

Clinical Pharmacology Applications of Real-World Data and Real-World Evidence in Drug Development and Approval—An Industry Perspective

Rui Zhu^{1,*,*}, Bianca Vora¹ , Sujatha Menon², Islam Younis³ , Gaurav Dwivedi⁴, Zhaoling Meng⁵, Amita Datta-Mannan⁶, Pooja Manchandani⁷, Satyaprakash Nayak² , Brinda K. Tammara², Parag Garhyan⁸, Shahed Iqbal⁹, Simon Dagenais¹⁰ , Pascal Chanu¹¹ , Arnab Mukherjee², Cyrus Ghobadi^{6*,*} and International Consortium for Innovation and Quality in Pharmaceutical Development (IQ) Real-World Data Working Group

Since the 21st Century Cures Act was signed into law in 2016, real-world data (RWD) and real-world evidence (RWE) have attracted great interest from the healthcare ecosystem globally. The potential and capability of RWD/RWE to inform regulatory decisions and clinical drug development have been extensively reviewed and discussed in the literature. However, a comprehensive review of current applications of RWD/RWE in clinical pharmacology, particularly from an industry perspective, is needed to inspire new insights and identify potential future opportunities for clinical pharmacologists to utilize RWD/RWE to address key drug development questions. In this paper, we review the RWD/RWE applications relevant to clinical pharmacology based on recent publications from member companies in the International Consortium for Innovation and Quality in Pharmaceutical Development (IQ) RWD Working Group, and discuss the future direction of RWE utilization from a clinical pharmacology perspective. A comprehensive review of RWD/RWE use cases is provided and discussed in the following categories of application: drug–drug interaction assessments, dose recommendation for patients with organ impairment, pediatric plan development and study design, model-informed drug development (e.g., disease progression modeling), prognostic and predictive biomarkers/factors identification, regulatory decisions support (e.g., label expansion), and synthetic/external control generation for rare diseases. Additionally, we describe and discuss common sources of RWD to help guide appropriate data selection to address questions pertaining to clinical pharmacology in drug development and regulatory decision making.

Real-world data (RWD) and real-world evidence (RWE) have gained broad attention in recent years, given their potential and capability to inform clinical drug development and regulatory decision making. Sources of RWD have evolved and expanded from the traditional electronic health records (EHRs), medical and pharmacy claims, disease and medical product registries, and observational clinical study data to include unstructured data sources (e.g., physician notes processed by natural language processing), novel data types (e.g., genomics data and diagnostic imaging), and patient/individual-generated data from wearable devices and

social media.¹ RWE derived from RWD is considered a complement/supplement to the gold-standard randomized controlled trials (RCTs), adding valuable features, such as greater patient heterogeneity and long-term outcomes in a typical care setting.² RWD/RWE have long been used by health authorities to assess post-approval drug safety (e.g., Sentinel System) and, more recently, they have also attracted users and participants in other parts of the healthcare ecosystem, such as biopharmaceutical companies, payers, providers, and patients.³ Legislation and regulatory policies are key factors that contributed to the increased interest and

¹Clinical Pharmacology, Genentech, Inc., South San Francisco, California, USA; ²Clinical Pharmacology, Pfizer Inc., Groton, Connecticut, USA; ³Clinical Pharmacology, Gilead Sciences, Inc., Foster City, California, USA; ⁴Quantitative Clinical Pharmacology, Takeda Development Center Americas, Inc., Cambridge, Massachusetts, USA; ⁵R&D Data and Data Science, Clinical Modeling & Evidence Integration, Sanofi, Cambridge, Massachusetts, USA; ⁶Exploratory Medicine & Pharmacology, Eli Lilly and Company, Indianapolis, Indiana, USA; ⁷Clinical Pharmacology and Exploratory Division, Astellas Pharma Global Development, Northbrook, Illinois, USA; ⁸Global PK/PD/Pharmacometrics, Eli Lilly and Company, Indianapolis, Indiana, USA; ⁹Biomarker Sciences, Gilead Sciences, Inc., Foster City, California, USA; ¹⁰Real World Evidence Center of Excellence, Pfizer, Inc., New York, New York, USA; ¹¹Clinical Pharmacology, Genentech/Roche, Inc., Lyon, France. *Correspondence: Rui Zhu (zhu.rui@gene.com)

[†]Authors who made substantial, equal contributions to this work.

[‡]IQ Real-World Data Working Group co-leads.

Received March 29, 2023; accepted June 19, 2023. doi:10.1002/cpt.2988

applications of RWD/RWE. The 21st Century Cures Act (2016) and the subsequent sixth Prescription Drug User Fee Act (PDUFA VI; 2017) required the US Food and Drug Administration (FDA) to assess the potential use of RWE to support new indication approvals and post-approval study requirements and to initiate activities to address key issues in using RWE to make regulatory decisions, respectively.^{4,5} Pursuant to this federal mandate, the FDA released the “Framework for FDA’s Real World Evidence Program” in 2018, which outlines how the FDA plans to implement its RWE Program and provides formal definitions for RWD and RWE.⁶ Since then, the FDA has issued a series of RWD/RWE-related draft guidance documents focusing on data sources, data standards, and regulatory considerations, and published related commentaries as well.^{7–9} Most recently, the FDA announced its Advancing Real-World Evidence Program, where one of the primary goals is to identify approaches for generating RWE in support of labeling claims, including new indications, populations, or dosing.¹⁰ Beyond the United States, global health authorities, such as the European Medicines Agency (EMA) and the Japan Pharmaceuticals and Medical Devices Agency (PMDA), have also shared their vision and perspectives on the application of RWD/RWE for regulatory decision making.^{11–13}

Besides regulatory applications, from an industry perspective, RWD can also benefit drug development at various stages. A recent publication gave a comprehensive review of RWD/RWE applications and their potential to inform decision making throughout the drug development process.¹⁴ However, to our knowledge, there are very few publications reviewing RWD/RWE applications in clinical pharmacology,^{15–17} especially from a drug development perspective.¹⁷ In this paper, we review and discuss how RWD/RWE have been, and can be, used to address questions relevant to clinical pharmacology from an industry point of view, based on recent publications from member companies in the International Consortium for Innovation and Quality in Pharmaceutical Development (IQ) RWD Working Group. Specifically, this paper reviews case examples mainly in the following three categories: (i) utilization of RWD to address core clinical pharmacology questions (e.g., drug–drug interaction (DDI) risk assessment and dose selection in special populations, including patients with organ impairments and pediatrics), (ii) utilization of quantitative clinical pharmacology and model-informed drug development (MIDD) as tools to analyze RWD and generate RWE (e.g., natural history characterization via disease progression modeling and prognostic factors identification), and (iii) utilization of RWE as a parallel source of evidence supplementing quantitative clinical pharmacology/MIDD and/or RCTs to support drug development decisions and regulatory approvals (e.g., label expansion). We also include multiple case examples in which RWD/RWE were used to generate synthetic/external control for rare diseases, which is a topic of broader scientific interest but still pertinent to clinical pharmacology. Key elements of each of the case examples are summarized in **Table 1**. We also illustrate example RWD applications to inform decisions throughout product development stages with clinical pharmacology-related applications highlighted in **Figure 1**. Finally, we discuss strengths, limitations, and key data elements for commonly used sources of RWD (e.g., EHRs, medical and

pharmacy claims, and patient registries) to guide the selection of appropriate RWD to investigate clinical pharmacology questions in drug development and approval.

GUIDE DRUG–DRUG INTERACTION ASSESSMENT

In clinical practice, adverse events caused by DDIs may lead to increased morbidity, hospitalization, prolonged hospital stays, or worsened outcomes.¹⁸ Assessing the risk of DDIs is crucial when designing clinical trials, because it can impact concomitant medication inclusion/exclusion criteria, patient recruitment, exposure variability, etc. As such, it is important to contextualize the elimination profile and potential DDIs for an investigational drug in its target indication prior to dosing patients. Toward this end, real-world polypharmacy data can be applied to strategically integrate findings from early translational medicine studies that evaluate an investigational drug’s DDI risk profile to define the inclusion/exclusion criteria for concomitant medications in clinical trials in patients. The value of this approach has been illustrated in an investigational molecule being developed for the treatment of psoriasis (unpublished findings). In this case, *in vitro* data indicated that the investigational molecule, which had not been evaluated in humans yet, was a CYP3A4 substrate and also a strong/moderate CYP3A4 inducer. So, there were two major DDI considerations: (i) strong/moderate CYP3A4 inhibitor or inducer concomitant medications administered with the investigational molecule could potentially alter the exposure of the molecule and thus confound the evaluation of the molecule’s exposure-response relationship, and (ii) the investigational molecule as a potential strong/moderate CYP3A4 inducer could lower the exposure of sensitive CYP3A4 substrate concomitant medications, which may reduce their efficacy. To have a realistic DDI risk assessment of the investigational molecule and inform clinical pharmacology strategy in clinical trial design, RWD from IBM MarketScan (a large, administrative US claims database) were used to derive the frequency of prescription claims for drugs that are considered strong/moderate CYP3A4 inhibitors/inducers and sensitive CYP3A4 substrates in the target patient population, and also derive their length of treatment to determine the chronicity of polypharmacy. It showed that, in patients with psoriasis, the majority of prescriptions of sensitive CYP3A4 substrates included corticosteroids, contraceptives, statins, and antibiotics, of which > 50% were used for > 90 days. Thus, it indicated a risk for the investigational molecule as a perpetrator with the potential to limit the exposure and effectiveness of some chronically used CYP3A4 sensitive substrates (e.g., oral birth control pills and cardiovascular drugs). RWD also showed that up to 15% of patients with psoriasis had claims for at least one drug with CYP3A4 induction or inhibition potential, however, CYP3A4 victim potential of the investigational drug was of relatively less concern because of the low chronic use (< 1%) of strong/moderate CYP3A4 inducer/inhibitor concomitant medications within the psoriasis population. Given these findings, a clinical DDI substudy was included in the first-in-human healthy volunteer trial of this investigational molecule to evaluate its CYP3A4 victim and perpetrator potentials. Results demonstrated that the molecule was not a strong CYP3A4 inducer, but it was a CYP3A4 victim with the potential of > 3-fold change in exposure in the

