An Overview of Genomic Biomarker Use in Cardiovascular Disease Clinical Trials

Oluseyi Adeniyi[1,*], Anuradha Ramamoorthy[1], Robert Schuck[1], Jielin Sun[1], Jennifer Wilson[2,3], Issam Zineh[1] and Michael Pacanowski[1]

Clinical trial designs targeting patient subgroups with certain genetic characteristics may enhance the efficiency of developing drugs for cardiovascular disease (CVD). To evaluate the extent to which genetic knowledge translates to the CVD pipeline, we analyzed how genomic biomarkers are utilized in trials. Phase II and III trial protocols for investigational new drugs for CVD and risk factors were evaluated for prospective and exploratory genomic biomarker use; drug targets were evaluated for the presence of evidence that genetic variations can impact CVD risk or drug response. We identified 134 programs (73 unique drug targets) and 147 clinical trials. Less than 1% (n = 1/147) trials used a genomic biomarker prospectively for in-trial enrichment despite 32% (n = 23/73) of the drug targets having evidence of genetic variations. Additionally, 46% (n = 68/147) of the trials specified exploratory biomarker use. The results highlight an opportunity for more targeted CVD drug development by leveraging genomic biomarker knowledge.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?
☑ Prior studies have reported the decline of cardiovascular drug development and highlighted the role of biomarker-driven strategies to enhance drug development.

WHAT QUESTION DID THIS STUDY ADDRESS?
☑ This study examined the extent to which known genetic markers for cardiovascular disease (CVD) are applied in clinical trials of novel drugs for CVD and its risk factors.

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?
☑ Prospective use of genomic biomarkers in CVD clinical trials is minimal even when drug targets in development are

supported by evidence that genetic variants are linked to the CVD trait for which they are being developed. Exploratory genomic biomarker studies are relatively prevalent.

HOW MIGHT THIS CHANGE CLINICAL PHARMACOLOGY OR TRANSLATIONAL SCIENCE?
☑ These findings highlight a gap in clinical trials for CVD and the potential for knowledge of genetic variation to play a role in enhancing drug development for CVD.

Drug development for cardiovascular disease (CVD), the leading cause of morbidity and mortality in the United States and globally,[1] is widely reported to be in decline.[2–6] The decline has been attributed to multiple factors, including a paucity of novel therapeutic targets,[7] clinical trial failures due to lack of efficacy and safety,[2,8] and concerns of commercial viability.[7,9] Pharmaceutical pipeline evaluation across therapeutic areas has shown that drug targets validated through genetic association with the disease of interest were more likely to be successful in phase II clinical trials[10] and across successive phases of development.[11] In addition, use of biomarkers to molecularly select patients for clinical trial

enrollment has been associated with significantly higher probability of transition from phase I clinical trials to approval.[12] These findings suggest that using biomarker-based drug development strategies may enhance the efficiency of clinical trials and has the potential to stimulate drug development for CVD.

The ability of prognostic and predictive genomic biomarkers to enhance drug development has been most widely appreciated in oncology, where knowledge of genetic variants has aided the prediction of which drug candidates are likely to be efficacious in targeted subsets of patients. In addition, genomic biomarkers have aided the selection of patients who are likely to benefit from

[1]Office of Clinical Pharmacology, Office of Translational Sciences, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, Maryland, USA; [2]ORISE Fellow, Office of Clinical Pharmacology, Office of Translational Sciences, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, Maryland, USA; [3]Department of Bioengineering, Stanford University, Palo Alto, California, USA. *Correspondence: Oluseyi Adeniyi (oluseyi.adeniyi@fda.hhs.gov)

Table 1 Distribution of exploratory biomarker status by development phase and sample collection approach

Genomic study objectives	DNA collection	Trials, N		
		Phase II	Phase III	All
Yes				
Prespecified analyses	*Any*	9	4	13
	Mandatory	4	1	5
	Optional	5	3	8
Unspecified analyses	*Any*	32	23	55
	Mandatory	0	6	6
	Optional	32	17	49
None	*None*	59	20	79

"Any" is the total of mandatory and optional.

a specific drug or mechanism. Often, this has resulted in therapies targeted to patients whose tumors harbor certain genetic alterations.[13] Although genetic variants associated with susceptibility to many common chronic diseases have been discovered,[14–16] these discoveries have not translated into wider development of "targeted therapies" as they have in the oncology setting.

