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

Akinyemi Oni-Orisan[1,2,3] | Nithya Srinivas[4] | Krina Mehta[5] | Jesmin Lohy Das[6] | Thu T. Nguyen[7] | Geoffrey H. Tison[8] | Scott R. Bauer[9,10,11] | Maria Burian[12] | Ryan S. Funk[13] | Richard A. Graham[14] | Biomarkers and Translational Tools Community Working Group of the American Society for Clinical Pharmacology and Therapeutics

[1]Department of Clinical Pharmacy, University of California, San Francisco, California, USA

[2]Institute for Human Genetics, University of California, San Francisco, California, USA

[3]Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, California, USA

[4]Incyte Corporation, Wilmington, Delaware, USA

[5]qPharmetra, Cary, North Carolina, USA

[6]Janssen Pharmaceuticals, Raritan, New Jersey, USA

[7]Department of Family and Community Medicine, University of California, San Francisco, California, USA

[8]Division of Cardiology, University of California, San Francisco, California, USA

[9]Department of Medicine, University of California, San Francisco, California, USA

[10]Department of Urology, University of California, San Francisco, California, USA

[11]Veterans Affairs Medical Center, San Francisco, California, USA

[12]Translational Medicine Neuroscience, UCB Biopharma SRL, Braine-l'Alleud, Belgium

[13]Department of Pharmacy Practice, University of Kansas Medical Center, Kansas City, Kansas, USA

[14]Theravance Biopharma, South San Francisco, California, USA

**Correspondence**

Akinyemi Oni-Orisan, Institute for Human Genetics, University of California, San Francisco, 513 Parnassus Ave. S966B, San Francisco, CA 94143-0794, USA.
Email Akinyemi.Oni-Orisan@ucsf.edu

## 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

LEVERAGING INNOVATIVE TECHNOLOGY TO GENERATE DRUG RESPONSE PHENOTYPES FOR
THE ADVANCEMENT OF BIOMARKER-DRIVEN PRECISION DOSING

785

> 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.

## 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.[1] 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.[2] 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
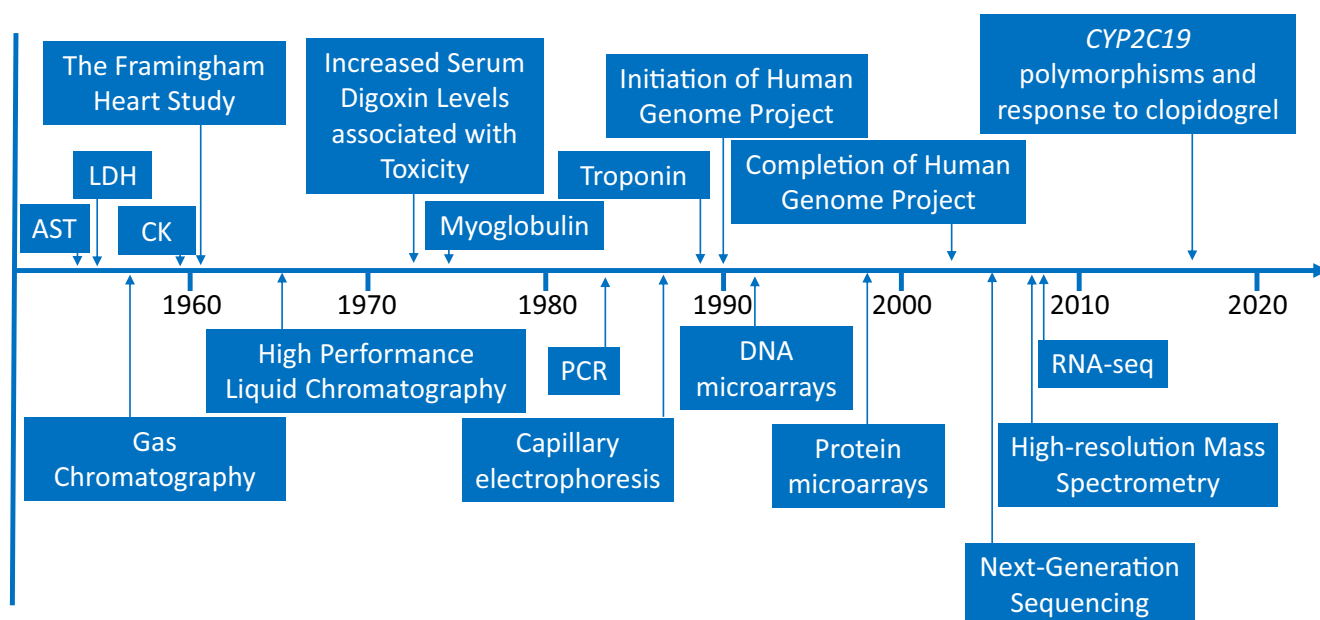