Table 1 Summary of case examples in RWD clinical pharmacology applications

Category	Reference	Objective	RWD source	Insight/evidence generated
Guide DDI assessments	Unpublished data from Eli Lilly and Company	Determine the frequency and treatment duration of concomitant medication use in the target patient population	IBM MarketScan Claims Database	Informed realistic DDI evaluations based on the target patient population in the real-world setting and informed inclusion/exclusion criteria of concomitant medications use for patient studies
	Duke <i>et al.</i> (2012) ¹⁹	Identify and evaluate novel DDIs by combining a literature discovery approach with RWD analysis	Indiana Network for Patient Care database	Five new drug–drug pairs were identified with increased risk of myopathy, thus indicating clinically relevant DDIs through CYP3A4 and CYP2D6 enzymes
	Lorberbaum <i>et al.</i> (2016) ²⁰	Develop a data-driven pipeline for discovering QT-DDIs using a combination of adverse event reports, EHRs, and laboratory experiments	FAERS and an EHR database at New York-Presbyterian/Columbia University Irving Medical Center	Unanticipated QT-DDIs can be efficiently identified via data mining and laboratory experiments. Combination therapy with ceftriaxone and lansoprazole was associated with increased risk of acquired long QT syndrome
	Yee <i>et al.</i> (2021) ²¹	Identify transporter-mediated DDIs of 25 small molecule drugs being evaluated for COVID-19 treatment via in-vitro experiments and real-world data	EHR data from UCSF Research Data Browser and Cerner Real World COVID-19 Database	Majority of the drugs were predicted to cause at least one clinical DDI. COVID-19 patients should be carefully monitored for adverse reactions likely to result from these DDIs
Inform dose recommendation for patients with organ impairment	Sane <i>et al.</i> (2022) ²⁷	Assess the prevalence of hepatic impairment prior to first-line therapy in patients diagnosed with mCRPC or HR+/HER2– mBC	Flatiron Health EHR	Provided evidence to justify the conduct of a pharmacokinetic study of ipatasertib in at least participants with mild hepatic impairment
	Lu <i>et al.</i> (2020) ²⁸	Assess the feasibility of conducting dedicated organ impairment studies in patients diagnosed with DLBCL	Flatiron Health EHR	Indicated a very challenging enrollment for a dedicated organ impairment study due to prevalence, leading to a waiver of the dedicated organ impairment studies from regulatory agencies
	Spillane <i>et al.</i> (2020) ²⁹	Evaluate characteristics and outcomes in patients with organ impairment diagnosed with advanced melanoma and treated with an immune checkpoint inhibitor	Flatiron Health EHR	Patients with advanced melanoma and baseline organ impairment have poorer clinical outcomes than patients with normal organ function
	Sybing <i>et al.</i> (2022) ³⁰	Characterize the time course of glomerular hyperfiltration in pediatric and adult patients with SCD	Optum EHR	Showed evidence of glomerular hyperfiltration in pediatric SCD patients and the rate of decline in adult SCD patients. Results could help clinicians in anticipating the need for dose adjustment due to renal impairment in SCD patients
Provide insights for pediatric plan development and study design	Chanu <i>et al.</i> (2020) ³²	Use RWD/RWE to supplement modeling work based on RCTs to optimize C.E.R.A. development in the confirmatory trial of the pediatric plan	IPDN database	Confirmed the model simulated treatment outcomes in pediatric patients receiving C.E.R.A. i.v. and s.c. and provided a strong rationale for applying the C.E.R.A. S.c. dosing regimen only in pediatric patients rather than both i.v. and s.c., leading to a simplified confirmatory trial in pediatrics
	Zhang (2021) ³³	Evaluate the dose–response relationship of vedolizumab in a pediatric population with IBD using RWD to support dose selection for etrolizumab in pediatric clinical trials	Pediatric registry ImproveCareNow	Indicated that the majority of pediatric patients from the RWD database were treated with the equivalent adult labeled dose of vedolizumab and efficacy was similar or slightly better in the pediatric study cohort compared with that observed in adult clinical trials
	Lukka <i>et al.</i> (2021) ³⁴	Support lacosamide dosing in the younger pediatric population (<4 years old) where dose recommendations are not yet available	Real-world therapeutic drug monitoring data	Simulation-based dosing regimens for pediatrics derived via pharmacokinetic modeling of RWD assessed during clinical care to generate RWE and provide a rational basis for exposure-matched lacosamide pharmacotherapy in children <4 years of age

(Continued)

Table 1 (Continued)

Category	Reference	Objective	RWD source	Insight/evidence generated
Enable and enrich MIDD notably disease progression modeling	Doler <i>et al.</i> (2013) ³⁶ Jamalian <i>et al.</i> (2020) ³⁷	Develop a disease progression model to describe natural progression of Alzheimer's disease	ADNI database	Disease progression model was developed to quantitatively characterize the progression of the disease and can be used to predict natural disease progression in Alzheimer's disease patients
	Boucher <i>et al.</i> (2018) ³⁸	Investigate disease progression and treatment effect in patients diagnosed with hereditary transthyretin-mediated amyloid polyneuropathy, a rare disease	THAOS	Relevant and consistent disease progression and treatment effects (of tafamidis) were estimated in an independent clinical trial and in patients from RWD
	Wang <i>et al.</i> (2019) ³⁹	Model disease progression and identify risk factors for patients diagnosed with DMD, a rare disease	Cooperative International Neuromuscular Research Group DMD Natural History Study	Models adequately described disease progression for key end points in ambulatory and nonambulatory DMD boys
	Abrams <i>et al.</i> (2020) ⁴⁰	Develop a QSP model to predict response for patients diagnosed with GD1	ICGG Gaucher Registry	Generated virtual patients that captured the appropriate disease phenotypes of interest with more accurate representation of their variability, which enabled the QSP modeling that captured specific clinical attributes of the disease, incorporated markers of disease severity, and informed relevant treatment strategies
	Chanu <i>et al.</i> (2021) ⁴²	Use M-protein dynamics as an early time biomarker to predict OS for patients diagnosed with multiple myeloma	Flatiron Health EHR	Model built with RWD can inform drug development in multiple myeloma; e.g., predict survival outcomes of multiple independent Phase iii trials leveraging M-protein dynamics collected in a smaller early phase trial
	Kotani <i>et al.</i> (2021) ⁴³	Use RWD to check survival distribution from clinical trial used to develop a disease model for HER2-/HR+ mBC	Flatiron Health EHR	Survival data from RWD is consistent with the one from the clinical trial used to develop the disease model
Identify prognostic and predictive biomarker/factors	Julian <i>et al.</i> (2022) ⁴⁴	Investigate prognostic factors of OS in patients with advanced aNSCLC and develop a novel prognostic model	Flatiron Health EHR	A prognostic model in patients with aNSCLC receiving anti-PD1/PD-L1 immune checkpoint inhibitors as second line monotherapy was developed using RWD with 42 important prognostic factors identified. The prognostic model was able to discriminate overall survival and perform well in real-world and clinical trial cohorts
	Yun <i>et al.</i> (2021) ⁴⁵	Identify patients who may respond better to a specific drug or mechanism of action	ACR's RISE registry	Supported better understanding of disease and treatment response and identified/confirmed treatment benefit differentiation factors
Support regulatory decision making for label expansion	Erdman <i>et al.</i> (2021) ⁵⁰	Provide evidence of efficacy and safety of tacrolimus-based immunosuppressive regimens in adult lung transplant recipients in the United States	SRTR database	Supported the use of two tacrolimus combinations as maintenance immunosuppressive regimens in adult lung transplant recipients and supported the expansion of the product label to include lung transplantation
	Wedam <i>et al.</i> (2020) ⁵¹ Kraus <i>et al.</i> (2020) ⁵⁰	Provide evidence of safety and effectiveness of palbociclib plus ET in men with HR+/HER2- mBC	Flatiron Health EHR, Claims databases (Pharmacy and Medical), and Post-marketing surveillance	RWD indicated that men with mBC benefit from palbociclib plus ET, with a safety profile consistent with previous observations in women with mBC; combined with the collective data in women supported by RWE, the palbociclib indication was expanded to include men with HR+/HER2- mBC in the United States
	Lamba <i>et al.</i> (2017) ⁵⁴ Cohen <i>et al.</i> (2021) ⁵⁵	Provide comparative RWE for the IR and PR formulations of tofacitinib in RA patients	CorEvitas registry (formerly Corrona)	RWD provided key supportive evidence, in addition to model-based bridging and clinical trial data, for the marketing application and subsequent approval of a PR formulation of tofacitinib in the European Union (EU) as part of the totality of evidence to bridge efficacy and safety data from a previously approved IR tablet formulation

(Continued)

Table 1 (Continued)

