

REVIEW

Innovation at the Intersection of Clinical Trials and Real-World Data Science to Advance Patient Care

Brandon Swift^{1,*}, Lokesh Jain^{2,*}, Craig White³, Vasu Chandrasekaran⁴, Aman Bhandari⁴, Dyfrig A. Hughes⁵ and Pravin R. Jadhav⁶

While efficacy and safety data collected from randomized clinical trials are the evidentiary standard for determining market authorization, this alone may no longer be sufficient to address the needs of key stakeholders (regulators, providers, and payers) and guarantee long-term success of pharmaceutical products. There is a heightened interest from stakeholders on understanding the use of real-world evidence (RWE) to substantiate benefit–risk assessment and support the value of a new drug. This review provides an overview of real-world data (RWD) and related advances in the regulatory framework, and discusses their impact on clinical research and development. A framework for linking drug development decisions with the value proposition of the drug, utilizing pharmacokinetic–pharmacodynamic–pharmacoeconomic models, is introduced. The summary presented here is based on the presentations and discussion at the symposium entitled *Innovation at the Intersection of Clinical Trials and Real-World Data to Advance Patient Care* at the American Society for Clinical Pharmacology and Therapeutics (ASCPT) 2017 Annual Meeting.

Clin Transl Sci (2018) 11, 450–460; doi:10.1111/cts.12559; published online on 16 May 2018.

The fundamental goal of advancing patient care through precision and translational medicine is to provide targeted treatments enabling favorable treatment outcomes while minimizing the risk. Traditional clinical trials and regulatory approval processes focus on “does the drug work?” under a(n) selected/ideal design. While this is reasonable, it may not provide sufficient information on how well the drug works under real-world conditions in varied contexts (e.g., polypharmacy or comorbidities) and across patient subpopulations.¹ Consequently there has been an increasing focus on inclusion of real-world data (RWD) in healthcare decisions as well as in the development and commercialization of new medicines.^{2,3} Further, with the heightened attention to value-based pricing, pharmaceutical companies are under increased pressure to demonstrate the value of new drugs in the context of their routine use. It is more pressing than ever, therefore, to understand the value (cost-effectiveness) of new drugs early in the development process using novel predictive approaches such as pharmacokinetic–pharmacodynamic–pharmacoeconomic (PK-PD-PE) models.

Advances in digital technology and analytics are making use of RWD more feasible than ever; however, important challenges remain to be resolved. Privacy laws, technical complications, and evolving regulations have all hindered access and implementation of RWD to improve the efficiency of the drug development cycle. Uptake of RWD to inform development of the target product profile and/or design of clinical studies is still a work in progress.⁴ Nevertheless,

as discussed in this review, many of these hurdles can be overcome to improve clinical research, support application for marketing authorization, conduct postmarketing safety surveillance, and expand patient access.

The aim of this article is to provide an overview of RWD and recent advances in the regulatory framework, highlight gaps and limitations, and discuss implementation opportunities in the areas of drug development, regulatory assessments, medical practice, and payer assessments to readers who are familiar with clinical development applications of PK-PD models. More important, a method for integrating RWD within the paradigm of model-based drug development, namely PK-PD-PE modeling, is introduced, to highlight opportunities for contribution in early value assessment for clinical pharmacology and pharmacometric scientists.

REAL-WORLD DATA AND REAL-WORLD EVIDENCE

Real-world data (RWD) and real-world evidence (RWE) are typically used interchangeably. The following section will introduce, discuss, and provide clarity around these terms.

Types and sources of real-world data

When used in the healthcare context, the term “Real-World Data” usually refers to patient-level data gathered outside the conventional clinical trial setting. Such data may be generated in the course of normal clinical practice or administrative claims processing, or may be reported directly by patients. Examples include data from: patient charts,

¹Advisory Analytics, IQVIA, Durham, North Carolina, USA; ²Quantitative Pharmacology and Pharmacometrics, Merck & Co., Inc., Rahway, New Jersey, USA; ³Harvard PhD program in Health Policy, Cambridge, Massachusetts, USA; ⁴Center for Observational and Real World Evidence, Merck & Co., Inc., Boston, Massachusetts, USA; ⁵Centre for Health Economics and Medicines Evaluation, Bangor University, Bangor, Gwynedd, UK; ⁶Corporate Projects, Research & Development (R&D) Innovation, Otsuka Pharmaceutical Development and Commercialization (OPDC), Princeton, New Jersey, USA. *Correspondence: Brandon Swift (brandon.swift@quintiles.com); Lokesh Jain (lokesh.jain@merck.com)

The first two authors contributed equally to this work.

Received 13 December 2017; accepted 29 March 2018; published online on: 16 May 2018. doi:10.1111/cts.12559

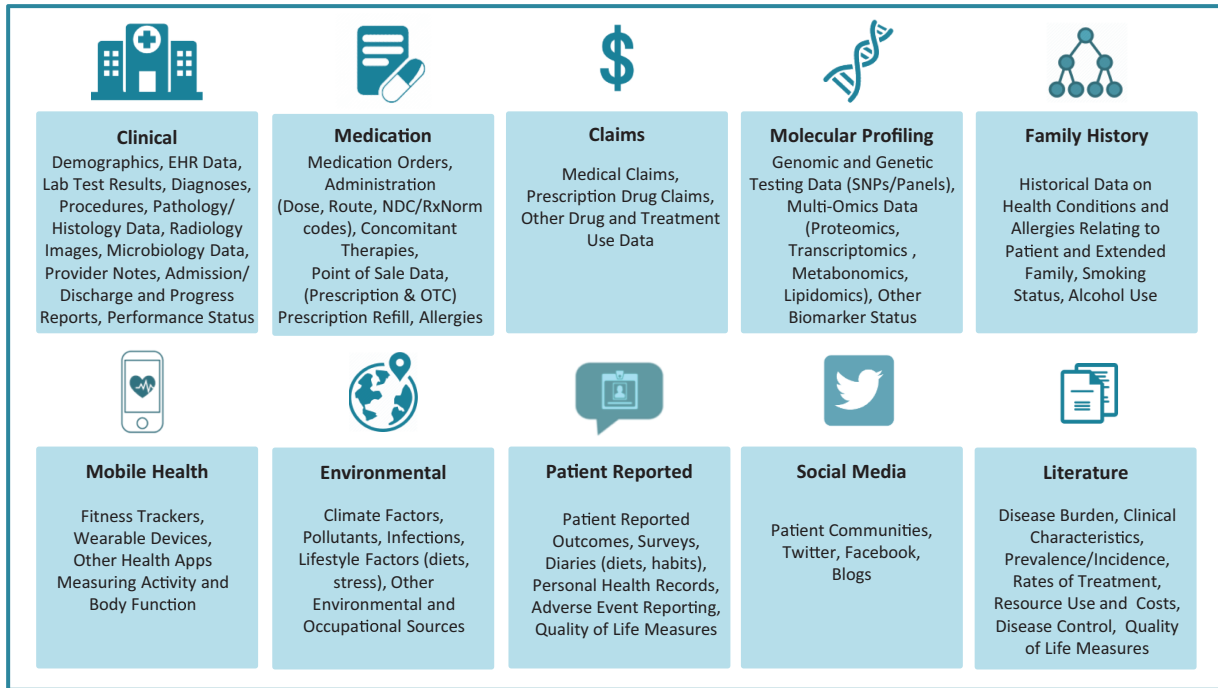


Figure 1 Types of real-world data.

laboratory reports, prescription refills, patient registries, patients treated on- and off-label, patients treated through expanded access, pragmatic clinical trials, surveys, and mobile health devices, as well as other data from existing secondary sources used to support decisions concerning safety, quality, care coordination, coverage, and reimbursement.⁵