Genetic association studies have identified numerous gene variants that alter susceptibility to or prognosis in CVD and its risk factors,[14] but how they have translated into genetic or genomic biomarkers used in clinical trials has not been evaluated. We sought to evaluate the extent to which genomic biomarkers are used in clinical trials of treatments for CVD and its associated risk conditions and to examine how known genetic associations between drug targets and CVD traits are reflected in the use of genomic biomarkers in clinical trials.

RESULTS

Prospective and exploratory genomic biomarker utilization in phase II and III clinical trials

We identified 134 development programs for CVD and relevant risk factors from PharmaProjects, 67 (50%) of which contained 147 relevant clinical trial protocols for evaluation. We identified 1 of 147 trials (0.7%) employing a genomic biomarker. In this lone trial, a predictive genomic biomarker was used to select patients for enrollment into a phase III clinical trial.

Use of exploratory genomic biomarkers for hypothesis generation beyond the primary and secondary objectives was more common. Of the 147 clinical trial protocols, 55 trials (37%) stated an intention to study unspecified biomarkers or a general class of genes (e.g., those related to drug metabolism or inflammatory genes), 13 trials (8.8%) prespecified analysis of genes harboring a genomic biomarker of interest, and the remaining 79 trials (54%) did not specify the application or analysis of genomic biomarkers. Overall, 33 (25%) of the 134 development programs included exploratory biomarker aims in their protocols. The collection of DNA for exploratory genomic biomarker analyses did not differ by study phase (P = 0.16).

DNA sample collection rates and purpose are shown in **Table 1**. Forty-six percent (n = 68) indicated sample collection in their protocols. Of these 68 trials that indicated sample collection, 57 (84%)

made the collection of samples for pharmacogenomic analysis optional, whereas 11 (16%) mandated sample collection as a condition for study enrollment. Of note, where sample collection was mandatory (n = 11), it was nominally greater in phase III protocols (n = 7; 64%) than in phase II protocols (n = 4; 36%). Trials with specific exploratory biomarker aims were more likely to have DNA collection mandated in the trial (P = 0.02).

Determining genetic association for drug targets under development

A total of 93 of 134 (69%) programs publicly disclosed the drug target, accounting for 73 unique drug targets across the therapeutic areas. We determined that 23 of 73 targets (32%) had genetic variants related to the disease, which could support the feasibility of prospective genomic biomarker use in clinical trials. Of the 50 targets (68%) that had no evidence of genetic association with the disease for which they were being studied, 16 (33%) were associated with other CVD traits or risk conditions (**Table 2**).

To contrast the liberal criteria for designating genetic association from open targets above, we analyzed targets in the pipeline against a list of genes curated by an American Heart Association-commissioned expert working group[14] as being implicated in various CVD traits. Fifty-two targets were in development for the subset of CVDs in the study that overlapped with CVDs for which susceptibility loci were compiled (hypertension, lipid disorders, ischemic heart disease, type 2 diabetes, and stroke). Of the 52 targets, 4 (8%) were associated with the disease for which they were being developed.

DISCUSSION

Despite CVD association with high morbidity and mortality,[1] there has been declining drug development activity in this area[9] and high drug development attrition rates.[17] Opportunities exist to stimulate new drug development through the application of biomarkers.[10,11] Therefore, we examined the extent to which genomic biomarkers are used in clinical trials for CVD and its risk factors. Our analysis reveals a near absence of prospective genomic biomarker use in the examined clinical trials, even in cases where genetic variability in the drug target is relevant to the indication being pursued. However, half of the drug development programs include the genomic assessments in clinical trials that ostensibly would support future targeted drug development. Overall, the application of genomics in the clinical trials for CVD was limited.