Category	Reference	Objective	RWD source	Insight/evidence generated
Generate synthetic/external control for rare diseases	CDER (2017) ⁶⁰	Characterize the natural history of patients diagnosed with metastatic MCC treated with chemotherapy	Observational study consisting of a multicenter retrospective chart review of patients treated with chemotherapy for distant metastatic MCC	Provided response rates to chemotherapy treatment allowing exploratory characterization of the risk:benefit profile of avelumab in the context of the natural history of MCC and treatment outcomes with cytotoxic chemotherapy
	CDER (2018) ⁶¹	Compare blinatumomab treated patients with historical control with respect to RFS	A noninterventional retrospective analysis of RFS and OS among patients with Philadelphia chromosome-negative ALL and MRD $\geq 0.01\%$ who received standard-of-care treatment	Improved survival probability on blinatumomab relative to historical control calculated based on propensity scores for each patient
	Popat <i>et al.</i> (2022) ⁶²	Assess comparative effectiveness of pralsetinib in NSCLC by combining RWD and trial data; performed sensitivity analyses to quantify effect of sources of bias	FMI CGDB	Provided evidence in favor of pralsetinib over other treatments as an effective first line treatment for RET fusion-positive aNSCLC. Bias assessments showed robustness to potential sources of bias and can be used as a template for future studies
	Ayodele <i>et al.</i> (2021) ⁶³ Chen <i>et al.</i> (2020) ⁶⁴	Investigate eGFR changes in hypoparathyroidism patients treated with rhPTH(1–84) relative to SOC treatment	US Explorys EHR database; Geisinger Healthcare database	eGFR was stable in rhPTH(1–84) treated patients from clinical trials but declined in the historical control/SOC group
	Gosmanova <i>et al.</i> (2021) ⁶⁵	Explore risk of CKD in hypoparathyroidism patients treated with rhPTH(1–84)	US Explorys EHR database	Patients treated with rhPTH(1–84) in long-term clinical trials had lower risk of CKD than historical control

ADNI, Alzheimer's Disease Neuroimaging Initiative; ALL, acute lymphoblastic leukemia; aNSCLC, advanced non-small cell lung cancer; CDER, Center for Drug Evaluation and Research; C.E.R.A., Continuous Erythropoietin Receptor Activator; CGDB, Clinico-genomic database; CKD, chronic kidney disease; COVID-19, coronavirus disease 2019; DDI, drug-drug interaction; DLBCL, diffuse large B-cell lymphoma; DMD, Duchenne muscular dystrophy; eGFR, estimated glomerular filtration rate; EHRs, electronic health records; ET, endocrine therapy; FAERS, US Food and Drug Administration Adverse Event Reporting System; FMI, Flatiron Health Foundation Medicine; GD1, Gaucher disease type 1; IBD, irritable bowel disease; ICGG, International Collaborative Gaucher Group; IPDN, International Pediatric Dialysis Network; IR, immediate release; mBC, metastatic breast cancer; MCC, Merkel cell carcinoma; mCRPC, metastatic castrate-resistant prostate cancer; MIDD, model-informed drug development; MRD, minimal residual disease; OS, overall survival; PR, prolonged release; QSP, quantitative systems pharmacology; RA, rheumatoid arthritis; RCT, randomized controlled trial; RFS, relapse-free survival; RISE, Rheumatology Informatics System for Effectiveness; RWD, real-world data; RWE, real-world evidence; SCD, sickle cell disease; SOC, standard-of-care; SRTR, Scientific Registry of Transplant Recipients; THAOS, Transthyretin Amyloidosis Outcomes Survey; UCSF, University of California, San Francisco.

presence of strong CYP3A4 perpetrators. Combining the concomitant medications' information from RWD and findings from the human DDI substudy, it was decided to exclude the use of concomitant medications that were strong CYP3A4 inducers/inhibitors in clinical trials of this investigational molecule in patients with psoriasis. Taken together, RWD/RWE-informed early DDI evaluations can help generate a seamless clinical pharmacology strategy to inform the study design of clinical trials in healthy volunteers (e.g., DDI substudy) and in patients (e.g., concomitant medication exclusion).

In addition, RWD combined with other relevant data (e.g., *in vitro* or experimental data, published literature, and clinical databases) have the potential to identify clinically relevant DDI risks. For example, Duke *et al.* identified potential drug–drug pairs that could result in an adverse event of myopathy due to CYP enzymes-based interactions (as either a substrate or inhibitor) when co-administered. First, authors identified an initial set of 13,197

potential drug pairs predicted to have DDI potential based on published *in vitro* pharmacology experiments.¹⁹ Then, they identified 3,670 drug pairs taken by patients in the real world based on analysis of a clinical repository containing over 800,000 patients. Finally, based on rigorous statistical evaluation, five new drug–drug pairs were identified which had increased risk of myopathy, thus demonstrating clinically relevant DDIs through CYP enzymes. Authors conclude that similar automated search algorithm may be beneficial in identifying additional clinically significant interactions of the FDA approved drug by leveraging published data and large RWD clinical databases. Another example identified that combination therapy with ceftriaxone and lansoprazole increased the risk of acquired long QT syndrome.²⁰ Investigators identified 889 drug combination pairs signaling QT risk in the FDA Adverse Event Reporting System (FAERS) with 1.8 million QT prolongation adverse events that could not be attributed to individual effect of the drugs. An EHR database with an additional 1.6 million

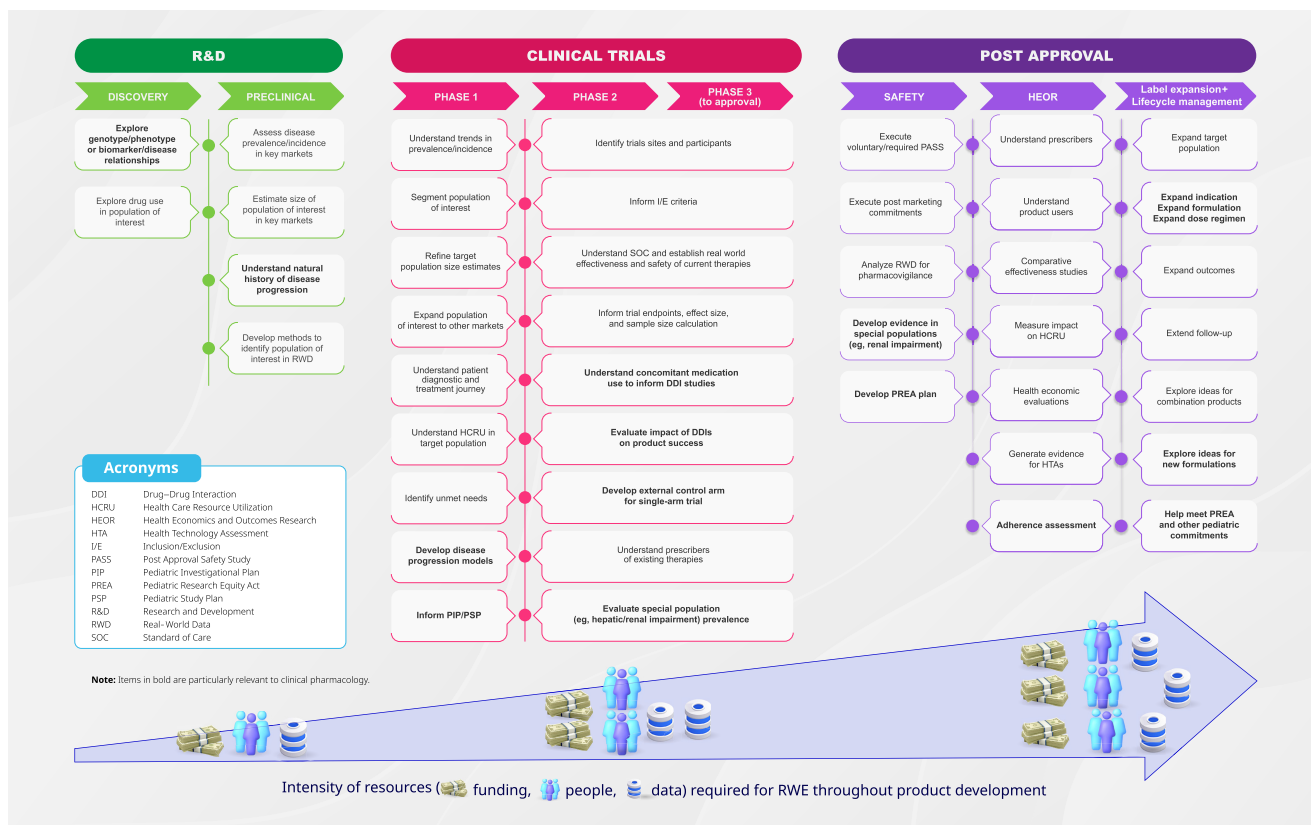


Figure 1 Example applications of real-world data to inform decisions throughout product development.

electrocardiogram records from 380,000 patients was used to narrow down the identified DDI pairs to 8 combinations, including the ceftriaxone and lansoprazole combination. Then, patch clamp experiments were conducted and revealed significant human ether-a-go-go related gene channel inhibition by the synergistic effect of this drug combination. This example highlights the use of RWD to identify a novel DDI risk and then laboratory experiments to elucidate the mechanisms of the DDI. Another recent example during the coronavirus disease 2019 (COVID-19) pandemic exemplifies how *in vitro* experiments followed by confirmation from RWD analysis identified clinically relevant transporter-mediated DDIs for 25 small molecule drugs evaluated in clinical trials for COVID-19.²¹ Authors analyzed EHR data from the University of California, San Francisco (UCSF) Research Data Browser and Cerner Real World COVID-19 Database to confirm that the *in vitro* transporters mediated DDI risks were consistent with the RWD. Thus, they recommended that vulnerable patients diagnosed with COVID-19 (i.e., geriatric patients with polypharmacy risk) should be carefully monitored for adverse drug reactions due to transporter-mediated DDIs for the drugs being evaluated expeditiously for COVID-19 at that time.