Advances in technology, data science, and healthcare policies have resulted in tremendous growth in the volume, sources, and utilization of RWD with collection of larger and more diverse data sets. The expansion in the use of electronic health records (EHRs) and the proliferation of consumer digital technologies including mobile devices, wearables, sensors, adherence tools, social media platforms, and online patient communities have provided new data sources as well as improved means of capturing, storing, and analyzing longitudinal RWD on patients. The current RWD landscape is characterized by enormous variety and complexity (**Figure 1**). It extends beyond traditional sources such as chart reviews, prescription, or claims data to include both structured and unstructured data from a host of heterogeneous sources. These data include, among others, phenotypic and genotypic data from discrete fields as well as clinical notes in electronic health records, multi-omics and other molecular profiling data from biospecimen banks, and patient-reported outcomes from surveys and prospective registries. Mobile health devices and other wearable applications comprise additional novel sources of previously unavailable patient-level data. These devices offer continuous monitoring, data collection, and real-time transmission capabilities that is rarely achieved in routine clinical care.⁶ Online patient communities such as PatientsLikeMe⁷ and initiatives like PCORnet⁸ as well as consumer genetic

services like 23andMe (Mountain View, CA) and uBiome (San Francisco, CA) have led to the rise of empowered patients who are more open and willing to share their health information for decision-making and research purposes.

Real-world evidence

Real-world evidence refers to the output of RWD analysis that is used to generate insights, using appropriate study design and scientific methods, to inform decision-making by healthcare stakeholders. Generating evidence from RWD therefore depends not only on capturing “big data”—large volumes of these diverse data—but in effectively integrating these multiple and often disparate sources of data to obtain meaningful insights. Most recently, a publication from the US Food and Drug Administration (FDA) broadened the definition of RWE to any data “generated from any study design (including RCTs) as long as the data source is from routine care and the design is highly pragmatic, meaning the trial design and conduct closely approximate the eventual use of the product in clinical practice.”⁹

Regional perspectives on real-world data

The emphasis on the use of RWE to inform and improve healthcare decisions reverberates across all major markets. However, the acceptance and applications of RWE for decision-making is variable across the globe. In the United States, the FDA has long been interested in using healthcare data generated in the real world to learn about medical products, particularly drug safety. In considering this, the FDA launched the Sentinel initiative in 2008, which allows monitoring of the safety of FDA-regulated products using RWD from sources such as EHRs, insurance claims data, and registries.¹⁰ More recently, the Sentinel capabilities

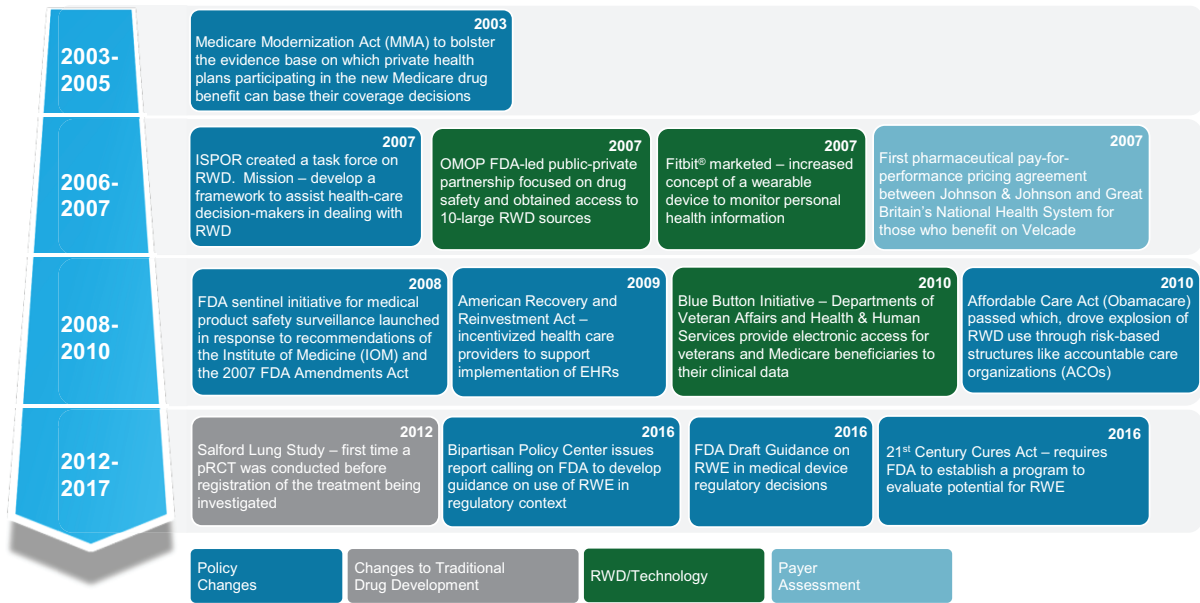


Figure 2 Key events leading to the importance of real-world evidence. International Society for Pharmacoeconomics and Outcomes Research (ISPOR); Observational Medical Outcomes Partnership (OMOP); pragmatic randomized clinical trial (pRCT); real-world evidence (RWE); real-world data (RWD); electronic health record (EHR).

were expanded under a public-private partnership in order to provide access of Sentinel data to private-sector entities, such as regulated industry, academic institutions, and nonprofit organizations. This program, named Innovation in Medical Evidence Development and Surveillance (IMEDS), is a national resource of big healthcare data, promoting the use of RWD for research related to broader public health benefit and medical evidence generation.¹¹ RWD is also used routinely in the EU, particularly for monitoring of safety and drug utilization for marketed products. Further, in the EU there is “increasing interest in the use of RWE for efficacy, outcomes for Health Technology Assessment (HTA), and for rapid cycle evaluation of medicines.”¹²

Government policies to address rising healthcare costs and the need for better ways to measure performance have also fueled a strong demand for RWD. In the United States, healthcare reforms intended to improve healthcare quality and reduce costs are a strong driver in this respect. Value-based payment reforms contain direct provider incentives that rely on the collection, reporting, and analysis of RWD to assess and improve provider performance based on approved quality metrics. In Europe, similar budget pressures are driving HTA bodies and payers to use RWE in conjunction with evidence from clinical research to inform reimbursement decisions.

PROMISE OF RWD FOR VARIOUS STAKEHOLDERS

As RWD has become more robust and ubiquitous, various stakeholders have become increasingly interested in its use and application in a variety of different settings (**Figure 2**).

Regulatory agencies

The passage of the 21st Century Cures Act (“The Act”) by the US Congress in December 2016 opened up a new

pathway for pharmaceutical companies to leverage RWD in the development and expansion of indications for their products in the United States, which was reinforced in the Prescription Drug User Fee Act (PDUFA) VI authorization. The Act clearly defines a role for RWE in the regulatory process by mandating that the US FDA must issue guidance describing how pharmaceutical companies may use RWE:

1. To help support the approval of a new indication for a drug approved under section 505(c); and
2. To help support or satisfy postapproval study requirements.

This has significant implications for both the cost and speed of drug development. Pharmaceutical companies have sought to utilize combinations of multiple existing registries to satisfy postapproval safety study (PASS) requirements in both US and EU contexts.¹²

In addition to increased use due to better data availability and regulatory accommodation, the role of RWD/RWE only appears poised to increase as drugs target smaller and smaller populations of patients. These niche populations limit pharmaceutical companies’ attempts to undertake randomized controlled trials (RCTs). RWD could potentially provide another axis of data to consider how evidence of efficacy and safety are substantiated in niche populations. RWD may be combined with traditional data to increase efficiency and reduce costs of clinical development without lowering the standard of evidence. RWE can be leveraged for regulatory decision-making related to label extension or to support a new indication for an approved drug, as well as to substantiate confirmatory evidence for drugs approved under the expedited regulatory programs. Other authors have described opportunities for single-arm trials

to be implemented and evaluated through the concept of “threshold-crossing.”¹³ In this model, an efficacy or safety threshold is established as a benchmark for a new drug via the use of RWD, and if this threshold is achieved the new drug can be considered successful and can forgo RCT evaluation. If unsuccessful, a traditional RCT is established. Finally, RWD/RWE is being used to fulfill postmarketing commitments in many instances, and increased uptake of RWE will further improve the efficiency of such monitoring and surveillance activities.