Understanding of the molecular and genetic basis for a disease is an important limiting factor for using genomics prospectively in clinical trials.[13] Drug developers could use such genomic characteristics to identify subsets of patients that may receive different drug doses, may be excluded from trials, or for whom studies can be enriched. CVD is pathophysiologically complex and polygenic. However, numerous genetic susceptibility markers for CVD and metabolic traits have been identified that could support targeted drug development.[14] In addition, genomic attributes for classifying patient populations do not have to be linked to a mechanistic or pathophysiologic characteristic;[18] important genetic mediators of drug response can also reside in genes relevant to the drug's

Table 2 Targets in development pipeline of CVD therapeutic areas

Disease queried in PharmaProjects	Genetic association derived from Open Targets 18.10 [17]		Genetic association derived from AHA working group [14]		Targets not genetically associated with any CVD trait from either source
	Targets that have genetic association with pursued indication[a]	Targets that have genetic association with other CVD traits[b]	Targets that have genetic association with pursued indication[a]	Targets that have genetic association with other CVD traits[b]	
Atrial fibrillation	KCNJ5	ADRB1	N/A	–	ADRB2, KCNJ3, KCNA5
Chronic heart failure	–	ENPEP, NPR1, NRG1, NR3C2, ADRB1	N/A	ENPEP, ADRB1	SLC9A3, VIPR1, ATP1A1, NPR2, TRPV4
Hypertension	NPR1, AGTR1	EDNRA, NR3C2	–	EDNRA	–
Ischemic heart disease	MAPK14, CETP, LPA, RAMP1, CALCRL, APOA1	LCAT	LPA	–	BRD4, FGFR1
Lipid disorders[c]	APOC3, PPARA, CETP, APOE, PCSK9	PPARD, APOA, PON1	ANGPTL3	PCSK9, APOE	ACLY, DGAT1, SREBF2
Metabolic syndrome	–	–	N/A	–	DGAT1
Obesity	MC4R	SLC5A1	N/A	MC4R	GLP1R, SLC6A4, SLC6A3, SLC6A2, SLC5A2, FGFR4, NR3C1
Peripheral vascular disease	PLG	PDE5A	N/A	PLG	PTGER1, PLA2G1B, TBXAS1
Stroke	PLG	ABCC8	–	–	EGLN2, PLAT, MAOB, CPB2, TBXAS1, PROC
Thrombosis	VWF, F11		N/A	–	–
Type 2 diabetes	SLC5A1, GCK, GCGR, GIPR, GLP1R, NFKB1	PPARA, MTTP, LYN, PTPN1, SLC22A12, SLC10A2	GCK, GIPR	–	IL1A, SLC5A2, DPP4, BRD4, FBP1, FGFR4, DGAT1, GPR119, FFAR1, NLRP3, FGF21

Genes targeted by at least one drug in development for the queried disease, and genes encoding a heteromeric complex are separated as individual targets.
AHA, American Hospital Association; CVD, cardiovascular disease; N/A, Comparison not applicable as the AHA-working group publication did not compile susceptibility loci for the disease.
[a]Integrated evidence points to genetic association with disease. [b]No genetic association with the queried disease but genetic association with other CVD traits or risk conditions. [c]The lipid phenotypes hypercholesterolemia, hypertriglyceridemia, and hyperlipidemia were queried and assembled in this category. Programs in development for cardiomyopathies had no drug target identified.

pharmacokinetics, genes that may predispose patients to toxicities,[19] or multiple genes that in aggregate influence risk or prognosis (e.g., polygenic risk scores).[20–22]

In our study, a genomic biomarker was used to select patients for a clinical trial in only one case; this biomarker was for the primary target of the drug and gene variant effects on related drugs that had been extensively studied. Despite the availability of biomarkers, biomarker-based trial enrichment or stratification maneuvers were otherwise absent in the CVD pipeline. It is possible that the availability of measurable biomarkers, such as blood pressure, ejection fraction, hemoglobin A1C, and low-density lipoprotein for the studied traits, facilitate individualized assessments of patient prognostic risk and, therefore, reduce the incentive to use predictive or prognostic biomarkers in clinical trials. These findings may also suggest that genomic enrichment strategies are not considered a practical approach to improve the efficiency of clinical trials in the cardiovascular space, or that such biomarkers are not robust enough for prospective testing due to inadequate understanding of the role of the biomarker in the context of the disease.

