmacology & Therapeutics (ASCP) Biomarkers & Translational Tools (BTT) community will discuss innovative approaches, mostly adapted from domains outside of biomedical research, for gathering and accommodating drug phenotypic data from real-world data sources in a time- and cost-efficient manner. We will touch on how these innovations can be used as data sources (electronic health records, direct-to-consumer genetic testing, and social media), data collection tools (mobile devices), and data processing tools (machine learning) for drug-response phenotypes with the ultimate objective of advancing precision dosing.

TRADITIONAL DATA SOURCES

Traditionally, retrospective substudies of RCTs and observational studies have been used to generate drug response phenotypes for biomarker studies. These have been successful in establishing the clinical validity of candidate biomarkers. The majority of these biomarker success stories have been in pharmacogenetics. For example, compelling evidence for the clinical validity of CYP2C19 as a biomarker of clopidogrel efficacy was generated from a genetic substudy of the phase III TRITON-TIMI 38 trial. Additionally, clinical validity of HLA-B*5801 genotype for allopurinol-induced severe cutaneous adverse reactions was generated largely from observational case-control studies. The association between HLA-B*5701 genotype and risk of abacavir hypersensitivity reactions was first observed in two independent observational studies from 2002. These gene-drug pair examples have the highest level of evidence for clinical implementation and are among the best examples of biomarker-driven precision medicine. In contrast, for precision dosing, some examples exist, but further advances are necessary. For instance, current guidelines recommend alternative antiplatelet therapy in CYP2C19 poor metabolizers for whom clopidogrel is indicated. However, variation in CYP2C19 metabolizer status (poor, intermediate, extensive, and ultrarapid) suggests that clopidogrel dose optimization may be a viable therapeutic strategy in addition to alternative therapies. Indeed, evidence from an RCT demonstrated that tripling maintenance dose (to 225 mg daily) in intermediate metabolizers achieved similar platelet reactivity reduction as that achieved from standard doses in normal metabolizers. For poor metabolizers, however, adequate platelet reactivity reduction was not achieved despite a broad range of doses (75, 150, 225, and 300 mg daily). Furthermore, the trial was small (333 patients) and thus did not have the power to investigate clinical outcomes. A follow-up study to determine the precise dosing requirements necessary for adequate platelet reactivity reduction in poor metabolizers would require a broader range of doses and adequate sample size in each dosing group (overall and within each CYP2C19 metabolizer type). Thus, whereas these examples highlight the advances made toward precision medicine using traditional approaches, they currently demonstrate the need for large study populations and diverse dose-response phenotypes in order to reach the resolution required for precision dosing biomarker studies. Innovative approaches that can gather phenotypic data in an efficient manner are necessary. In pursuit of this goal, we highlight several relatively new and emerging innovations. Excellent comprehensive reviews of these technologies, including the broad benefits, potential challenges, and further applications, have been covered elsewhere; this unique perspective is focused on their promising application to precision dosing.