Payers

Besides the use of RWD in a regulatory context, novel types of RWD are creating efficiencies and permitting uses heretofore unavailable to pharmaceutical companies and payers. In many indications, end points for efficacy and effectiveness are subjective, or occur so infrequently in early stages of the disease progression that drug developers choose to use proxies for the true clinical end points. Examples include cholesterol as a surrogate for adverse cardiovascular events, HbA1c for diabetes end points such as nerve damage or kidney disease, and progression-free survival in many types of cancer. In addition to surrogate-to-real end-point comparisons, payers have concerns about small sample sizes among drugs approved to treat rare diseases. RWD can facilitate the creation of larger cohorts of patients with rare diseases where patients are extremely difficult to identify, or those slow-progressing conditions, otherwise not amenable to prospective clinical trial-based analyses. This is being achieved via aggregation of multiple RWE data sets to enable creation of “virtual control arms” or “synthetic patient cohorts” that combine smaller cohorts of similar indication and stage of disease multiple data sets. This combined group of patients can then be used as both test and counterfactual arms for rare-disease analyses.

Increasingly, payers are concerned about the validity of surrogate end points among their populations, and the smaller sample sizes used to test new medications in targeted and genetic therapies.¹⁴ Payers are turning to their own or third-party RWD repositories to undertake analyses of the relationships between these surrogate end points reported from randomized clinical trials, and the true clinical end points available in larger RWD repositories that are most relevant to the disease being treated. Large data sets are also being used to collect data from prospective drug trials to provide estimates of drug treatment effects in broader, more heterogeneous populations using a variety of end points. The use of these data sets can overcome the limitations of clinical trials, where end points have to be valid within the (relatively short) duration of the trial, and the set of end points captured is often limited by financial and patient and provider burden considerations. This is a concept exemplified by the Salford Lung Study.¹⁵

These questions of surrogate to real-end point consistency and efficacy in heterogeneous patient populations are significant in the context of medication value, where outcomes among the highly selected patient participants in an RCT can be markedly different when seen in a real-world patient population with comorbidities and imperfect medication adherence. Payers have begun to use these RWD-enabled

analyses to drive conversations with pharmaceutical companies regarding prices for therapies and appropriate payment models.¹⁶ The availability of RWD enables these new payment models by allowing pharmaceutical companies and payers to identify performance in individual patients and pay for good outcomes, or avoid paying for bad ones for drugs while they are being used by patients in real practice. This parallels the pay-for-performance models that have been, and continue to be used, for healthcare providers.

Care delivery system

The use of RWD may also permit more accurate ways to align drugs to patients than is possible with clinical trial or anecdotal physician experience data. Using RWD, population health analysts and epidemiologists are able to identify large subsets of patients within specific disease populations, and using phenotypic, genotypic, or laboratory data available within the patient record, more effectively and appropriately assign treatments. Examples of this in practice include analysis of RWD on diabetic patients' use of specific drugs to identify inappropriate use of medications with renal dosing implication among patients with kidney disease, through estimated glomerular filtration rate calculation via electronic medical record (EMR) data.¹⁷ The application of these results at a population-level and via feedback to physicians through clinical decision support systems embedded within the EMR and data collection tools can lead to better patient management.

The potential for RWD to improve the provision of health care is enormous. It remains to be seen, though, to what extent RWD can significantly change the healthcare landscape.

BARRIERS AND SOLUTIONS TO ENABLE USE OF REAL-WORLD DATA

While RWD offers significant opportunities for improving healthcare research, innovation, and decision-making, as with any rapidly evolving field there are challenges to leveraging its full potential. These challenges range from technical (e.g., claims data are collected for administrative billing purposes but are envisioned now to support drug development decisions), to ethical (a risk to privacy as health information and patient data sets are aggregated), to analytical (selecting which RWD are appropriate to informing a particular decision, and reducing the risk of bias). The growing emphasis on the use of RWD is expected to improve our understanding of these data sets as well as fill existing knowledge gaps.

Data quality

Since most RWD are not collected for research purposes, the data collection is episodic, reactive, and at best offers a partial picture. As a result, RWD is in general messy and sparse, and requires statistically rigorous and valid methods to clean the data and correct inconsistencies. Careful data curation, using both structured and unstructured data, is especially important for precision therapeutics in oncology, where often crucial information related to molecular biomarkers or end-points data can be missing. Missing data may also need to be filled by linking to alternative data sources.¹⁸

Analysts must also identify and adjust for confounding factors such as demographics, socioeconomic and insurance status, disease severity, comorbidities, concomitant medications, and genetic predispositions to certain conditions before conducting in-depth analyses. RWD is also subject to selection bias, as cohort selection and treatment decisions in clinical practice are not random. Therefore, following appropriate guidelines on design and validation of RWE studies can help in minimizing some of the sources of bias and inconsistencies.¹⁹

Interoperability

In addition, standards for the development and maintenance of data assets have not yet caught up with the rapid evolution of RWD. A lack of interoperability between real-world databases creates difficulties for combinatorial analysis and collaboration between data holders. Even within individual organizations there is often a lack of consolidated or centralized data storage, leading to difficulties in analyzing data across different data sets. Overall, there is a need to implement standardization and maintain robust quality assurance (QA) / quality control (QC) practices to support data robustness. The not-for-profit organizations such as Health Level 7 (<http://www.hl7.org/>) and the IMEDS initiative¹¹ are creating standards for electronic health data and promoting interoperability among systems. In the future, advances in data standardization, interoperability, and linkage techniques are anticipated to further enable disparate data sources to converge into a single platform for more seamless and efficient analytics. One such example is the Administrative Data Research Network (<https://adrn.ac.uk>) established in the United Kingdom in 2012, which allows access to linked, deidentified government data for social, economic, environmental, and health research. Commercial organizations like QuintilesIMS and Flatiron Health have linked numerous community healthcare practice data sets to provide larger, more robust data analytic platforms for research.

Analytical platforms

Even though the adoption and use of EHRs has grown significantly, extracting meaningful data from EHRs in an accurate and efficient manner remains challenging. This is due to the fact that a significant portion of high-value clinical information in EHRs is often stored in unstructured, free-text clinical documents that are inaccessible to algorithms and requires layers of preprocessing. For example, even a frequently used metric such as the ankle-brachial index (ABI)—a “quantitative” data point for defining peripheral arterial disease (PAD), is typically embedded in the text of radiology reports, hidden from structured data analytics tools.

Natural Language Processing (NLP) methods provide one approach for extraction and conversion of unstructured information from clinical text data to structured observations—such as Karnofsky and Mini Mental State Examination scores for determination of disease severity and functional status in oncology and Alzheimer’s disease patients, extraction of findings such as ejection fraction from laboratory reports, biomarker information from pathology reports, as well as in the assessment of patient characteristics such as emotional and social behaviors from physician notes. Further,

predefined fields in EHR (e.g., problem lists, past history, or test result fields) capture only certain disease information and may miss the trends of other prevalent, but unlisted, health conditions. NLP can be a powerful tool to extract symptoms from physician notes or textual data from lab reports to enable identification of those trends/conditions, thus complementing the assessments using structured data.