Approximately half of the programs specified DNA collection for exploratory genomic studies, reflecting an aspiration to generate new hypotheses on potential drug targets, genomic biomarkers associated with disease, and drug responses that may impact success rates and longer-term outcomes. The rate of DNA collection observed in our study may be indicative of reported variability in global DNA collection practices, which has been attributed to factors including differing country requirements and prompted efforts to harmonize regulations to collect genetic samples.[23,24] Although our study was restricted to drug development programs in the United States, the global nature of clinical trials exposes trials in our study to such a challenge.

In a 2013 guidance for industry, the US Food and Drug Administration (FDA) provided recommendations for assessing the effect of genomic variations in exploratory studies and their use in the study of a drug's pharmacokinetics, pharmacodynamics, efficacy, and safety.[19] These recommendations were issued with the expectation that the interrogation of genomic information throughout the drug development process would enable discovery

of clinically important genomic differences before a product is marketed. Regardless of whether a development program is influenced by early phase genomic assessments, analysis of data gathered from exploratory studies can generate hypotheses and provide further insight into disease classification or classification of drug exposure and response by defined genomic biomarkers.

By our derivation of genetic association via open targets and the expert-curated lists of CVDs susceptibility loci, 32% and 8%, respectively, of targets in the pipeline were genetically associated with the CVD trait for which they are being developed. Another study reported a low proportion of genetically supported drug targets for approved indications in this therapeutic area.[11] Considering a recent report that interindividual genome variability can predict drug withdrawals,[25] it is plausible that the possibility of variable drug response has influenced the pursuit of drug targets harboring significant genetic variability. However, some targets may not harbor clinically relevant genetic variability and may contribute to the seeming lack of genetic support in some drug targets in development. Additionally, despite evidence for disease-risk loci, other considerations, including where the genetic variation manifests in the disease pathway, the nature of the genetic variation, and the druggability of the encoded protein, support the actualization of a genetically validated drug target. With the increasing ability of technologies to better characterize the impact of genomic variations on diseases, the landscape of disease-associated targets will continue to evolve.

Our findings have certain limitations, particularly those inherent to the databases and sources used. First, 31% of the CVD development programs we identified did not have targets listed in PharmaProjects; therefore, the drug targets that are listed may limit our estimation of targets in development and consequent association with disease. Second, we did not appraise the scoring algorithm used to determine the evidence for genetic association between the drug target and disease on the open targets platform or examine the standard used to select disease susceptibility loci compiled by the expert-working group. In the latter case, for example, the expert-reviewed list may represent an underestimate of validated targets, as only novel targets were reviewed and listed for some traits (e.g., the lipid phenotypes). Finally, genetic associations may have been recently discovered but potential evaluable products from such associations are yet to mature in to the phases of development we studied.

In conclusion, we examined the CVD pipeline to characterize the use of genomic biomarkers in clinical trials and how known disease-associated genetic variants, from an open database and an expert-curated list, are represented in the pipeline. Our efforts reveal a near absence of prospective genomic biomarker utilization in CVD clinical trials, a finding bolstered by limited representation of genetically supported targets in the pipeline. These findings underscore a potential for knowledge of genetic-variation to play a role in enhancing drug development for CVD traits, whether in pursuing new targets or conducting more genetically targeted clinical trials.

METHODS

Inclusion of diseases in the study
CVDs were initially selected for inclusion in the study from the list of top 30 diseases in the United States.[1] The list was supplemented with diseases and conditions that are inter-related with CVD pathophysiology and its associated risk conditions, some of which also account for the high global disease burden. Ultimately, drugs in development for ischemic heart disease, type 2 diabetes, hypercholesterolemia, hypertriglyceridemia and hyperlipidemia, hypertension, metabolic syndrome, obesity, thrombosis, peripheral vascular disease, congestive heart failure, cardiomyopathies, and atrial fibrillation were included.

Data sources and extraction
We queried the commercial database PharmaProjects (Citeline, New York, NY) for drugs in development phases II and III in the United States for the selected diseases. Additional search terms were included to improve the yield of the search. For example, the query for ischemic heart disease included the additional search terms atherosclerosis, coronary artery disease, acute coronary syndrome, angina, myocardial infarction, and heart attack (**Table S1**).

Cell therapies, gene therapies, reformulations of approved drugs for the same disease, and imaging agents were excluded from the search results. For each result, we extracted the following information: drug name, drug target, mechanism of action, and disease for which the drug is being developed. Each result was manually checked to ensure its inclusion conformed to criteria in the search query. Citeline provided real-time updates to results of search queries. Results and updates of queries from December 2015 to March 2017 were included in the study.