INFORM DOSING RECOMMENDATIONS IN PATIENTS WITH ORGAN IMPAIRMENT

Understanding the impact of organ impairment on drug exposure, safety, and efficacy is crucial to guide dosing recommendations and adjustments in these populations. These dose adjustments

are typically based on exposure changes due to organ impairment, which can be characterized using either dedicated pharmacokinetic (PK) studies or modeling and simulation approaches. It is a typical practice in drug development to explicitly exclude participants with advanced organ impairment from phase II and phase III trials. In a few cases, dedicated efficacy and safety studies in patients with organ impairment are conducted.^{22–26} In general, enrolling participants with organ impairment is challenging irrespective of the study type, especially for those with moderate and severe organ impairment. RWD/RWE can be used in this space in multiple ways: (i) informing the need for organ impairment studies, (ii) evaluating the feasibility of characterizing exposure changes of the investigational drug in target patient populations with organ impairment, (iii) generating post-approval efficacy and safety data to guide dosing in patients with organ impairment, and (iv) assessing the time course of organ impairment progression.

Dosing recommendations for patients with organ impairment are typically required if the drug is likely to be used in patients with organ impairment. RWD can be used to assess the prevalence of organ impairment in the target patient population in cases where organ impairment is not a comorbidity in a significant portion of the target population, which can be used to inform the need to conduct dedicated organ impairment studies. Sane *et al.*²⁷ conducted a retrospective analysis using Flatiron Health EHR data to assess the prevalence of hepatic impairment prior to first-line therapy in patients diagnosed with metastatic castrate-resistant prostate cancer (mCRPC) and hormone receptor positive/human epidermal

growth factor receptor 2 negative (HR+/HER2-) metastatic breast cancer (mBC). Using the National Cancer Institute Organ Dysfunction Working Group (NCI-ODWG) classification criteria, the proportions of patients with mild and moderate-to-severe hepatic impairment were 12.4% and 0.7%, respectively, in the mCRPC cohort, and 18.8% and 1.4%, respectively, in the HR+/HER2- mBC cohort. Based on this, the study team concluded that mild hepatic impairment is a relatively common comorbidity in patients with mCRPC and HR+/HER2- mBC, whereas the prevalence of moderate to severe hepatic impairment is low in this patient population. This analysis provided the evidence to inform internal decision making regarding the need to conduct a PK study for the investigational drug in participants with at least mild hepatic impairment.

RWD can also be used to assess the feasibility of characterizing drug exposure changes in patients with organ impairment in the target patient population when it is not ethical to administer the drug to participants with organ impairment who are otherwise healthy. This can be achieved by either conducting a dedicated PK study in patients with organ impairment or allowing the enrollment of patients with organ impairment in phase II and phase III studies. For example, to assess the feasibility of conducting dedicated organ impairment studies in patients diagnosed with diffuse large B-cell lymphoma (DLBCL), an exploratory analysis was performed using RWD on previously untreated patients with DLBCL ($n = 1,341$) extracted from the Flatiron Health database.²⁸ A relatively low percentage of patients with DLBCL with moderate or severe hepatic impairment (<6%) or with severe renal impairment (<6%) were found in the database, indicating a very challenging enrollment for a dedicated organ impairment study. In addition, the analysis showed that the study population in phase II/III trials cover the majority (>90%) of the previously untreated patients with DLBCL in the Flatiron RWD, and a dedicated hepatic/renal impairment study would be of limited added value. Therefore, regulatory agencies agreed with the sponsor's proposal to not conduct dedicated organ impairment studies, which reduced costs, patient burden, and development timelines. However, one limitation of this RWD study was that the data included only patients with DLBCL who received first-line therapy because there were no data available for relapsed/refractory patients, which was the target population. In addition, it should be noted that the eventual labeling recommendation states that drug administration should be avoided in patients with moderate or severe hepatic impairment due to lack of data in this special population.

In addition, RWD can be used to provide cumulative evidence about safety and efficacy outcomes in patients with organ impairment following approval of an investigational agent. This can apply to the above example where such information can be generated for patients with moderate to severe hepatic impairment as well as severe renal impairment after the initial approval, which can be used to inform dosing recommendations in these patients. Spillane *et al.*²⁹ used RWD to evaluate the use of immune checkpoint inhibitors (ICIs) approved for advanced melanoma management in patients with organ dysfunction. They conducted a retrospective analysis using Flatiron Health EHR to identify patients with melanoma who received ICIs as the first line of treatment. A total

of 2,407 patients were identified, of which 2.4% had a baseline of moderate or severe renal impairment, and 2.8% had a baseline of moderate or severe hepatic impairment. The analysis showed that patients with advanced melanoma and baseline organ impairment have poorer clinical outcomes (i.e., shorter real-world time to treatment discontinuation and overall survival) than patients with normal organ function. This information may be used to evaluate the need of dose adjustment in the organ impairment population.

RWD can also be an effective tool in assessing the time course of organ impairment progression, especially in certain rare diseases, and subsequently guiding clinicians on the timing of dose adjustment with respect to organ impairment. Sybing *et al.*³⁰ used RWD to characterize the time course of glomerular hyperfiltration (defined as higher-than-normal renal filtration rate) in pediatric and adult patients with sickle cell disease (SCD), which is a driver of renal impairment in later years. The onset and peak of glomerular hyperfiltration, and subsequent decline in renal function in patients with SCD, were characterized using RWD from the Optum EHR database. The analysis showed that hyperfiltration was observed in hemoglobin (Hb) SS genotype (the common type of SCD) patients with SCD as early as 1 year of age and peaked between 8 and 10 years of age. Hyperfiltration declined steadily with age in Hb SS patients, and after 40–50 years of age, the estimated glomerular filtration rate fell below that for the non-SCD population. This was the first analysis showing evidence of glomerular hyperfiltration in pediatric patients with SCD and the rate of decline in adult patients with SCD using RWD and is deemed useful for clinicians by helping anticipate the need for dose adjustment due to renal impairment in patients with SCD.

SUPPORT PEDIATRIC PLAN DEVELOPMENT AND DOSING OPTIMIZATION

Pediatric dosing is often extrapolated from adults using exposure matching approach with limited information on the physiological, anatomic, and ontogeny-based differences between pediatric and adult populations. The recent draft International Conference on Harmonization (ICH) harmonized guideline E11A on pediatric extrapolation recommends leveraging multiple sources of information to contribute to the clinical evidence package, and the use of RWD is encouraged and discussed.³¹ In the case studies detailed below, RWD were leveraged to support pediatric drug development and dose selection in multiple ways by confirming model-based inferences, by leveraging competitor data, or by exploring pediatric population for which no dose recommendations were yet available.

Chanu *et al.*³² used RWD/RWE to supplement modeling based on RCTs to optimize pediatric development plan and the confirmatory trial design for Continuous Erythropoietin Receptor Activator (C.E.R.A.). The initial pediatric plan was designed as a phase II dose-finding study using C.E.R.A. administered intravenously (i.v.) followed by a large confirmatory trial using C.E.R.A. dosed via both i.v. and subcutaneous (s.c.) injection along with a comparator arm. The plan was then optimized using a model-based PK/pharmacodynamic (PK/PD) analysis confirmed by RWD to reduce the confirmatory trial to a smaller, single arm trial with only s.c. C.E.R.A. The PK/PD model was built with data from a phase

II pediatric i.v. study and phase II/III adult studies to determine the PK/PD characteristics of C.E.R.A. administered i.v. and s.c. in a broader population. RWD on C.E.R.A. doses and Hb levels obtained from the International Pediatric Dialysis Network (IPDN) registries confirmed the model-predicted treatment outcomes in pediatric patients receiving C.E.R.A. i.v. and s.c. This provided a strong rationale for testing the C.E.R.A. s.c. dosing regimen only rather than both i.v. and s.c. in pediatric patients. Therefore, the confirmatory trial was re-designed and simplified, which reduced unwarranted drug exposure and treatment burden for children and shortened timelines to bring C.E.R.A. sooner to pediatric patients. One limitation of this study is that only summary level (not patient level) RWD were used to compare with model predicted data (i.e., model-based simulations as median and 90% prediction interval were compared with median of observed value in RWD). For further evaluation purpose, the sponsor has also launched a prospective RWD study to collect patient level data.

In another example, RWD from pediatric patients with off-label use of a competitor drug were used to inform pediatric dose selection and the development plan for an investigational drug. Etrolizumab, a humanized monoclonal antibody, was under clinical development for inflammatory bowel disease (IBD). One of etrolizumab's mechanisms of action is shared with vedolizumab, which has already been approved for adults with IBD. RWD from pediatric patients treated with vedolizumab (off-label use) from the ImproveCareNow (ICN) registry (the largest pediatric IBD registry) were used to characterize the dose–response relationship and provide additional evidence in support of dose selection for etrolizumab in pediatric clinical trials.³³ The results indicated that the majority of pediatric patients from the RWD database were treated with doses equivalent to the adult labeled dose of vedolizumab, and the efficacy was similar or slightly better in pediatric cohorts compared with that observed in adult pivotal clinical trials. The key limitation of this study was related to the dosing data. This is because vedolizumab doses (off-label use) in pediatric patients are not standardized in real-world settings (e.g., some doses are fixed while others are body weight-based) and dose disparity as dose can be confounded by disease severity (i.e., more severe patients are likely to be given higher doses as compared with less severe patients), both of which may introduce bias in the analysis.