NLP uses a combination of linguistics, pattern recognition, and machine-learning techniques to extract context-appropriate information from reports to derive insights from a clinical text narrative. In addition to routine text-mining and extraction capabilities, more sophisticated machine-learning algorithms can also be layered on top of NLP for automated assignment of diagnosis codes as well as in the identification of patient cohorts, using predefined inclusion/exclusion criteria, based on information contained in clinical notes. More recently, application of deep learning methods have resulted in impressive advances in NLP, especially in the development of unsupervised models using recurrent neural networks and autoencoders that reduce dependence on high-quality, manual annotations of text data.^{20–22} These methods allow algorithms to learn high-level abstractions from clinical data and notes when concepts are not mentioned explicitly. Availability of large volumes of real-world clinical data enables the training, development, and validation of new algorithms. However, before clinical notes can be used for research, adequate precautions must be taken to ensure all HIPAA (Health Insurance Portability and Accountability Act)-defined protected health information (PHI) elements are removed or anonymized to deidentify the data set. Overall, NLP methods can add significant analytical capabilities that can increase the utility of EHR data.

Beyond NLP, advances in machine learning have enabled new approaches for prediction of disease onset, future diseases,^{23,24} and risk scores^{25,26} from longitudinal EHRs and lab tests, identification of treatment course based on patient outcomes,²⁷ and in image recognition for classification of radiology and pathology images.^{28,29} Methods for assessing disease heterogeneity and predicting patient outcomes, given the information about a patient, their history, and individual-specific variability, have demonstrated capabilities to include both observed as well as latent features extracted from messy, multivariate EHR data.³⁰ Advanced analytics using machine learning on longitudinal RWD has the potential to inform and reframe drug development and clinical trial design strategy—through patient stratification into subgroups based on disease subtypes, drug treatment efficacy, progress, side effects, and toxicity profiles—by shifting from presumption of a single disease to multiple, related diseases. As machine-learning algorithms and frameworks continuously advance, there will be improvements in the ability of these models to learn continuously as new information emerges either in the form of additional data sources or updated treatment guidelines.³¹

APPLICATIONS OF RWD FOR PHARMACEUTICAL R&D

The following section will discuss application of RWE in clinical development and introduce PK-PD-PE modeling to

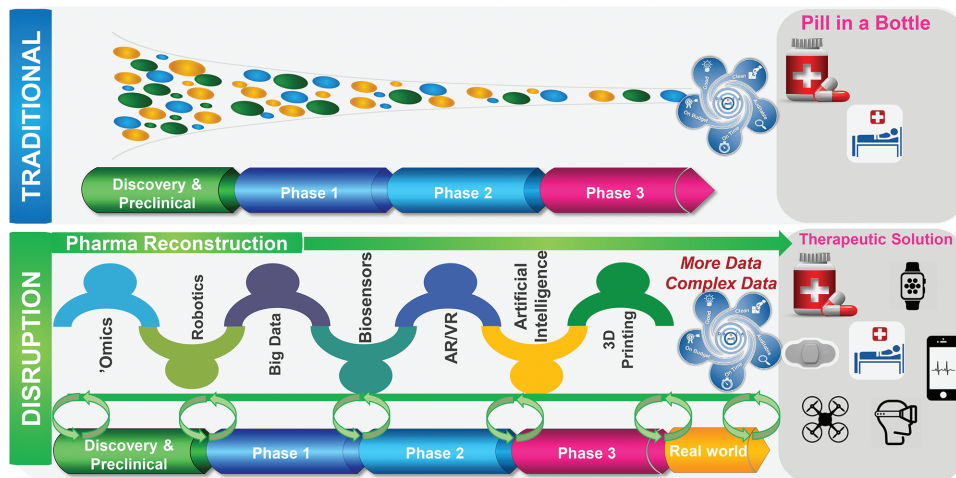


Figure 3 Innovation in the traditional drug development paradigm moving from the randomized controlled trial to gain regulatory approval to an all-encompassing collection of real-world evidence in the context of a therapeutic solution.

highlight opportunities for early value assessment as part of clinical development.

Transforming clinical development through RWE

RWD presents an opportunity to disrupt an inefficient pharmaceutical business model stifled by a high failure rate, rising development costs,³² and increased pressure from various stakeholders including regulators, payers, prescribers, and patients.³³

To discuss future opportunities for leveraging RWD, a traditional drug discovery/development model is contrasted to propose a new model for drug discovery/development in the digital era (**Figure 3**). In the traditional model, the goal is to obtain regulatory approval for a “pill-in-a-bottle” followed by real-world considerations (differentiation, value, integration with other modalities, etc.). The proposed new model leverages the advances in big data such as multi-omics, sensor devices and technology, imaging, and other relevant data in health ecosystem wellness apps, social networking, etc., to garner the RWE while still gaining the necessary data for regulatory approval. The new types of data can also allow pharmaceutical companies to change the product identity from merely selling a pill-in-a-bottle, to offering a comprehensive *therapeutic solution*. The new goal of drug developers should be to develop an integrated therapeutic strategy that takes into consideration real-world usage of a therapeutic solution in the context of digital devices, behavioral interventions, and other therapeutic options, which may or may not include a pill.

In addition to eventually changing the treatment paradigm from a product identity to a therapeutic solution that the pharmaceutical industry can deliver in conjunction with health-care providers, there are some immediate benefits to the current business model that can be leveraged by using RWD during clinical development:

Improve clinical trials execution and success

- *Optimize drug dosing through adherence measurement:* Adherence in clinical trials remains largely

unmonitored and is assumed to be high, contrary to evidence.^{34,35} There have been several developing approaches to monitor and potentially improve adherence.³⁶ As an example, with a combination of sensors and a mobile technology interface, Otsuka Pharmaceuticals and Proteus Digital Health developed the first FDA-approved real-time adherence monitoring platform that will allow physicians, caregivers, and others involved in care delivery to monitor adherence for a patient, as well as relate drug intake to activity and other vital measures.³⁷ The use of electronic monitoring devices coupled with provision of feedback to patients of their recent dosing histories is an evidence-based approach to enhancing patient adherence to medications.³⁸ These types of data are rarely available during clinical development and could be used to optimize dosing based on adherence profiles observed in clinical trials. Collection of adherence data in clinical trials and the real world would strengthen the predictive analytics on generating the patient response given relevant data with respect to drug intake and action.^{39,40}

- *Leverage digital biomarkers/phenotypes to increase trial success by enriching clinical trial population:* Understanding the determinants of placebo and drug response has been of interest to pharmaceutical sponsors and regulators.⁴¹ However, traditionally these determinants have been limited to demographics and certain baseline variables collected in a typical clinical study. With the availability of dynamic measures, as well as detailed data on activity, sleep, vitals, circadian rhythm, and behavior, speech, etc., there exists an opportunity to better understand these determinants and optimize the clinical trial population. For example, one of the primary reasons for psychiatry trial failures is believed to be inclusion of refractory patients and partial responders. However, traditional inclusion/exclusion criteria are unable to differentiate these types of patients. It is possible that digital phenotype of patients based