Prospective and exploratory genomic biomarker utilization in phase II and III clinical trials
Drugs in development were linked to investigational new drug (IND) numbers in the FDA's Documents Archiving, Reporting, and Regulatory Tracking System (DARRTS). Clinical trial protocols and investigator brochures submitted under the INDs to the FDA before March 2017 were evaluated. IND applications submitted in nonelectronic format, programs with no relevant protocols, and programs with unidentifiable IND numbers were excluded.

Phase of development was recorded as identified by the developer on the protocol. For this analysis, protocols spanning two phases were assigned the phase of the most progressive stage referenced (e.g., a phase Ib/IIa study was designated as phase II, whereas a phase IIb/III study was designated as phase III). The protocols of phase II and phase III studies were reviewed for genomic biomarker strategies as follows: (i) the use of predictive or prognostic genomic biomarkers[26] to select the trial population, (ii) the use of a genomic biomarker to stratify randomization, and (iii) the analysis of treatment effect in subgroups of patients defined by a genomic biomarker. An overview of the data sources and filtering process is presented in **Figure 1**.

The protocols were further reviewed for exploratory genomic biomarker assessment, defined as biomarkers studied outside of the primary and secondary objectives. Exploratory genomic biomarkers were determined to be prespecified if the drug developer explicitly identified the biomarker to be studied or unspecified if the biomarkers to be studied were unidentified. Information on the approach to germline DNA collection for pharmacogenomic studies was determined as either mandatory (i.e., collection required for a participant to be a part of the trial) or optional (enrollment in the study possible without consenting to DNA collection). The schedule for genomic sample collection schedule was also derived from the protocol.

DNA collection time was designated as at baseline when the sample was collected prior to treatment initiation but was designated as flexible when the trial permitted collection at any point in the study, including when protocol called for baseline collection but permitted collection if the patient consented to it after the treatment was initiated. Studies indicating that sample collection could take place at any time or which did not specify the timing of sample collection were deemed flexible as well.
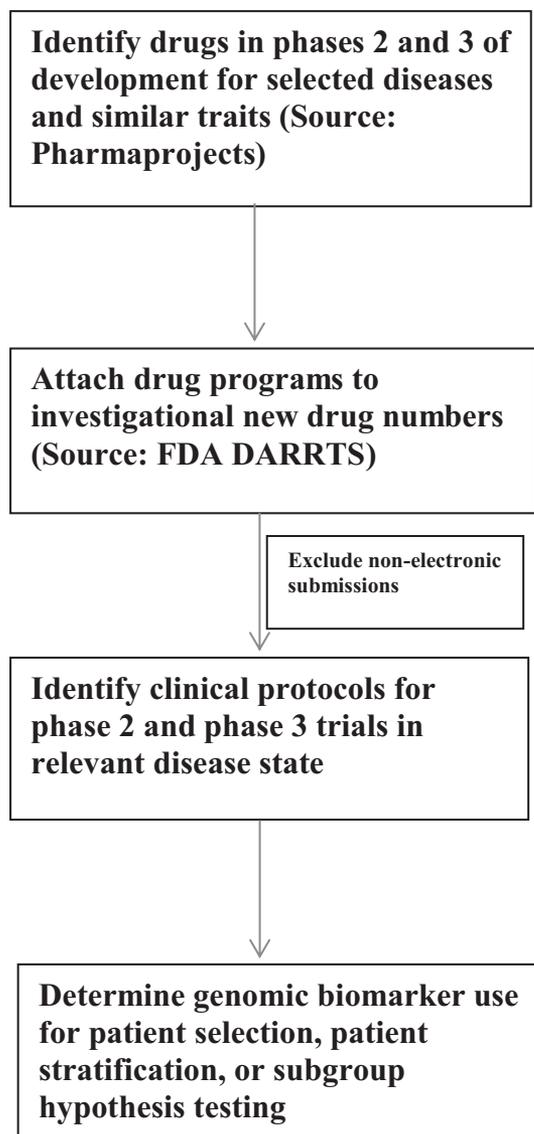
Figure 1 Evaluation of genomic biomarker use in the clinical trials. DARRTS, Documents Archiving, Reporting, and Regulatory Tracking System; FDA, US Food and Drug Administration.