**FIGURE 1** Timeline for technological innovation and the identification of select clinical biomarkers in cardiovascular therapeutics. Shown are select biomarker developments (top half) and accompanying technological innovations (bottom half) for cardiovascular therapeutics. Similar progress has been 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/scien ce-research/science-and-research-special-topics/real-world -evidence). This multidisciplinary perspective on behalf of the American Society for Clinical Pharmacology & Therapeutics (ASCPT) 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.
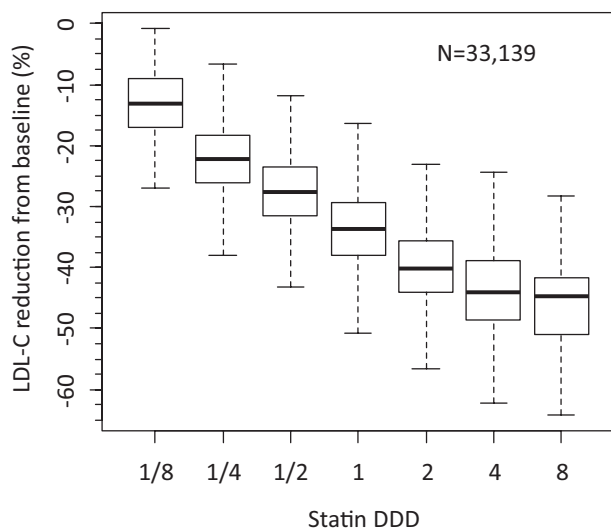
## TRADITIONAL DATA SOURCES

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.[2] Additionally, clinical validity of *HLA-B*5801* genotype for allopurinol-induced severe cutaneous adverse reactions was generated largely from observational case-control studies.[3] The association between *HLA-B*5701* genotype and risk of abacavir hypersensitivity reactions was first observed in two independent observational studies from 2002.[3] 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 exists,[4] but further advances are necessary. For instance, current guidelines recommend alternative antiplatelet therapy in CYP2C19 poor metabolizers for whom clopidogrel is indicated.[2] However, variation in CYP2C19 metabolizer status (poor, intermediate, extensive, and ultrarapid[2]) 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.[5] 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 concurrently 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[6-10]; 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.[6] 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

LEVERAGING INNOVATIVE TECHNOLOGY TO GENERATE DRUG RESPONSE PHENOTYPES FOR
THE ADVANCEMENT OF BIOMARKER-DRIVEN PRECISION DOSING

787

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).[11] 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.[12] Although less common, there are also successful examples related to drug response using this resource that can serve as models for further studies.[13] 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,[14] another resource with promise for precision dosing. There are a number of direct-to-consumer genetic testing companies.[14] 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.[14] 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.[14] 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.[15] Consequently, novel variants of genetic susceptibility to a number of health conditions have been identified.[16] 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.[17] 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.



**FIGURE 2** Statin low-density lipoprotein cholesterol (LDL-C) dose response curve generated from electronic health record (EHR)-linked biobank data. Despite highly heterogenous data from the EHRs of a real-world population, a dose response phenotype (with resolution at the level of each incremental dose change) was generated that replicated findings previously reported in randomized controlled trials (RCTs). Large sample size ($N = 33,139$) and rigorous statistical methodology was necessary to produce this robust phenotype. This phenotype is useful for precision dosing because a biomarker (from biobank data linked to the EHR) found to be associated with statin LDL-C lowering response at different doses could inform the best dose to prescribe for a patient. Reproduced from reference 7 (authors of published articles from American Heart Association journals may reuse figures without requesting permission). DDD, defined daily dose.

## SOCIAL MEDIA

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

dosing, platforms, such as Facebook, Twitter, and Instagram, have been used to answer a variety of questions in health research.[18] Social media is increasing in popularity; Twitter has over 300 million active users globally as far back as 2015.[18] Moreover, social media data can be used to collect timely information compared with traditional data sources due to the real-time nature of the content. In addition, social media can also be used as a platform for interventions with targeted messages for specific audiences or as a recruitment tool for studies.[18] Despite these remarkable advantages, social media remains an untapped resource for biomarker-driven precision dosing research. A few challenges must be addressed for this to become a reality. First, researchers need to explore creative approaches for characterizing drug response using social media data. For example, there have been multiple techniques used to extract publicly reported adverse drug reactions for pharmacovigilance studies.[18] Data collected from the social media platform PatientsLikeMe also could be a valuable tool; participants share their experience (effectiveness, side effects, adherence, and burden) with over 3500 prescription, over-the-counter, and supplement therapies for a variety of diseases.[19] Second, privacy and security of health information are always a concern in the digital age; platforms can become Health Insurance Portability and Accountability Act(HIPAA) compliant on top of usual safeguards for additional protection. Third, significant effort needs to be put toward linking social media data with external sources (such as EHR-linked biobanks) for molecular biomarker studies. Fortunately, evidence suggests that patients are willing to link their social media data with EHRs[20] and projects, such as Sync for Science, can facilitate this process. Moreover, crowdsourcing through social media platforms has already identified a biomarker for disease.[21] This same approach could be applied for drug response biomarkers.