RWD were also used to support dosing in younger pediatric populations where dose recommendations are not yet available. For instance, lacosamide, an anti-epileptic drug, is approved for the treatment of focal seizures in children ≥ 4 years of age and adults. Researchers used real-world therapeutic drug monitoring data from 315 pediatric patients (> 1 month to < 18 years of age) who received lacosamide to build a population PK model with allometric scaling of body weight and covariate analysis (age included).³⁴ The model was used to simulate lacosamide exposure for age-associated doses to match the exposure in children ≥ 4 years of age with the weight-based dosing recommendations provided by the FDA. Using this approach, the authors provided dose recommendations for children aged 1 month to < 4 years old. Potential limitations of this study included imbalance in patient representation across age groups, co-medications, and a higher variability in drug dosing and follow-up timing in the RWD.

ENABLE AND ENRICH MODEL-INFORMED DRUG DEVELOPMENT NOTABLY DISEASE PROGRESSION MODELING

The use of quantitative methodologies and MIDD in clinical pharmacology has exponentially increased in the past decade to guide decision making in drug development and approval.³⁵ Approaches such as pharmacometric modeling and quantitative systems pharmacology (QSP) modeling have allowed researchers to answer questions that cannot be addressed by traditional clinical pharmacology or statistical approaches. The increasing availability and accessibility of RWD has expanded the capabilities of modeling in a multifaceted manner. RWD can be used for model development (e.g., alone or in combination with clinical trial data), model validation (e.g., external data source to validate a model built with only clinical trial data), and in support of model building and evaluation (e.g., inform virtual population generation in QSP modeling).

Notably, RWD/RWE provide insights into disease progression by capturing information on real-world patients in the real-world setting. Unlike clinical trials, which provide a snapshot of a patient's journey, RWD with long follow-up can better capture the natural history of a disease and/or disease progression as well as potential risk factors which may influence disease progression. For instance, researchers developing models to complement drug development efforts for Alzheimer's disease used RWD from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database alone or together with placebo data from interventional clinical trials, to predict disease progression.^{36,37} By predicting individual trajectories of disease progression in the absence of treatment, the model can reliably assess drug effects for molecules in development and can be leveraged for study design of future clinical trials. Use of an RWD patient cohort allowed for a larger sample size, more diverse patient population, and longer follow-up compared with a clinical trial. Combining clinical trial data with RWD allowed a richer dataset for model building, and conducting external validation with clinical trial data also helped assess predictive performance of the model.³⁷ RWD may be particularly valuable to characterizing disease progression in rare diseases, where clinical trial size is limited by the small patient pool. Boucher *et al.*³⁸ used RWD from the Transthyretin Amyloidosis Outcomes Survey (THAOS) to investigate the natural history of disease progression as well as treatment effect in patients diagnosed with hereditary transthyretin amyloid polyneuropathy. Similarly, using an existing Duchenne muscular dystrophy (DMD) natural history database, researchers built independent longitudinal models for the North Star Ambulatory Assessment total score for ambulatory boys and for the forced vital capacity percent predicted value for non-ambulatory boys to characterize disease progression of DMD and identified risk factors having a significant effect on outcomes in boys diagnosed with this disease.³⁹ These models informed the quantitative understanding of disease progression in rare diseases, and simulation results from the models could then contribute to the study design of future clinical trials to accelerate clinical drug development.

Besides disease progression models, RWD can be used to enrich and enable other types of models to inform drug development. For example, a QSP model for Gaucher disease type 1 (GD1) was built

with the intention to predict treatment response within the heterogeneous GD1 patient population.⁴⁰ Researchers in this study were able to use data from a variety of sources, including RWD from the International Collaborative Gaucher Group (ICGG) Gaucher Registry. This registry consists of more than 6,000 patients with Gaucher disease across more than 60 countries, allowing for integration and evaluation of response from more diverse and representative patient populations relative to those enrolled in clinical trial settings. Although differences between registry and clinical trial data were identified, steps were taken to reduce bias when using the registry data to generate virtual patients that capture the appropriate disease phenotypes of interest with more accurate representation of their variability. This in turn enabled the QSP modeling that captured specific clinical attributes of the disease, incorporated markers of disease severity, and allowed development of relevant treatment strategies.

In addition, there is a well-established modeling and simulation framework in oncology that links dose, exposure, tumor dynamics, and overall survival (OS).⁴¹ The link between tumor dynamic and OS is disease-specific and drug-independent. Tumor growth inhibition (TGI) metrics can be used to capture treatment effect and predict the OS benefit in this TGI-OS disease modeling framework. In recent years, RWD have been used to model TGI and its relationship to OS, enabling early efficacy predictions based on tumor dynamics for molecules of interest. For example, Chanu *et al.*⁴² recently presented data where researchers utilized a serum-based marker for tumor burden (M-protein) and OS data from Flatiron Health to build a drug-independent disease model for multiple myeloma. This disease model can support drug development in multiple myeloma by predicting OS based on early M-protein dynamics data obtained for a new investigational drug. For instance, it could be used to simulate OS benefits vs. standard of care complementary to observed progression-free survival obtained in a randomized pivotal trial, or assess the probability of success of a trial investigating a new treatment over a reference treatment leveraging M-protein dynamics collected in a smaller phase I or phase II trial. In a similar context, for the development of a TGI-OS model in patients with HR+/HER2- mBC based on clinical trial data, real-world OS data from Flatiron Health were used to ensure the OS distribution used to develop the model was representative of the real world population, thus verifying model relevance for drug development in the target population and real world setting.⁴³ One challenge with the use of RWD in oncology for modeling purposes is that patients are followed along several lines of therapies and it is important to have a systematic and consistent way to select the right line of therapy in each patient that is relevant to the drug development question.

IDENTIFY PROGNOSTIC AND PREDICTIVE BIOMARKERS/FACTORS

RWD can be leveraged to identify factors for disease prognosis or to elucidate increased treatment benefit via modeling approaches, which can inform patient subpopulation classification of either high-medical need or high-benefit potential. To enhance clinical utility, data from clinical studies are often combined with data from real-world sources to either establish or validate prediction

models. For example, Julian *et al.*⁴⁴ utilized a retrospective observational patient cohort with advanced non-small cell lung cancer (NSCLC) from the Flatiron Health database to develop a novel prognostic model to identify disease prognostic factors for OS. A Cox proportional hazards survival model was built with RWD from 4,049 patients with advanced NSCLC on selected second line anti-PD1/PDL1 monotherapy (atezolizumab, nivolumab, or pembrolizumab). The ability of the established prognostic model to predict OS differentiation was then validated using a patient cohort treated with second line atezolizumab monotherapy from an independent RCT. With this validated prognostic model of OS, multiple prognostic factors (e.g., baseline demographics, clinical characteristics, and laboratory results) that enhance or reduce risk of death were identified, which can be used to support future research on patient prognostic prediction.

In addition, Yun *et al.*⁴⁵ presented an example of using RWD to validate a previously established treatment responder identification model built with clinical trial data. Based on sarilumab clinical trial data, a laboratory results-based rule to identify patients with rheumatoid arthritis (RA) with a more favorable response to sarilumab was established via a machine learning approach. Two hundred five sarilumab users were then identified from the American College of Rheumatology (ACR)'s Rheumatology Informatics System for Effectiveness (RISE) data from 2017 to 2021 based on prespecified inclusion/exclusion criteria. Logistic regression indicated that better sarilumab responses were demonstrated in rule-positive patients compared with those in rule-negative patients in the real-world setting. This RWD application illustrated the use of RWE to understand disease and treatment response and identify/confirm treatment benefit differentiation factors.

SUPPORT REGULATORY DECISION MAKING AND LABEL EXPANSION

Currently, substantial evidence from adequate and well-controlled clinical studies, typically consisting of RCTs, is required for drug licensing and regulatory approval.⁴⁶ However, other types of clinical studies, including single-arm trials, open-label trials, and meta-analyses supplemented by RWE, have been utilized for drug approval in limited instances, such as in rare diseases. Under the Cures Act, the FDA's RWE Program evaluates the potential use of RWD to generate RWE on product effectiveness to help support label expansion with new indications, populations, or dosing for approved drugs, and the role of RWE in supporting label changes has been discussed in the literature.^{46–49} In this context, the examples below illustrate recent utilization of RWE for label expansion of new indications and new product formulation, with significant clinical pharmacology input in the totality of evidence required for such regulatory approvals.

RWD have been used to assess effectiveness of tacrolimus (TAC)-based immunosuppressive regimen combinations in adult lung transplant recipients, which is a patient population challenging to recruit in clinical trials. In this case, RWE was generated using retrospective analysis of the Scientific Registry of Transplant Recipients (SRTR) database (the most comprehensive database of transplant recipients in the United States). Analyses indicated

that lung transplant recipients who received a combination of TAC with mycophenolate mofetil or azathioprine at hospital discharge had high survival rates at 1-year post-transplant, supporting the use of TAC in these combinations as maintenance immunosuppression in adult lung transplant recipients. This RWE, along with evidence from pivotal RCTs of TAC and a meta-analysis report of three RCTs in patients with a lung transplant with TAC-based treatment, provided totality of effectiveness and safety evidence that led to label expansion of TAC to include lung transplantation.⁵⁰

Based on RWD, men with HR+/HER2- mBC were shown to benefit from palbociclib plus endocrine therapy (ET).^{51,52} Palbociclib treatment in men with mBC was evaluated using three independent RWD sources: IQVIA Insurance database (pharmacy and medical claims, $n = 1,139$), Flatiron Health breast cancer database (EHRs, $n = 59$), and a global safety database. RWE indicated that men with mBC benefit from palbociclib plus ET, with a safety profile consistent with previous observations in women with mBC. In addition to RWE, similar palbociclib exposures were demonstrated in men and women using a population PK analysis.⁵³ Therefore, RWE together with previous pivotal clinical trials data and a well-established benefit/risk profile of palbociclib led to the palbociclib label expansion to include men with HR+/HER2- mBC in the United States.