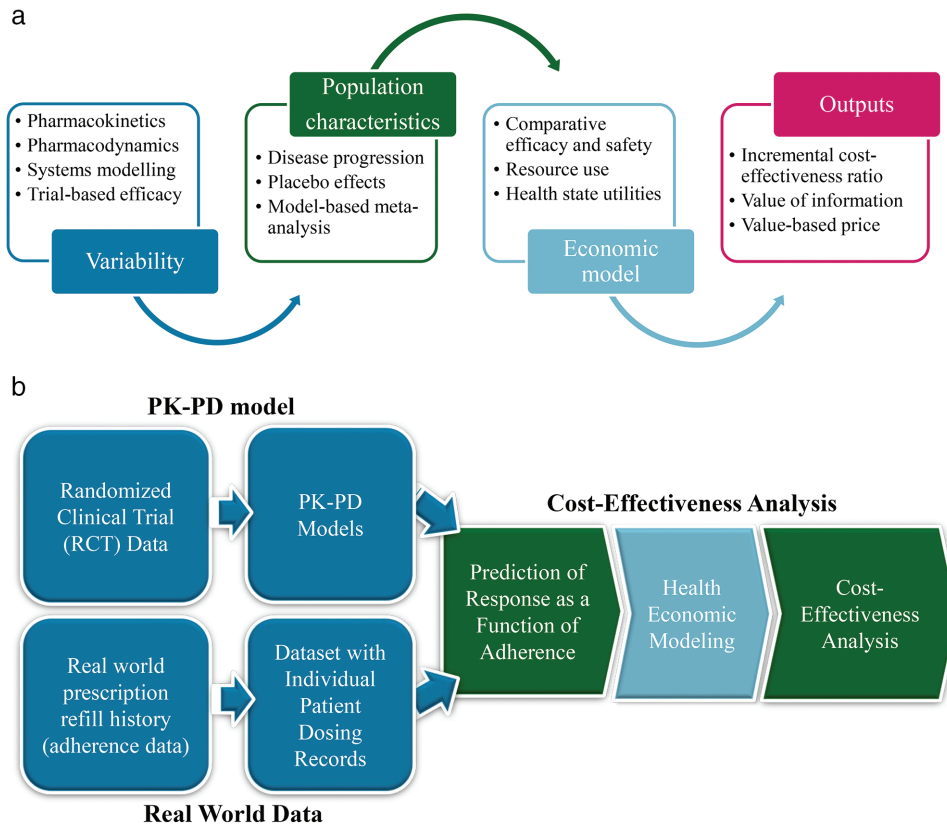


Figure 4 (a) Conceptual framework of a pharmacometric-pharmacoeconomic model; and (b) an illustrative application in the context of medication adherence.

on activity, sleep, behavior, etc., might be able to offer a better view of potential response. These data types have the potential to be leveraged to enrich the appropriate population studied in a given clinical trial.⁴²

Real-world differentiation for products

- *Predictive analytics of real-world effect given clinical trial response:* Real-world effect of drugs has always been debated with regard to the generalizability of the response observed in a clinical trial conducted with a specific and limited patient population.⁵ Modeling and simulation is an often-overlooked tool that can be used early on in the development to project real-world performance of drugs under different combinations of patient characteristics and/or conditions (such as adherence) not explicitly studied in clinical trials. One such example is an integration of adherence rates taken from RWD with PK-PD modeling to inform go/no-go decisions.⁴³ In this example, as shown in the schematic in **Figure 4b**, dose and response data from RCTs were modeled using the PK-PD modeling approach. Separately, RWD on patient adherence from a large database with prescription refill history was transformed into an individual patient level database (by making some assumptions) to bring it in the same format as the PK-PD database (to allow simulations by using the PK-PD

modeling software). Subsequently, the PK-PD model based on RCT was applied to the individual patient database created from the prescription refill history to simulate the clinical responses under different degrees of adherence in the real-world setting. These response predictions were entered into a health economic model to assess what degree of adherence improvement would result in a clinically meaningful improvement in clinical outcomes that are also cost-effective. Such approaches can be used to analyze the economic value of therapeutic interventions directed at improving adherence, as well as to assess the need of acquiring more information. Predictive analytics is potentially the only way to project real-world performance before a drug is released to a broader population. Such model-based outputs for future performance can be effectively leveraged to support realistic future value during clinical development.

- *Leverage digital biomarkers/phenotypes to generate hypotheses for real-world differentiation:* With a broader definition of RWD noted above, the ability to collect data in both a healthy and diseased state could enable generation of new hypotheses during clinical trials. The measurement of activity, vitals, sleep, behavior, speech, etc., in real time makes it possible to learn more about the therapeutic or adverse effect of drugs and can be utilized to show differentiation among drugs. While these

data can be collected on an exploratory basis and likely have minimal impact on the primary results of a clinical trial, they provide a big opportunity to generate additional hypotheses and test those in prospective clinical trials.

PK-PD-PE modeling for early prediction of outcomes

Model-based clinical drug development uses pharmacometric (quantitative pharmacology) approaches to inform trial design and optimize compound development strategies.⁴⁴ This is achieved by integrating PK, PD, and clinical evidence using empirical or mechanism-based modeling to predict efficacy and safety outcomes from simulated clinical trials. Such approaches have been used to facilitate the identification of trial design inefficiencies, adjust for nonadherence and dropout, for exploration of the effects of different dosing regimens, for consideration of specific populations,⁴⁵ and offer valuable insights for future studies with the aim of reducing late-stage failure and improving the efficiency of drug development.^{46–48}

A natural extension to pharmacometric analyses, exploiting the structural relationship between dose and response and accounting for the statistical uncertainty, is to link with PE models (**Figure 4a**) that consider the resource constraints of payers of health care.⁴⁹ Collectively, these models can be defined as PK-PD-PE models. PE models assess the incremental costs per quality-adjusted life years (QALY) gained for a given intervention. If the cost per QALY is below a predefined threshold (often accepted as £30,000; €50,000; \$100,000 per QALY in the United Kingdom, European Union, and United States, respectively) then the intervention is considered cost-effective. This evaluative framework has the potential to improve methods for strategic, clinical, and pricing decisions during phase II/III drug development and offers advantages over standard (empirical) PE models during these phases of clinical drug development.^{50,51} The determination of a value-based price, for instance, will inform whether further development is commercially viable—it may be appropriate to halt the development of a drug with no prospect of achieving a value-based price. Alternatively, a value of information analysis might be carried out, taking advantage of the PK/PD uncertainty as well as economic parameters derived from RWE. This can inform go/no-go decisions on whether a trial is worth undertaking, based on the expected net trade-off of the benefits of the trial in relation to its costs.^{52,53}

Published examples of PK-PD-PE models are limited to application in relation to rituximab for follicular lymphoma,⁵⁴ pharmacogenetics-guided warfarin dosing,⁵⁵ eribulin for castration-resistant prostate cancer,⁵⁶ a hypothetical drug (potentially representative of a drug in development) for chronic obstructive pulmonary disease,⁵⁷ and oseltamivir for the management of influenza pandemics.⁵⁸ While these studies support the proof of principle of the method, the application of pharmacometrics in pharmacoeconomic evaluation and HTA is very much in its infancy. However, the limited published evidence to date suggests utility in many contexts, most evidently in integrating RWE in early-stage evaluation. These relate to: (i) providing early indications of cost-effectiveness before large-scale trial data become

available; (ii) directing future research based on the cost of reducing uncertainty; (iii) assessing subgroups, dosing schedules, and protocol deviations; (iv) informing strategic research and development along with pricing decisions; and (v) estimating the cost-effectiveness of complex pharmaceutical interventions (such as pharmacogenetics testing).

Key challenges to the further advancement of PK-PD-PE model development and application, however, include overcoming different modeling paradigms in pharmacometrics and health economic evaluation, a need for increased acceptance of model-based drug development through to pricing and reimbursement to inform critical-stage decision-making, and further evidence on the validity and reliability of complex and computationally intensive models. These may be overcome through closer integration and collaboration between the disparate disciplines of pharmacometrics, clinical pharmacology, and health economics / outcomes research, which may be achieved through colocation and appropriate training of discipline-agnostic biostatisticians and mathematical modelers. This should ultimately lead to greater acceptance of model-based drug development incorporating RWE and PK-PD-PE, specifically, in the drug development process.

DISCUSSION/FUTURE DIRECTIONS

The application of RWD is growing in several areas including, but not limited to, supporting decisions in drug development and medical practice, as well as decisions by HTA agencies and regulators. Greater adoption of RWE in these areas is closing the gaps between evidence used to support drug approval or reimbursement decisions and evidence used by the medical community. RWE is also augmenting the information healthcare providers' use for clinical decision-making by adding information that is not collected as part of RCTs such as benefit–risk in underrepresented patients with comorbidities or in the elderly. Overall, enhanced use of RWE is anticipated to enable the development of medicines with unambiguous value to patients, make the selection of medical treatments more effective and efficient, and reduce the time for bringing novel drugs to market.