Statistical analysis

Association between phase of development and classification of exploratory biomarker aim, exploratory biomarker status, and genomic sample collection were all determined using Fisher's exact test. Multiplicity-corrected statistical significance was claimed at an alpha level of 0.025 (two-sided). Statistical analyses were performed using SAS 9.4 (Cary, NC).

Determining genetic association for drug targets under development

We identified diseases/traits with the genetic association scores > 0 as scored on the Open Targets Platform (release 18.10; http://www.target-validation.org),[27] a peer-reviewed,[28] publicly available platform that integrates evidence of association of a drug target with diseases. On the Open Targets Platform, the functional consequence of a variant is an important factor in scoring the strength of genetic association. The final score of a target-disease association ranged from zero, no evidence supporting association; to one, representing the strongest evidence of target-disease association.[28] No limit was placed on the strength of evidence and a target

with a genetic association score > 0 to a cardiovascular trait was deemed to have genetic support.

A list of genetically validated, disease-relevant targets was compiled from the latest scientific statement from the American Heart Association on the genetics and genomics for the prevention and treatment of CVD.[14] We made comparisons only among diseases in this study that corresponded to a disease reviewed by the working group (e.g., neither thrombosis nor atrial fibrillation was reviewed by the working group and they are not compared). Type 2 diabetes–related genes combined expert-reviewed lists of type 2 diabetes mellitus, fasting insulin, 2-hour glucose, and fasting glucose. Genes related to lipid phenotypes were derived from expert-reviewed lists of genes associated with total cholesterol, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, and triglycerides. The list of targets compared for stroke were those compiled for sporadic cardioembolic ischemic stroke, sporadic large-vessel atherosclerotic stroke in the expert-curated list, and for hypertension were those compiled for hypertension, systolic blood pressure, diastolic blood pressure, mean arterial pressure, and pulse pressure in the expert-curated list. Genes with no matching Human Genome Organization (HUGO) symbol, genes that were withdrawn as determined from genenames.org, and those with no gene associated with the identified loci were excluded from the published list and not used for comparison.

CONFLICT OF INTEREST
The authors declared no competing interests for this work.

AUTHOR CONTRIBUTIONS
O.A., A.R., R.S., J.S., J.W., I.Z., and M.P. wrote the manuscript. O.A., A.R., R.S., and M.P. designed the research. O.A. performed the research. O.A., A.R., R.S., and M.P. analyzed the data.

DISCLAIMER
The article reflects the views of the authors and should not be construed to represent the FDA's views or policies.

1. Murray, C.J. *et al.* The state of US health, 1990–2010: burden of diseases, injuries, and risk factors. *JAMA* **310**, 591–608 (2013).
2. Hwang, T.J., Lauffenburger, J.C., Franklin, J.M. & Kesselheim, A.S. Temporal trends and factors associated with cardiovascular drug

development, 1990 to 2012. *JACC Basic Transl. Sci.* **1**, 301–308 (2016).

3. Kaitin, K.I. & DiMasi, J.A. Pharmaceutical innovation in the 21st century: new drug approvals in the first decade, 2000–2009. *Clin. Pharmacol. Ther.* **89**, 183–188 (2011).

4. Beierlein, J.M., McNamee, L.M., Walsh, M.J., Kaitin, K.I., DiMasi, J.A. & Ledley, F.D. Landscape of innovation for cardiovascular pharmaceuticals: from basic science to new molecular entities. *Clin. Ther.* **39**, 1409–1425 (2017).

5. Pammolli, F., Magazzini, L. & Riccaboni, M. The productivity crisis in pharmaceutical R&D. *Nat. Rev. Drug Discov.* **10**, 428–438 (2011).

6. Hay, M., Thomas, D.W., Craighead, J.L., Economides, C. & Rosenthal, J. Clinical development success rates for investigational drugs. *Nat. Biotechnol.* **32**, 40–51 (2014).

7. Ringel, M.S., Shah, N.A., Baedeker, M., Lim, C.T., Lamichhane, A. & Schulze, U. Occlusion in the flow of new drugs for cardiovascular disease. *Clin. Pharmacol. Ther.* **102**, 246–253 (2017).