## 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

LEVERAGING INNOVATIVE TECHNOLOGY TO GENERATE DRUG RESPONSE PHENOTYPES FOR
THE ADVANCEMENT OF BIOMARKER-DRIVEN PRECISION DOSING

789

**TABLE 1** Executive summary: achieving precision dosing with technologies used to generate phenotypic data from real-world study populations

| |
| --- |
| EHR-linked biobanks: |
| 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 |
| Direct-to-consumer genetic testing: |
| 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 |
| Social media: |
| There is evidence demonstrating patient willingness to link data from health-specific social media platforms with EHR records as promising phenotypes for precision dosing |
| Mobile devices: |
| 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 |
| Machine learning: |
| Machine learning provides the much-needed analytic counterpart applicable to the big data drug response phenotypes constructed from the aforementioned approaches |

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.[25] 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 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

1. Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults Executive summary of the third report of the national cholesterol education program (NCEP) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (adult treatment panel III). *JAMA*. 2001;285:2486-2497.
2. Scott SA, Sangkuhl K, Stein CM, et al. Clinical Pharmacogenetics Implementation Consortium guidelines for CYP2C19 genotype and clopidogrel therapy: 2013 update. *Clin Pharmacol Ther*. 2013;94:317-323.
3. Martin MA, Klein TE, Dong BJ, et al. Clinical pharmacogenetics implementation consortium guidelines for HLA-B genotype and abacavir dosing. *Clin Pharmacol Ther*. 2012;91:734-738.
4. Relling MV, Schwab M, Whirl-Carrillo M, et al. Clinical pharmacogenetics implementation consortium guideline for thiopurine dosing based on TPMT and NUDT15 genotypes: 2018 update. *Clin Pharmacol Ther*. 2019;105:1095-1105.
5. Mega JL, Hochholzer W, Frelinger AL, et al. Dosing clopidogrel based on CYP2C19 genotype and the effect on platelet reactivity in patients with stable cardiovascular disease. *JAMA*. 2011;306:2221-2228.
6. Bowton E, Field JR, Wang S, et al. Biobanks and electronic medical records: enabling cost-effective research. *Sci Transl Med*. 2014;6:234cm3.
7. Horton R, Gillian C, Lindsey F, et al. Direct-to-consumer genetic testing. *BMJ*. 2019;367:l5688.
8. George DR, Rovniak LS, Kraschnewski JL. Dangers and opportunities for social media in medicine. *Clin Obstet Gynecol*. 2013;56:453-462.
9. Steinhubl SR, Muse ED, Topol EJ. The emerging field of mobile health. *Sci Transl Med*. 2015;7:283rv3.
10. Obermeyer Z, Emanuel EJ. Predicting the future – big data, machine learning, and clinical medicine. *N Engl J Med*. 2016;375:1216-1219.
11. Oni-Orisan A, Hoffmann TJ, Ranatunga D, et al. Characterization of statin low-density lipoprotein cholesterol dose-response using electronic health records in a large population-based cohort. *Circ Genom Precis Med*. 2018;11:e002043.
12. Hoffmann TJ, Passarelli MN, Graff RE, et al. Genome-wide association study of prostate-specific antigen levels identifies novel loci independent of prostate cancer. *Nat Commun*. 2017;8:14248.
13. McCarty CA, Wilke RA. Biobanking and pharmacogenomics. *Pharmacogenomics*. 2010;11:637-641.
14. Khan R, Mittelman D. Consumer genomics will change your life, whether you get tested or not. *Genome Biol*. 2018;19:120.
15. Tung JY, Do CB, Hinds DA, et al. Efficient replication of over 180 genetic associations with self-reported medical data. *PLoS One*. 2011;6:e23473.

16. Nalls MA, Blauwendraat C, Vallerga CL, et al. Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet Neurol*. 2019;18:1091-1102.

17. Abbasi J. 23andme develops first drug compound using consumer data. *JAMA*. 2020;323:916.

18. Sinnenberg L, Buttenheim AM, Padrez K, Mancheno C, Ungar L, Merchant RM. Twitter as a tool for health research: a systematic review. *Am J Public Health*. 2017;107:e1-e8.

19. Wicks P, Mack Thorley E, Simacek K, Curran C, Emmas C. Scaling PatientsLikeMe via a "generalized platform" for members with chronic illness: web-based survey study of benefits arising. *J Med Internet Res*. 2018;20:e175.

20. Padrez KA, Ungar L, Schwartz HA, et al. Linking social media and medical record data: a study of adults presenting to an academic, urban emergency department. *BMJ Qual Saf*. 2016;25:414-423.

21. Chancellor MB, Bartolone SN, Veerecke A, Lamb LE. Crowdsourcing disease biomarker discovery research; the IP4IC study. *J Urol*. 2018;199(5):1344-1350.

22. Burke LE, Ma J, Azar KMJ, et al. Current science on consumer use of mobile health for cardiovascular disease prevention: a scientific statement from the American Heart Association. *Circulation*. 2015;132:1157-1213.

23. Kronish IM, Alcántara C, Duer-Hefele J, et al. Patients and primary care providers identify opportunities for personalized (N-of-1) trials in the mobile health era. *J Clin Epidemiol*. 2017;89: 236-237.

24. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521:436-444.

25. Gianfrancesco MA, Tamang S, Yazdany J, Schmajuk G. Potential biases in machine learning algorithms using electronic health record data. *JAMA Intern Med*. 2018;178:1544-1547.