ELECTRONIC HEALTH RECORD-LINKED BIOBANKS

Biobanks are biorepositories that store biospecimens primarily for research purposes. Electronic health records (EHRs) are intended for the documentation of clinical care delivery and billing in health care systems. Large biobanks integrated with EHRs as long-term resources for research studies address some of the limitations of the aforementioned traditional approaches in biomarker discovery and development for precision dosing. EHR-linked biobanks are more cost-effective than biomarker substudies of RCTs because a single biobank can be used for multiple studies across many disease states and therapies. Moreover, the cost of generating drug response phenotypes is reduced to the minimal expense of extracting data already in place from routine clinical care and billing. Another advantage of EHR-linked biobanks is the real-world nature of this data source; participants are less limited by strict inclusion/exclusion criteria allowing for a demographically heterogeneous sample more generalizable to the true population of patients receiving a particular therapy. Importantly, as drug response phenotypes represent patient outcomes in the course of routine clinical care, patients will be treated with a variety of drugs and doses, thereby enhancing variation in therapeutic regimens and facilitating precision dosing. Thus, EHR-linked biobanks can identify biomarkers of dose selection for drugs with a large range of approved doses. This advantage of EHRs can also present as a challenge: the lack of uniformity between
participants in recorded dose, frequency, and timing of response for a particular drug may lead to doubt in the robustness of a drug response phenotype. However, large sample size and rigorous study design can overcome this potential shortcoming. For example, a high-resolution dose response for statin-induced low-density lipoprotein cholesterol (LDL-C) lowering was demonstrated using EHRs in a pattern similar to what has been reported in clinical trials (Figure 2). This is despite ostensibly messy and highly heterogenous data: unlike a clinical trial, the drug type, drug dose, timing of LDL-C measurements relative to statin initiation, year of statin initiation (spanning almost 2 decades of EHR records), and patient characteristics were highly variable. Adjustment for confounders, imputation to handle missing data, a defined daily dose (DDD) algorithm to account for LDL-C lowering potency differences across statin type, and other methods were necessary to provide a reliable phenotype. Furthermore, the large study population (e.g., \( N > 10,000 \)), broad dose range covered, and statin LDL-lowering variability within each dose group also facilitates precision dosing: the phenotype could be interrogated for biomarkers that may predict differential drug response by dose. In general, molecular biomarker studies derived from EHR-linked biobanks have demonstrated significant success in the discovery of genetic markers associated with various disease outcomes, such as cancer and cardiovascular disease. Although less common, there are also successful examples related to drug response using this resource that can serve as models for further studies. An additional consideration is that the large majority of these examples are genetic biobanks; an important next step for precision dosing is the utility of EHR-linked biobanks for the discovery of circulating biomarkers of drug response.

**DIRECT-TO-CONSUMER GENETIC TESTING**

Declining genotyping costs in tandem with heightened public interest in genomics has resulted in high demand for direct-to-consumer genetic testing, another resource with promise for precision dosing. There are a number of direct-to-consumer genetic testing companies. Perhaps the best-known is 23andMe, which is the only company with FDA approval for its genetic tests. The 23andMe data consist of genomewide genotypes generated from deoxyribonucleic acid collected with home-based saliva kits linked to phenotypes collected through self-reported surveys. The company serves individual consumers by providing personal ancestry and health information and also serves third parties (drug companies and academic researchers) by granting access to their data. Together, this large-scale collection of genetic and phenotypic data represent a disruptive and enabling innovation in the conduct of genetic health research. The concept of self-reported data has generated doubt among health scientists about the accuracy of phenotypes using this approach. However, the replication of numerous previously established genetic associations has blunted skepticism. Consequently, novel variants of genetic susceptibility to a number of health conditions have been identified. Recently, the company licensed its first drug compound developed from its own consumer data: a bispecific antibody targeting the IL-36 cytokine family for the treatment of psoriasis. This major milestone illustrates the promise of the direct-to-consumer genetic testing model in discovering biomarkers as pharmacotherapeutic targets. Considering the large sample size (~ 10 million) and easy-to-deploy surveys for phenotyping, the potential utility for advancing precision dosing with postapproval drugs is also enormous. More attention directed at generating drug and dose response phenotypes from consumer self-reported data through additional survey questions would provide additional avenues for biomarker discovery and personalized medicine.

**SOCIAL MEDIA**

Although social media data is not commonly perceived as a tool to generate drug response phenotypes for precision...
MOBILE DEVICES

Mobile devices, such as smartphones, smartwatches, fitness bands, and other “wearables,” are increasingly used by consumers and produce a wealth of data that can be harnessed to advance precision dosing. Data from mobile devices can be collected using both passive and active user input. Passive input can be used individually or in combination to track health-related behaviors, such as physical activity and smoking, or estimate physiologic signs, such as heart-rate and sleep.22 Active input is used to collect patient-reported outcomes, such as mood or stress.22 Passive input entail minimal engagement from the patient, but require investigators to use complex analytics to distinguish signal from noise and identify clinically meaningful phenotypes. Conversely, active input requires engagement from the user, but allows for customization of which outcomes are most important for a particular patient or study. Importantly, data output from mobile devices can be responsive to drug therapy at varying doses. Furthermore, the feasibility (e.g., in-home setting, repeated measures) and user satisfaction of these devices allow for the collection of multiple data points from many users. These features enable for the resolution necessary to parse out the effects of varying doses in the generation of precision dosing phenotypes using mobile devices. Another application of mobile devices that would be of particular benefit for precision dosing is within the context of N-of-1 trials, which are multiple-crossover trials conducted within a single individual.23 N-of-1 trials can serve as precision medicine tools for clinical research or routine clinical care to optimize treatment for an individual. Although N-of-1 trials have had slow uptake due to investigator and participant burden of data collection, a mobile device used in this setting could help facilitate the efficiency in reporting outcomes (e.g., an electronic diary) and receiving instructions from the study coordinator. Similar to social media, privacy and security are also challenges to data generated from mobile devices and require attention, but are beyond the scope of this paper. Active collaborations between researchers and the companies that produce these devices may be needed to address challenges and leverage the promise of mobile devices for precision dosing.