8. Sacks, L.V., Shamsuddin, H.H., Yasinskaya, Y.I., Bouri, K., Lanthier, M.L. & Sherman, R.E. Scientific and regulatory reasons for delay and denial of FDA approval of initial applications for new drugs, 2000–2012. *JAMA* **311**, 378–384 (2014).

9. Fordyce, C.B. *et al.* Cardiovascular drug development: is it dead or just hibernating? *J. Am. Coll. Cardiol.* **65**, 1567–1582 (2015).

10. Cook, D. *et al.* Lessons learned from the fate of AstraZeneca's drug pipeline: a five-dimensional framework. *Nat. Rev. Drug Discov.* **13**, 419–431 (2014).

11. Nelson, M.R. *et al.* The support of human genetic evidence for approved drug indications. *Nat. Genet.* **47**, 856–860 (2015).

12. Thomas, D.W., Burns, J., Audette, J., Carroll, A., Dow-Hygelund, C. & Hay, M. *Clinical Development Success Rates 2006-2015*. (Bio, Biomedtracker, Amplion, 2016).

13. US Food and Drug Administration. Targeted drug development: why are many diseases lagging behind? <https://www.fda.gov/downloads/AboutFDA/ReportsManualsForms/Reports/UCM454996.pdf> (2015).

14. Ganesh, S.K. *et al.* Genetics and genomics for the prevention and treatment of cardiovascular disease: update: a scientific statement from the American Heart Association. *Circulation* **128**, 2813–2851 (2013).

15. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014).

16. Cuyvers, E. & Sleegers, K. Genetic variations underlying Alzheimer's disease: evidence from genome-wide association studies and beyond. *Lancet Neurol.* **15**, 857–868 (2016).

17. Cottingham, M.D., Kalbaugh, C.A. & Fisher, J.A. Tracking the pharmaceutical pipeline: clinical trials and global disease burden. *Clin. Transl. Sci.* **7**, 297–299 (2014).

18. US Food and Drug Administration. Guidance for industry: enrichment strategies for clinical trials to support approval of human drugs and biological products <https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM332181.pdf> (2012).

19. US Food and Drug Administration. Guidance for industry: clinical pharmacogenomics: premarket evaluation in early-phase clinical studies and recommendations for labelling <https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM337169.pdf> (2013).

20. Khera, A.V. *et al.* Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat. Genet.* **50**, 1219–1224 (2018).

21. Knowles, J.W. & Ashley, E.A. Cardiovascular disease: the rise of the genetic risk score. *PLoS Med.* **15**, e1002546 (2018).

22. Natarajan, P. *et al.* Polygenic risk score identifies subgroup with higher burden of atherosclerosis and greater relative benefit from statin therapy in the primary prevention setting. *Circulation* **135**, 2091–2101 (2017).

23. Franc, M.A., Warner, A.W., Cohen, N., Shaw, P.M., Groenen, P. & Snapir, A. Current practices for DNA sample collection and storage in the pharmaceutical industry, and potential areas for harmonization: perspective of the I-PWG. *Clin. Pharmacol. Ther.* **89**, 546–553 (2011).

24. Ricci, D.S. *et al.* Global requirements for DNA sample collections: results of a survey of 204 ethics committees in 40 countries. *Clin. Pharmacol. Ther.* **89**, 554–561 (2011).

25. Lee, K.H., Baik, S.Y., Lee, S.Y., Park, C.H., Park, P.J. & Kim, J.H. Genome sequence variability predicts drug precautions and withdrawals from the market. *PLoS One* **11**, e0162135 (2016).

26. FDA-NIH Biomarker Working Group. BEST (Biomarkers, EndpointS, and other Tools) Resource (Food and Drug Administration, Silver Spring, MD, and National Institutes of Health, Bethesda, MD, 2016). <https://www.ncbi.nlm.nih.gov/books/NBK326791/>.

27. Open Targets platform <https://www.targetvalidation.org/>.

28. Koscielny, G. *et al.* Open Targets: a platform for therapeutic target identification and validation. *Nucleic Acids Res.* **45**, D985–D994 (2017).