RWE also supported the approval of a prolonged release (PR) formulation of tofacitinib in the European Union as part of totality of evidence approach to bridge efficacy and safety data from a previously approved immediate release (IR) tablet formulation. Tofacitinib is an oral Janus kinase inhibitor that was initially approved globally as an IR oral tablet, at a dose of 5 mg twice daily, for adult patients with moderate to severe active RA, based on a traditional development program that included confirmatory phase III RCTs. The PR tablet formulation of tofacitinib was developed, at a dose of 11 mg, to provide a more convenient once daily dosing alternative. Following the approval of the PR tablet in the United States, based on comparative phase I PK data and an MIDD approach to bridge efficacy and safety in patients with RA,⁵⁴ the PR formulation was included in the ongoing US registry (CorEvitas, formerly Corrona) initially used for the IR formulation. Limited clinical trial data with the PR formulation and the MIDD-based bridging were considered to be insufficient evidence for consideration to approve the PR formulation in the European Union. The comparative effectiveness of the IR and PR formulations in patients with RA from the CorEvitas registry provided key supportive evidence that subsequently resulted in product approval for the PR formulation in the European Union.⁵⁵

GENERATE SYNTHETIC/EXTERNAL CONTROL FOR RARE DISEASE

RCTs, by definition, require a control intervention arm to accompany the experimental intervention arm. The control intervention may be placebo or standard-of-care, if one exists. Randomization supports an unbiased comparison between the intervention arm and the control arm. Various practical and ethical factors can make patient recruitment and retention challenging for the control arm. This is especially true for rare diseases, where the eligible patient pool may be too small to adequately recruit for even the

treatment arm. Some rare diseases may not have a standard-of-care established, and patients may be reluctant to join or remain in the placebo arm. In some cases, treatment with placebo may be unethical. Moreover, emerging evidence from an ongoing study could affect clinical equipoise due to, for instance, belief of clinical benefit in the intervention arm.⁵⁶ Because of these logistical and ethical challenges, external control arms based on external data (e.g., RWD) may be considered as an alternative. An external control is based, at least in part, on external data, which refers to any relevant source of clinical data not from a concurrently randomized control group within the same study.⁵⁷ The external control may be selected patient cohorts from one or more external data sources and also modified using statistical methodologies, and it is also referred to as the synthetic control.⁵⁸

External control arms based on RWD have been used to support regulatory filings.⁵⁹ Avelumab was approved for the treatment of metastatic Merkel cell carcinoma, a rare skin cancer. Data from an EHR database were used to create an external control arm and were considered in the FDA review as an exploratory component to “further characterize the risk/benefit profile of avelumab.”⁶⁰ Another relevant example is blinatumomab, which is indicated for Philadelphia chromosome-negative relapsed or refractory B-cell precursor acute lymphoblastic leukemia (ALL). Retrospective ALL patient data were used to construct an RWD-based external control using propensity matching and compared with RCT data from a single-arm efficacy study conducted for blinatumomab.⁶¹ Various challenges (e.g., non-contemporaneous historical control and difference in follow-up duration), were identified as limiting factors for a reasonable comparison. The review concluded, “the propensity score analysis was not sufficient to determine the estimate of the benefit of blinatumomab.” Both reviews provided insights into how external controls may be used to support regulatory filings and highlighted data quality and analytical approaches which may maximize the utility and acceptance of external controls by health authorities.

Use of external control arms has also been reported for other oncology drugs. Only 1–2% of all patients with NSCLC are rearranged during transfection (RET) gene fusion-positive, making this a rare mutation. A nonrandomized, open-label, uncontrolled study demonstrated that pralsetinib, a selective RET inhibitor, was efficacious in treatment-naïve RET fusion-positive patients. However, in this trial, efficacy of pralsetinib was not assessed in relation to other therapies. Popat *et al.*⁶² used RWD to construct an external control and demonstrated pralsetinib offered a survival benefit relative to other therapies. Importantly, the authors performed various sensitivity analyses to quantify how various sources of bias in the RWD may impact the conclusions. Recent studies have also used external control arms to assess the long-term impact of treatment in rare endocrine diseases when a simultaneous long-term control arm is either unethical or impractical.^{63–65}

As with any other application of RWD, the quality of the data is a key factor in determining its usability, especially for regulatory purposes.⁶⁶ Use of evolving statistical and analytical methods that quantify the impact of potential bias in the external control data are likely to improve the robustness of conclusions and increase confidence in the evidence generated.⁵⁸

CONSIDERATIONS FOR SELECTING COMMON SOURCES OF RWD TO ADDRESS CLINICAL PHARMACOLOGY QUESTIONS

Key information needed from RWD to address research questions related to clinical pharmacology often includes: (i) patient characteristics (e.g., age and sex), (ii) disease status (e.g., diagnosis, severity, duration, and staging), (iii) medications used (e.g., name, formulation, dose amount, frequency, duration, and prior lines of therapy), (iv) laboratory tests (e.g., test type, methods, and results), and (v) other diagnostic tests (e.g., biomarkers, genotypes, and tumor size). Common types of RWD that may provide this information include medical and pharmacy claims, patient-level data from EHRs, diagnostic laboratory tests, hospital chargemaster records, disease or product registries, and patient surveys, as well as data from medical and mobile/wearable devices. Each type of RWD has strengths and limitations, and tradeoffs are often necessary when selecting a source of RWD for a study. Key attributes, strengths, and limitations of the three common RWD sources (i.e., claims, EHRs, and patient registries) are summarized in [Table 2](#) and discussed below. Common data elements found in different types of RWD are also summarized in [Table 3](#).

Databases with claims from large employers or health plans (e.g., MarketScan and PharMetrics) can provide rich information about working-age adults but often exclude those covered by Medicaid, Medicare, or other government-sponsored programs in the United States. Conversely, claims data from government sources (e.g., Medicare 5% national sample) may contain important information about adults over the age of 65 years, with permanent disability, or with certain chronic health conditions (e.g., end-stage renal disease) but often exclude those under the age of 65 years. To study diseases that span across ages and types of payers, such as infections, it may be necessary to combine findings from multiple sources of RWD with different patient populations (e.g., commercial claims + Medicare claims + Medicaid claims). Alternatively, researchers can consider an open network claims database that includes all payers (e.g., IQVIA Rx/Dx and Symphony Health), although these data sources often lack details about patient health insurance coverage status and contain only unadjudicated claims that may later be rejected by payers. A more general limitation for all claims databases is that they rely exclusively on clinical coding (e.g., International Classification of Diseases (ICD)-10 diagnosis codes) to convey health status, because claims are designed primarily for billing purposes, and such coding may be inaccurate or imprecise to identify or confirm a specific disease of interest. It can also be difficult to attribute specific healthcare encounters (e.g., office visit and hospital admission) to a single health condition when multiple diagnoses are present in the claims data.

Data from EHRs generally offer richer information about patient demographics (e.g., race, ethnicity, marital status, and occupation), lifestyle (e.g., smoking status, use of alcohol, and recreational drugs), medical history (e.g., vaccinations, family history, and lines of therapy), drug allergies, use of over-the-counter medications, vital signs (e.g., body mass index (BMI) and blood pressure), symptoms (e.g., severity, frequency, and duration), examination findings, clinical reasoning (e.g., differential diagnosis), and treatment plan (e.g., specialist referral).⁶⁷ However, researchers interested in using EHR data should be aware of several limitations. A general

drawback of EHR data is that specific data elements (e.g., BMI) are often missing and it is unclear if this missingness pattern is random. For example, a study found that BMI was more likely to be recorded in EHR data for patients with type 2 diabetes mellitus (T2DM) than those without T2DM.⁶⁸ Because individuals with T2DM generally have a higher BMI, differences in data availability could bias comparisons of BMI between these two groups. EHR data may also contain conflicting or inaccurate information which is exacerbated with repeated measurements (e.g., large fluctuations in body weight over short time intervals may indicate a data entry error).⁶⁹ Additionally, EHR databases are often limited to health-care providers who use a specific type of EHR system or have a data sharing agreement and may therefore not include all health care received by a patient (e.g., EHR data may not include data from allied health providers who use a different EHR system). Data on medication use in EHRs generally reflects medications that are prescribed rather than dispensed and consumed. Although some EHR databases include results of basic laboratory tests (e.g., complete blood cell counts and metabolic panel), limited information is available about the specific test methods used, and EHR databases generally do not contain detailed results about more advanced diagnostic tests (e.g., imaging, biopsies, and genotyping). It should also be noted that the unstructured EHR data (i.e., detailed clinical notes) that may be of greatest interest to researchers often cannot be shared by data vendors due to privacy concerns. Biopharmaceutical companies must therefore rely on EHR data vendors to review, interpret, analyze, and summarize this information on their behalf according to a study protocol and analytical plan; such studies can be very costly and must often be repeated for each disease of interest. Natural language processing is increasingly used to more efficiently analyze large volumes of unstructured EHR data but such techniques are still evolving.