Impact on drug development

RWE will not replace the need for data from traditional trials; however, technologies supporting RWD are enabling far richer and more diverse information to be collected during drug development. Traditional drug development restricts capturing of quality of life measurements within RCTs to when instruments like patient-reported outcomes (PROs) are administered, whereas RWD allows for more efficient and unrestricted capturing of these data in larger volumes and novel settings. Incorporation of technologies for RWD collection (e.g., wearables) (**Figure 5a**) and use of RWE in drug development decision-making may also inspire innovation in clinical programs and trial designs. This has the potential to favorably impact the efficiency of planning and operations, and ultimately pave the path for reductions in the cost of drug development. To realize the full impact of RWE on drug development, the healthcare community will need to make fundamental changes in the way clinical information is collected. Due to its historical use for billing and claims management

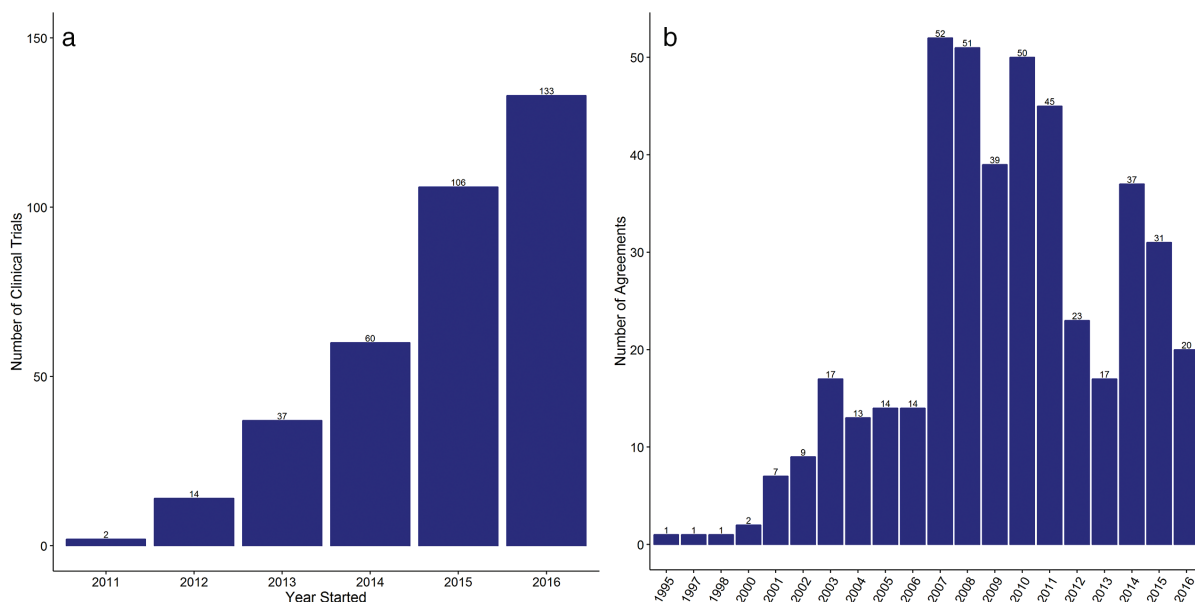


Figure 5 Increasing application of RWD in drug development through (a) utilization of wearables in clinical trials and with payer assessments through (b) the use of value-based pricing agreements. (a) Data extracted from <https://clinicaltrials.gov> using the search term “wearables.” (b) Data extracted from the University of Washington School of Pharmacy Performance Based Risk Sharing database.

the current system was not designed for research purposes. It therefore lacks some of the critical details that researchers require, for example, detailed clinical information on such things as tumor biomarkers, PROs or similarly nuanced, but valuable clinical insights. Often if this detailed clinical information exists, it resides in the unstructured data, making analytical tools such as NLP even more important. The scientific community will need to continue working on improving the methodologies for handling complex and unstructured data.

Impact on regulatory decisions

The use of RWE in regulatory decisions about efficacy has been historically very limited; however, the narrative is changing. This has been spurred by the recent series of publications and presentations, as well as statutory changes enforced by the FDA. In the short period from December 2016 to the present, the FDA has published opinion articles^{5,9} and actively participated in public discussions⁵⁹ to define and promote the use of RWD and RWE as an evidentiary standard for drug approval. Additionally, the recently legislated 21st Century Cures Act and the new user fee laws (Prescription Drug User Fee Act VI) provide enough opportunity for expanding the use of RWE for demonstrating efficacy for drug approval and mandates the FDA to hold public workshops and develop draft guidance documents aimed at enhancing the use of RWE in regulatory decision-making within the next 5 years. In fact, the current statutes provide enough latitude for incorporation of RWE in decision-making, as remarked by the FDA commissioner Dr. Scott Gottlieb at a recently held workshop: “there is nothing in our statute or regulations that prevent FDA from using a broad range of informative sources of evidence. On the contrary many of our statutory responsibilities boil down to one principal calculus, what do we know and how do we balance benefits and

risks based on the fullest possible information.”⁵⁹ Dr. Gottlieb also remarked on the important role RWD can play in meeting the postmarket study requirements and approval of new indications for already marketed drugs. RWE is also playing a key role in approval of medical devices strengthened by the recently finalized FDA guidance. Since 2015 alone, approval of at least eight new medical devices and expanded use of at least six technologies have relied on evidence derived from RWD.⁵⁹ Other regulatory agencies such as the EMA⁶⁰ and PMDA⁶¹ are also taking initiatives to expand the use of RWD. For example, the adaptive pathways approach implemented by the EMA aims to provide timely access for patients to new medicines and seeks to involve patients and health technology assessment agencies in discussions during drug development. Overall, the evolving changes in the regulatory landscape regarding use of RWD hold promise for a new future.

Impact on medical practice

Experiences with medication use in the real world has always influenced decisions in medical practice and is the foundation of evidence-based medicine (EBM). However, over reliance on inferences from RCT and the ensuing inability to individualize for specific clinical scenarios has often been seen to limit the application of EBM.⁶² The ability to collect richer data from the real world both pre- and postapproval and advanced analytics has promise to overcome these limitations. This change would offer the opportunity for integration of multiple sources of data and from many more patients than one provider sees on an individual basis, which in turn could bring medical decisions specific to the individual characteristics of patients, with the potential of making health care more personalized and effective. Access to RWD can allow exploration of important clinical outcomes which may not be possible otherwise and it can also help with development and validation of

surrogate instruments for such outcomes for incorporation in clinical practice. RWD can also provide insights on rare safety events which could potentially be prevented. To capitalize on these obvious advantages, healthcare stakeholders including Aetna, Kaiser-Permanente, and Geisinger to name a few are increasingly applying “big data” to improve patient care guidelines and manage formularies.^{63,64}

Impact on payer assessments

The need to price drug products based on the value offered to patients is clearly reflected in value-based contracts between pharmaceutical companies and payers to link the price of a prescription drug to its clinical and economic performance (Figure 5b). As discussed in this article, collection of RWD is helping payers in economic assessments by providing information on outcomes in real-world settings (as opposed to controlled settings in clinical trials), by providing evidence of effectiveness (or lack of it) in population subgroups not adequately represented in randomized trials, and by aiding to the assessments of comparative effectiveness. A repository of patient characteristics data built using RWD also enables payers to run simulations for effectiveness or budget impact in populations that are not clinically tested or for which only limited data exist.

In conclusion, RWD and RWE have promise to strengthen the current ecosystem of data supporting healthcare decisions, and support transition into a new era of personalized, more effective, and more efficient health care. Collaboration among providers, patients, payers, drug companies, and other players in the healthcare system, aided by technology, would be necessary to unleash a new age of medicine.