MACHINE LEARNING

The aforementioned novel technologies for precision dosing would presumably generate a large volume of rich, high-dimensional raw phenotypic data that pose difficulties for traditional algorithmic and statistical approaches without significant data summarization, often based on a priori assumptions, in order to function. This requires leaving out a substantial amount of data from the discovery process, which may decrease predictive power and/or fail to capture biologic mechanisms beyond those prespecified. In recent years, improvements in algorithms used for data analysis have been largely driven by advances in machine learning, a subcategory within artificial intelligence.24 These algorithms have not only fostered performance improvements in many domains, but, in some cases, also offered entirely new approaches to analyze large-scale data and make powerful predictions that were previously not possible. These strengths of machine learning lend themselves to biomarker-driven precision dosing. The predominance of machine learning applications for healthcare thus far have largely been comprised of supervised learning, whereby machine learning models are “trained” using data that has been labeled by human experts for the task of interest. However, unsupervised learning, which finds patterns within data without predefined labels, may have particular utility toward the application of biomarker discovery from large amounts of data across various modalities in precision dosing. Whether supervised or unsupervised
TABLE 1 Executive summary: achieving precision dosing with technologies used to generate phenotypic data from real-world study populations

<table>
<thead>
<tr>
<th>Technology</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>EHR-linked biobanks:</td>
<td>The potential to generate robust drug response phenotypes that incorporate a variety of drug regimens is a unique feature of EHR-linked biobanks that has not been used to its full potential.</td>
</tr>
<tr>
<td>Direct-to-consumer genetic testing:</td>
<td>This resource provides an opportunity to create drug and dose response phenotypes via consumer self-reported data from millions of patients through easy-to-deploy surveys.</td>
</tr>
<tr>
<td>Social media:</td>
<td>There is evidence demonstrating patient willingness to link data from health-specific social media platforms with EHR records as promising phenotypes for precision dosing.</td>
</tr>
<tr>
<td>Mobile devices:</td>
<td>Studies investigating how physiologic processes and patient-reported outcomes are impacted by drugs using this technology have been conducted in N-of-1 trials: expansion into larger-scale studies with multiple drug regimens is a logical next step.</td>
</tr>
<tr>
<td>Machine learning:</td>
<td>Machine learning provides the much-needed analytic counterpart applicable to the big data drug response phenotypes constructed from the aforementioned approaches.</td>
</tr>
</tbody>
</table>

Abbreviation: EHR, electronic health record.

learning models are implemented, one limitation of machine learning is that results are biased to the underlying data. For example, a machine learning-generated finding that a biomarker is associated with drug efficacy at a particular dose may not be generalizable to populations that are not represented in the data source that was used. This caveat underscores the need for human involvement in the clinical interpretation of any results generated from machine learning algorithms. Ultimately, machine learning approaches provide the much-needed analytic counterpart that is applicable to various novel technologies. If the challenges are adequately met, there is great potential of this approach to advance biomarker-driven precision medicine in the coming decades.

CONCLUSION

A number of cutting-edge technologies from domains outside of science have been used successfully to advance biomedical research, but have not been capitalized on for precision dosing. There are no reports that have provided insight into this topic. On behalf of the ASCPT BTT community, and with the input of experts from outside of clinical pharmacology, we provided a multidisciplinary perspective of how these technologies can be used as data sources (EHRs, direct-to-consumer genetic testing, and social media), data collection tools (mobile devices), and data processing tools (machine learning) to generate phenotypic data from large real-world study populations in a matter that advances precision dosing beyond traditional approaches (Table 1). An initial future next step is to provide a foundation for the continued use of these technologies for precision dosing through proof-of-concept examples. Collaboration between the private and public sector may be necessary for this key step.

CONFLICT OF INTEREST

All authors declared no competing interests for this work.

REFERENCES