Patient registries (i.e., disease or product registry), on the other hand, often contain standardized and longitudinal information about patients with specific diseases (e.g., DMD) or who received specific therapies (e.g., gene therapy for SCD).⁷⁰ These data can provide insights about changes in outcomes (e.g., patient-reported outcomes and physician assessments) over time and on how patient covariates (e.g., genotypes, phenotypes, comorbidities, prior therapies, and concomitant medications) can impact such changes. For example, the Collaborative Trajectory Analysis Project (cTAP) combines data from multiple registries related to DMD that contain standardized measures such as the North Star Ambulatory Assessment taken at regular intervals.⁷¹ Data from such patient registries are often used in clinical pharmacology to develop pharmacometric models (e.g., disease progression model), create external control for single-arm studies or *in silico* (simulated) trials. Although patient registries focused on specific diseases or therapies may have a much smaller total patient population (e.g., hundreds or thousands) than those found in large claims or EHR databases (e.g., tens of millions), they may nevertheless represent the single largest available data source for that disease or therapy. One particular limitation of patient registries is that they lack a comparable control group (i.e., patients without the disease or therapy of interest); as a result, comparative studies using registry data often combine them with other sources for an appropriate control group.

Table 2 General strengths and limitations of common RWD sources for clinical pharmacology applications

RWD source	Major data contributors	Limitations	Strengths	Example vendors
Claims (Medical and Pharmacy)	- Large, self-insured employers - Commercial insurance companies - Centers for Medicare and Medicaid - RWD aggregators	- Use clinical coding (e.g., ICD-10 diagnosis codes) to represent all health conditions - ICD codes can have varying degrees of predictive value depending on disease types or type of stays (outpatient vs. inpatient) - Complex data driven definitions sometimes needed to capture actual cases - Patients can be lost to follow-up with change in insurance provider - No measure of disease severity unless specified in diagnosis code - Represent only health conditions deemed relevant for billing purposes - No physician notes (i.e., no unstructured data) - No information on laboratory test results - Limited information on race, ethnicity, smoking status, family history, etc. - No over-the-counter or self-pay medications - No information about medications prescribed but not filled - No information about medications administered during inpatient stays - Not representative of all age patient ages and third-party payer types - Deaths not captured for majority of cases	- Very large sample sizes - Include all health conditions and diseases - Include all healthcare services including procedures, covered by third-party payers - Comprehensive longitudinal patient history captured - Include all types of healthcare providers and care settings - Good representation of working age adults with employer-sponsored health insurance - Allow patients to be followed over several years - Include information about costs of healthcare services - Pharmacy claims include medication name, formulation, strength, route of administration, quantity, and days supplied	IBM/Truven MarketScan, IQVIA Pharmetrics Plus, Optum Clinformatics, Komodo
EHRs	- General EHR vendors (e.g., Cerner) - Specialty EHR vendors (e.g., Flatiron) - Large health systems (e.g., Truveta) - Specialty health networks (e.g., US Oncology)	- Include only health encounters with healthcare providers who use specific EHR software or are part of specific network so there could be unknown gaps in patient history - Do not contain information about charges or costs of healthcare services - Do not include detailed information about health plan coverage status - Information about medications prescribed but not if they are dispensed or taken - Self-reported health information (e.g., medications used, immunization history) may be inaccurate or missing - Detailed information in clinical notes and unstructured data may not be shared - Completeness of mortality data variable	- May include information about patient demographics (e.g., race, ethnicity, marital status) - May include information about lifestyle (e.g., smoking, alcohol) - May include vitals (e.g., blood pressure, BMI) and physical examination findings - May include basic laboratory tests (e.g., metabolic panel) but not advanced tests (e.g., genotyping) - May include clinical information beyond ICD-10 diagnosis codes - May include information about over-the-counter medications - May include reason why medication was prescribed - May provide more information on medications administered in hospital settings - May provide information about outcomes from clinical notes - May contain genomic/molecular test results - Multimodal EHRs contain clinical, genomic, and transcriptomic data appropriate for precision medicine or targeted drug development	Optum Panther, Flatiron Health, TriNetX, Clinical Practice Research Datalink (UK), Ontada iKnowMed, ConcertAI, Multimodal EHR (e.g., Tempus Labs)

(Continued)

Table 2 (Continued)

RWD source	Major data contributors	Limitations	Strengths	Example vendors
Patient disease or treatment registry	- Clinical research networks - Academic health centers - Commercial vendors (e.g., OM1) - Private-public partnerships (e.g., Critical Path Institute)	- Data and patient population captured depends on the original purpose/protocol of the registry - May not be suitable to use for other purposes - May not provide sufficient representation of a broad patient population - May have slow process, long wait time - Lack data on comparator groups: no controls for disease registries and no comparator data for drug registries - Important data elements (drug/comorbid medication dosing, duration, interruption data) may be missing unless registry development is not protocol or research hypotheses driven	- May represent single largest source of data for rare diseases - May have standardized data collection instruments and methods - May offer good longitudinal follow-up - May offer more detailed health information (e.g., disease severity) than claims or EHR - May span across regions and countries - Data collection is usually disease or drug focused and more comprehensive and granular	ImproveCareNow registry, Cystic Fibrosis Foundation Registry, International Collaborative Gaucher Group Gaucher Registry, Collaborative Trajectory Analysis Project

BMI, body mass index; EHR, electronic health record; ICD-10, International Classification of Disease-10th revision; RWD, real-world data.

Table 3 Common data elements found in different sources of RWD for clinical pharmacology applications

Category	Variable	Claims		
		Medical	Pharmacy	EHR
Patients	Age	X	X	X
	Sex	X	X	X
	Race/ethnicity	?	?	?
	Insurance coverage/type	X	X	
	Family history			X
	BMI			X
Healthcare providers	Identifier	?		
	Specialty	X		X
	Location	X		?
Encounter	Date	X		X
	Type (e.g., inpatient, outpatient)	X		X
	Procedure codes	X		X
	Diagnosis codes	X		X
Prescription medication	Generic/brand name		X	X
	Description (e.g., strength, formulation)		X	X
	Quantity (e.g., number, days supply)		X	
	Indication (reason for prescribing)		X	X
Clinical records	Measurements (e.g., vitals)			X
	Observations (e.g., notes)			X
	Rationale (e.g., reason for prescribing)			?
Diagnostic laboratory	Description			?
	Code (e.g., LOINC)			?
	Results			?

Blank, data generally unavailable; ?, uncertainty of data availability (depending on the source data vendor).

BMI, body mass index; EHR, electronic health record; LOINC, logical observation identifiers names and codes; RWD, real-world data; X, data generally available.

Given these tradeoffs, when selecting a source of RWD for clinical pharmacology purposes, it may be necessary to combine information from multiple data sources to answer all research questions. However, this process can be challenging

for biopharmaceutical companies because RWD are generally de-identified by data vendors prior to being shared. The data vendors (e.g., IQVIA and Optum) who have access to patient identifiers (e.g., name, date of birth, and health plan identifier)

must therefore be involved with the process of combining RWD databases. Traditionally, this was accomplished by licensing multiple types of RWD (e.g., claims and EHR) from one data vendor who would link databases with a common patient identifier prior to de-identifying and sharing the data externally. More recently, patient tokenization, for example, creation of a unique patient identifier that cannot be used to re-identify the patient without proprietary software, offers new options to combine sources of RWD. For example, companies with tokenization technology (e.g., Health Verity and Komodo) can now offer access to different types and sources of RWD (e.g., claims and EHRs) that can be linked by a common patient token prior to de-identification and licensing. As more RWD vendors use tokenization technology, the number and types of databases that can be linked to provide a more comprehensive understanding of patient health will increase. Perhaps more importantly, this same tokenization technology will also allow biopharmaceutical companies and clinical research organizations that generate patient-level data in clinical trials to link trial participants to external sources of RWD, allowing for hybrid study designs (e.g., clinical trial extensions with RWD). Furthermore, if combining RWD from multiple sources is not possible, another potential alternative is to design a prospective RWD study and intentionally capture the data and information needed in the real-world setting.⁷²

Of note, data quality and completeness issues (e.g., missing data) are well-known challenges of using RWD. Because commercially available sources of RWD generally contain data where both the patient and Healthcare Provider are de-identified, it is generally not possible for a user to issue queries to the data provider related to incomplete, missing, or questionable data. As a mitigation plan, researchers using RWD should be aware of this limitation, acknowledge that it may limit generalizability beyond the study population, perform feasibility analyses to determine the availability of specific variables of interest, consider sensitivity analyses to deal with outliers (e.g., exclude based on number of standard deviations from the mean), or develop methods to impute missing data (e.g., last observation carryforward).

Additional considerations and challenges for analyzing RWD to inform specific clinical pharmacology related areas (i.e., DDIs and organ impairment) are included in [Supplementary Material S1](#).

CONCLUSIONS AND FUTURE DIRECTIONS

Recent advances in RWD/RWE have created valuable sources of data and information. In addition to evidence from RCTs, RWD/RWE have been at the forefront of pharmaceutical innovation and informed decision making across the entire life cycle of a drug product including, but not limited to, discovery, clinical, regulatory/safety, value and access, and commercial. From the perspective of clinical pharmacologists, it is important to seek greater understanding and appreciation of the opportunities present at the intersection of RWE and clinical pharmacology, in terms of both the utilization of RWE to address specific clinical pharmacology questions as well as the application of quantitative clinical pharmacology approaches to analyze RWD and generate RWE. Case examples reviewed and discussed here cover both aspects and

beyond. Further understanding of the current scope and extent of interactions between RWE and clinical pharmacology will help guide strategies for broader and more advanced applications in the future.