Acknowledgments. We thank Dr. Josh Carlson at the University of Washington School of Pharmacy for providing the data on value-based pricing agreements (Figure 5b) from the performance based risk sharing database (<https://sop.washington.edu/department-of-pharmacy/performance-based-risk-sharing-database/>).

Funding. D.H. acknowledges the funding contribution of the Medical Research Council North-West Hub for Trial Methodological Research (MR/K025635/1).

Conflict of Interest. P.J. is an employee of Otsuka Pharmaceutical Development and Commercialization. L.J. and V.C. are employees of and hold stock in Merck & Co., Inc., Kenilworth, NJ, USA.

- Rothwell, P.M. External validity of randomised controlled trials: “To whom do the results of this trial apply?” *Lancet* **365**, 82–93 (2005).
- Oyinlola, J.O., Campbell, J. & Kousoulis, A.A. Is real world evidence influencing practice? A systematic review of CPRD research in NICE guidances. *BMC Health Serv. Res.* [Internet]. (2016) Dec [cited 2017 Oct 17];**16**(1). Available from: <http://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-016-1562-8>
- Bate, A., Juniper, J., Lawton, A.M. & Thwaites, R.M.A. Designing and incorporating a real world data approach to international drug development and use: what the UK offers. *Drug Discov. Today*. **21**, 400–405 (2016).
- Heatherington, A.C., Kasichayanula, S. & Venkatakrishnan, K. How well are we applying quantitative methods to reverse translation to inform early clinical development? *Clin. Pharmacol. Ther.* **103**, 174–176 (2018).
- Sherman, R.E. et al. Real-World Evidence — What Is It and What Can It Tell Us? *N. Engl. J. Med.* **375**, 2293–2297 (2016 Dec 8).
- Izmailova, E.S., Wagner, J.A. & Perakslis, E.D. Wearable devices in clinical trials: hype and hypothesis. *Clin. Pharmacol. Ther.* [Internet]. (2017) Dec 5 [cited 2018 Mar 9]; Available from: <http://doi.wiley.com/10.1002/cpt.966>
- Wicks, P. et al. Sharing health data for better outcomes on PatientsLikeMe. *J. Med. Internet. Res.* **12**, e19 (2010).
- Fleurence, R.L. et al. Launching PCORnet, a national patient-centered clinical research network. *J. Am. Med. Inform. Assoc.* **21**, 578–582 (2014).
- Jarow, J.P., LaVange, L. & Woodcock, J. Multidimensional evidence generation and FDA regulatory decision-making: defining and using “real-world” data. *JAMA* **318**, 703–704 (2017).
- Platt, R. et al. The new Sentinel Network — improving the evidence of medical-product safety. *N. Engl. J. Med.* **361**, 645–647 (2009).
- Califf, R.M. Introducing IMEDS, a public-private resource for evidence generation [Internet]. Available from: <https://blogs.fda.gov/fdavoices/index.php/2017/01/introducing-imes-a-public-private-resource-for-evidence-generation/>
- Real World Evidence Collection; Agenda item 4. In: 4th Meeting of the STAMP Expert Group [Internet]. Brussels: European Commission; (2016). Available from: https://ec.europa.eu/health/sites/health/files/files/committee/stamp/2016-03_stamp4/4_real_world_evidence_background_paper.pdf
- Eichler, H.-G. et al. “Threshold-crossing”: a useful way to establish the counterfactual in clinical trials? *Clin. Pharmacol. Ther.* **100**, 699–712 (2016).
- Bognar, K. et al. The role of imperfect surrogate endpoint information in drug approval and reimbursement decisions. *J. Health Econ.* **51**, 1–12 (2017).
- Vestbo, J. et al. Effectiveness of fluticasone furoate-vilanterol for COPD in clinical practice. *N. Engl. J. Med.* **375**, 1253–1260 (2016).
- Carlson, J., Garrison, L.P. & Sullivan, S.D. Paying for outcomes: innovative coverage and reimbursement schemes for pharmaceuticals. *J. Manag. Care Pharm.* **15**, 683–687 (2009).
- AstraZeneca’s CVD-REAL study shows SGLT-2 inhibitors significantly reduced hospitalizations for heart failure and death versus other type-2 diabetes medicines [Internet]. (2017) [cited 2017 Oct 6]. Available from: <https://www.astrazeneca-us.com/media/press-releases/2017/astrazeneca-cvd-real-study-shows-sgl-2-inhibitors-significantly-reduced-hospitalizations-for-heart-failure-and-death-versus-other-type-2-diabetes-medicines-03192017.html>
- Mikhsad, R.A. & Abernethy, A.P. Harnessing the power of real-world evidence (RWE): a checklist to ensure regulatory-grade data quality. *Clin. Pharmacol. Ther.* **103**, 202–205 (2018).
- White, R. Building trust in real-world evidence and comparative effectiveness research: the need for transparency. *J. Comp. Eff. Res.* **6**, 5–7 (2017).
- Dubois, S. & Romano, N. Learning effective embeddings from medical notes [Internet]. Stanford University; [cited 2017 Sep 25]. Available from: <https://web.stanford.edu/class/cs224n/reports/2744372.pdf>
- Jacobson O, Dallanis H. Applying deep learning on electronic health records in Swedish to predict healthcare-associated infections. Proceedings of the 15th Workshop on Biomedical Natural Language Processing [Internet]. Association for Computational Linguistics; (2016) [cited 2017 Sep 25]. Available from: <http://aclweb.org/anthology/W16-2926>
- Shickel, B., Tighe, P., Bihorac, A. & Rashidi, P. Deep EHR: a survey of recent advances on deep learning techniques for electronic health record (EHR) analysis. *Rev. J. Biomed. Health Inform.* [Internet]. (2017 Jun 12) [cited 2107 Sep 25]; Available from: <https://arxiv.org/abs/1706.03446>
- Miotto, R., Li, L., Kidd, B.A. & Dudley, J.T. Deep patient: an unsupervised representation to predict the future of patients from the electronic health records. *Sci. Rep.* [Internet]. (2016) [cited 2017 Sep 25];**6**. Available from: <http://www.nature.com/articles/srep26094>
- Razavian, N., Marcus, J. & Sontag, D. Multi-task prediction of disease onsets from longitudinal laboratory tests. Los Angeles, CA (2016) [cited 2017 Sep 25]. Available from: <https://arxiv.org/abs/1608.00647>
- Dyagilev, K. & Saria, S. Learning (predictive) risk scores in the presence of censoring due to interventions [Internet]. *Mach. Learn. J.* (2015) [cited 2017 Sep 25]. Available from: <https://arxiv.org/abs/1507.07295>
- Henry, K.E., Hager, D.N., Pronovost, P.J. & Saria, S. A targeted real-time early warning score (TREWScore) for septic shock. *Sci. Transl. Med.* **7**, 299ra122–299ra122 (2015).
- Zhao, Y. et al. Exploration of machine learning techniques in predicting multiple sclerosis disease course. Ramagopalan SV, (ed.). *PLoS One* **12**, e0174866 (2017).
- Gulshan, V. et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* **316**, 2402 (2016).
- Esteve, A. et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **542**, 115–158 (2017).
- Schulam, P., Wigley, F. & Saria, S. Clustering longitudinal clinical marker trajectories from electronic health data: Applications to phenotyping and endotype discovery. In: Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence. Austin, TX; (2015).
- Dorajoo, S.R. & Chan, A. Implementing clinical prediction models: pushing the needle towards precision pharmacotherapy. *Clin. Pharmacol. Ther.* **103**, 180–183 (2018 Feb).
- Karamelic, J. et al. Financial aspects and the future of the pharmaceutical industry in the United States of America. *Mater. Socio. Med.* **25**, 286 (2013).
- van Harten, W. & IJzerman, M.J. Responsible pricing in value-based assessment of cancer drugs: real-world data are an inevitable addition to select meaningful new cancer treatments. *ecancermedicalscience* [Internet]. (2017)Sep 11

- [cited 2017 Oct 12];11. Available from: <http://www.ecancer.org/journal/editorial/71-responsible-pricing-in-value-based-assessment-of-cancer-drugs-real-world-data-are-an-inevitable-addition-to-select-meaningful-new-cancer-treatments.php>
34. Fossler, M.J. Patient adherence: Clinical pharmacology's embarrassing relative. *J Clin Pharmacol*. **55**, 365–367 (2015).
 35. Vrijens, B. & Urquhart, J. Methods for measuring, enhancing, and accounting for medication adherence in clinical trials. *Clin. Pharmacol. Ther.* **95**, 617–626 (2014).
 36. Whalley Buono, E. et al. Coming full circle in the measurement of medication adherence: opportunities and implications for health care. *Patient Prefer. Adher.* **11**, 1009–1017 (2017).
 37. FDA approves pill with sensor that digitally tracks if patients have ingested their medication [Internet]. FDA News Release. (2017) [cited 2018 Mar 15]. Available from: <https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm584933.htm>
 38. for the ABC project team, Demonceau, J. et al. Identification and assessment of adherence-enhancing interventions in studies assessing medication adherence through electronically compiled drug dosing histories: a systematic literature review and meta-analysis. *Drugs* **73**, 545–562 (2013).
 39. Peters-Strickland, T. et al. Usability of a novel digital medicine system in adults with schizophrenia treated with sensor-embedded tablets of aripiprazole. *Neuropsychiatr. Dis. Treat.* **12**, 2587–2594 (2016).
 40. Hill-McManus, D., Soto, E., Marshall, S., Lane, S. & Hughes, D. Impact of non-adherence on the safety and efficacy of uric acid-lowering therapies in the treatment of gout: Impact of non-adherence on uric acid-lowering therapies in gout. *Br. J. Clin. Pharmacol.* [Internet]. (2017) Oct 10 [cited 2017 Oct 31]; Available from: <http://doi.wiley.com/10.1111/bcp.13427>
 41. Jain, L., Jadhav, P. & Gobburu, J. Implications of choosing a correlation structure on model selection and parameter estimation. *Int. J. Clin. Pharmacol. Ther* [Internet]. (2014) Apr 14 [cited 2017 Oct 12]; Available from: <http://www.dustri.com/index.php?id=8&artid=11405&doi=10.5414/CP202058>
 42. Temple, R. Complexities in drug trials: enrichment, biomarkers and surrogates. *Biomark. Med.* **2**, 109–112 (2008).
 43. Jain, L. et al. Integration of PK-PD and health economic modeling to assess cost-effectiveness of improving adherence in real world setting. In *American Society for Clinical Pharmacology and Therapeutics Annual Meeting*. San Diego, CA (2016).
 44. Kimko, H. & Pinheiro, J. Model-based clinical drug development in the past, present and future: a commentary: Modelling and simulations of clinical trials. *Br. J. Clin. Pharmacol.* **79**, 108–116 (2015).
 45. Holford, N., Ma, S.C. & Ploeger, B.A. Clinical trial simulation: a review. *Clin. Pharmacol. Ther.* **88**, 166–182 (2010).
 46. Milligan, P.A. et al. Model-based drug development: a rational approach to efficiently accelerate drug development. *Clin. Pharmacol. Ther.* **93**, 502–514 (2013).
 47. Grasela, T.H. & Slusser, R. Improving productivity with model-based drug development: an enterprise perspective. *Clin. Pharmacol. Ther.* **88**, 263–268 (2010).
 48. Lalonde, R.L. et al. Model-based drug development. *Clin. Pharmacol. Ther.* **82**, 21–32 (2007).
 49. Hughes, D.A. & Walley, T. Economic evaluations during early (phase II) drug development: a role for clinical trial simulations? *PharmacoEconomics* **19**, 1069–1077 (2001).
 50. Miller, P. Role of pharmacoeconomic analysis in R&D decision-making: when, where, how? *PharmacoEconomics* **23**, 1–12 (2005).
 51. Hartz, S. & John, J. Contribution of economic evaluation to decision-making in early phases of product development: A methodological and empirical review. *Int. J. Technol. Assess. Health Care*. **24**, 465–472 (2008).
 52. Pink, J., Lane, S. & Hughes, D.A. Mechanism-based approach to the economic evaluation of pharmaceuticals: pharmacokinetic/pharmacodynamic/pharmacoeconomic analysis of rituximab for follicular lymphoma. *PharmacoEconomics* **30**, 413–429 (2012).
 53. Breeze, P. & Brennan, A. Valuing trial designs from a pharmaceutical perspective using value-based pricing. *Health Econ.* **24**, 1468–1482 (2015).
 54. Pink, J., Pirmohamed, M., Lane, S. & Hughes, D.A. Cost-effectiveness of pharmacogenetics-guided warfarin therapy vs. alternative anticoagulation in atrial fibrillation. *Clin. Pharmacol. Ther.* **95**, 199–207 (2014).
 55. Hamberg, A.-K. et al. A pharmacometric model describing the relationship between warfarin dose and INR response with respect to variations in CYP2C9, VKORC1, and age. *Clin. Pharmacol. Ther.* **87**, 727–734 (2010).
 56. van Hasselt, J. et al. Integrated simulation framework for toxicity, dose intensity, disease progression, and cost effectiveness for castration-resistant prostate cancer treatment with eribulin: prostate cancer simulation framework. *CPT Pharmacometrics Syst. Pharmacol.* **4**, 374–385 (2015).
 57. Slejko, J.F., Willke, R.J., Ribbing, J. & Milligan, P. Translating pharmacometrics to a pharmacoeconomic model of COPD. *Value Health* **19**, 1026–1032 (2016).
 58. Kamal, M.A. et al. Interdisciplinary pharmacometrics linking oseltamivir pharmacology, influenza epidemiology and health economics to inform antiviral use in pandemics: Linking pharmacology to influenza epidemiology and health economics. *Br. J. Clin. Pharmacol.* **83**, 1580–1594 (2017).
 59. Workshop 1: Incentives. In: *Examining the Impact of Real-World Evidence on Medical Product Development: A Workshop Series* [Internet]. Washington DC: The National Academies of Sciences Engineering Medicine; Health and Medicine Division; Available from: <http://www.nationalacademies.org/hmd/Activities/Research/DrugForum/2017-SEP-19.aspx>
 60. Cave, A. What are the real-world evidence tools and how can they support decision-making? [Internet]. *European Medicines Agency-EuropaBio information day*; (2016) Nov 22 [cited 2017 Oct 6]; London, UK. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Presentation/2016/12/WC500217732.pdf
 61. Kondo, T. "Rational Medicine" Initiative [Internet]. (2017) [cited 2017 Oct 6]. Available from: <https://www.pmda.go.jp/english/about-pmda/0012.html>
 62. Sur, R. & Dahm, P. History of evidence-based medicine. *Indian J. Urol.* **27**, 487 (2011).
 63. Erskine, A., Karunakaran, B., Slotkin, J. & Feinberg, D. How Geisinger Health System uses big data to save lives. *Harvard Business Review* [Internet]. (2016) Dec 15 [cited 2017 Oct 6]; Available from: <https://hbr.org/2016/12/how-geisinger-health-system-uses-big-data-to-save-lives>
 64. Marcum, C. The Rise of Big Data in Health Care [Internet]. *Institute for Health Policy*. (2014) [cited 2017 Oct 6]. Available from: <https://www.kpihp.org/how-big-data-can-inform-healthcare-decisions/>

© 2018 The Authors. Clinical and Translational Science published by Wiley Periodicals, Inc. on behalf of American Society for Clinical Pharmacology and Therapeutics. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.