There are several approaches that the clinical pharmacology community could adopt to achieve broader and more advanced applications of RWD/RWE. First, facilitate effective collaborations with various functions, including epidemiology/biostatistics, clinical sciences, biomarker/translational sciences, and medical affairs, and leverage their expertise in feasibility assessment, RWD source selection, patient population identification, and data extraction and data cleaning to enable successful use of RWD for evidence and insights generation. Ideally, the analysis of RWD should be undertaken by a cross-functional team with expertise in different domains, as outlined in [Table 4](#). Second, develop user-friendly RWD dashboards/interfaces to increase RWD accessibility and usage. Furthermore, include RWD/RWE in clinical pharmacology and/or MIDD plans at early stages of drug development, along with continual updates and assessments throughout the process, to increase the impact of RWD/RWE. In addition, engage with the health authorities and leverage the FDA's Advancing Real-World Evidence Program to get valuable input and feedback to improve the quality and acceptability of RWD/RWE for new intended labeling claims.¹⁰ With increased awareness, accessibility, and impactful applications of RWD/RWE in clinical pharmacology, we can better leverage opportunities for using RWE to address key drug development questions from a clinical pharmacology perspective in the future.

Table 4 Suggested cross-functional team for RWD project design and analysis for clinical pharmacology applications

Function	Expertise
Biostatistics	Data analysis, study design, and selecting appropriate statistical techniques to answer research questions
Clinical development	Evidence expected from product clinical development program and feasibility of traditional/prospective data collection
Clinical pharmacology	Research question, common data sources, expectations from various stakeholders, and data interpretation
Data engineering	Importing, verifying, and optimizing large sources of external data for analyses
Epidemiology	Observational research methods and study design to minimize bias and confounding
Medical affairs	Clinical domain, standards of care, and areas of concerns to healthcare providers
Real-world evidence	Available data sources from different vendors, types of variables they contain, and common/permitted uses
Regulatory affairs	Regulatory guidance related to research questions, regulatory precedents for using RWD, expectations from regulators
RWD programming	Efficient programming approaches to selecting and analyzing patient cohort of interest from larger RWD sources

RWD, real-world data.

SUPPORTING INFORMATION

Supplementary information accompanies this paper on the *Clinical Pharmacology & Therapeutics* website (www.cpt-journal.com).

ACKNOWLEDGMENTS

The authors would like to thank Mr Casey Kar-Chan Choong, Dr Dana Lu, Dr Samira Jamalian, Dr Wenhui Zhang, Dr Amita Joshi, Dr Parul Dayal, Dr Gemma Louise Dickinson, Dr Jay Erdman, and Dr Amit Desai for their scientific contributions to this manuscript.

FUNDING

No funding was received for this work.

CONFLICT OF INTEREST

R.Z., P.C., and B.V. are employees of Genentech/Roche and hold stock in the company. C.G., A.D.M., and P.G. are employees of Eli Lilly and Company and hold stock in the company. S.M., A.M., S.D., S.N., and B.T. are employees of Pfizer and hold stock in the company. G.D. is an employee of Takeda Pharmaceutical Company and holds stock in the company. P.M. is an employee of Astellas Pharma. I.Y. and S.I. are employees of Gilead Sciences and hold stock in the company. Z.M. is an employee of Sanofi and holds stock in the company.

© 2023 IQ Consortium. *Clinical Pharmacology & Therapeutics* published by Wiley Periodicals LLC 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](https://creativecommons.org/licenses/by-nc-nd/4.0/) 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.

- Duke-Margolis Center For Health Policy: A Framework for Regulatory Use of Real-World Evidence <https://healthpolicy.duke.edu/sites/default/files/2020-08/rwe_white_paper_2017.09.06.pdf>20Accessed%20October%206 (2017). Accessed October 6, 2022.
- Kim, H.S., Lee, S. & Kim, J.H. Real-world evidence versus randomized controlled trial: clinical research based on electronic medical records. *J. Korean Med. Sci.* **33**, e213 (2018).
- Rudrapatna, V.A. & Butte, A.J. Opportunities and challenges in using real-world data for health care. *J. Clin. Invest.* **130**, 565–574 (2020).
- 21st Century Cures Act. Public Law 114-255 <<https://www.congress.gov/bills/114th-congress/house-bill/34/>> (2016). Accessed 2022 October 6, 2022.
- The US Food and Drug Administration. PDUFA VI: Fiscal Years 2018-2022.
- Framework for FDA's Real-World Evidence Program <<https://www.fda.gov/media/120060/download>> (2018). Accessed 2022 October 6, 2022.
- Concato, J., Stein, P., Dal Pan, G.J., Ball, R. & Corrigan-Curay, J. Randomized, observational, interventional, and real-world-What's in a name? *Pharmacoepidemiol. Drug Saf.* **29**, 1514–1517 (2020).
- Concato, J. & Corrigan-Curay, J. Real-world evidence – where are we now? *N. Engl. J. Med.* **386**, 1680–1682 (2022).
- The US Food and Drug Administration. Real-world data (RWD) and real-world evidence (RWE) are playing an increasing role in health care decisions <<https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence>> (2022). Accessed 2022 October 6, 2022.
- The US Food and Drug Administration. Advancing Real-World Evidence Program <<https://www.fda.gov/drugs/development-resources/advancing-real-world-evidence-program>> (2022). Accessed October 30, 2022.
- Arlett, P., Kjaer, J., Broich, K. & Cooke, E. Real-world evidence in EU medicines regulation: enabling use and establishing value. *Clin. Pharmacol. Ther.* **111**, 21–23 (2022).
- Flynn, R. *et al.* Marketing authorization applications made to the European Medicines Agency in 2018-2019: what was the contribution of real-world evidence? *Clin. Pharmacol. Ther.* **111**, 90–97 (2022).
- Nishioka, K., Makimura, T., Ishiguro, A., Nonaka, T., Yamaguchi, M. & Uyama, Y. Evolving acceptance and use of RWE for regulatory decision making on the benefit/risk assessment of a drug in Japan. *Clin. Pharmacol. Ther.* **111**, 35–43 (2022).
- Dagenais, S., Russo, L., Madsen, A., Webster, J. & Becnel, L. Use of real-world evidence to drive drug development strategy and inform clinical trial design. *Clin. Pharmacol. Ther.* **111**, 77–89 (2022).
- Liu, Q., Ramamoorthy, A. & Huang, S.M. Real-world data and clinical pharmacology: a regulatory science perspective. *Clin. Pharmacol. Ther.* **106**, 67–71 (2019).
- Ramamoorthy, A. & Huang, S.M. What does it take to transform real-world data into real-world evidence? *Clin. Pharmacol. Ther.* **106**, 10–18 (2019).
- Zhao, X., Iqbal, S., Valdes, I.L., Dresser, M. & Girish, S. Integrating real-world data to accelerate and guide drug development: a clinical pharmacology perspective. *Clin. Transl. Sci.* **15**, 2293–2302 (2022).
- Dechanont, S., Maphanta, S., Butthum, B. & Kongkaew, C. Hospital admissions/visits associated with drug-drug interactions: a systematic review and meta-analysis. *Pharmacoepidemiol. Drug Saf.* **23**, 489–497 (2014).
- Duke, J.D. *et al.* Literature based drug interaction prediction with clinical assessment using electronic medical records: novel myopathy associated drug interactions. *PLoS Comput. Biol.* **8**, e1002614 (2012).
- Lorberbaum, T. *et al.* Coupling data mining and laboratory experiments to discover drug interactions causing QT prolongation. *J. Am. Coll. Cardiol.* **68**, 1756–1764 (2016).
- Yee, S.W. *et al.* Drugs in COVID-19 clinical trials: predicting transporter-mediated drug-drug interactions using In vitro assays and real-world data. *Clin. Pharmacol. Ther.* **110**, 108–122 (2021).
- Bourliere, M. *et al.* Ledipasvir-sofosbuvir with or without ribavirin to treat patients with HCV genotype 1 infection and cirrhosis non-responsive to previous protease-inhibitor therapy: a randomised, double-blind, phase 2 trial (SIRIUS). *Lancet Infect. Dis.* **15**, 397–404 (2015).
- Cherney, D.Z.I. *et al.* Efficacy and safety of sotagliflozin in patients with type 2 diabetes and severe renal impairment. *Diabetes Obes. Metab.* **23**, 2632–2642 (2021).
- Parker, V.E.R. *et al.* Efficacy and safety of cotadutide, a dual glucagon-like peptide-1 and glucagon receptor agonist, in a randomized phase 2a study of patients with type 2 diabetes and chronic kidney disease. *Diabetes Obes. Metab.* **24**, 1360–1369 (2022).
- Wei, L. *et al.* Glecaprevir-pibrentasvir to treat chronic hepatitis C virus infection in Asia: two multicentre, phase 3 studies – a randomised, double-blind study (VOYAGE-1) and an open-label, single-arm study (VOYAGE-2). *Lancet Gastroenterol. Hepatol.* **5**, 839–849 (2020).
- Weinberg, E.M. *et al.* Multicenter, double-blind, randomized trial of Emricasan in hepatitis C-treated liver transplant recipients with residual fibrosis or cirrhosis. *Liver Transpl.* **27**, 568–579 (2021).
- Sane, R. *et al.* Pharmacokinetics of Ipatasertib in subjects with hepatic impairment using 2 methods of classification of hepatic function. *J. Clin. Pharmacol.* **62**, 171–181 (2022).
- Lu, D. *et al.* Application of a two-analyte integrated population pharmacokinetic model to evaluate the impact of intrinsic and extrinsic factors on the pharmacokinetics of Polatuzumab Vedotin in patients with non-Hodgkin lymphoma. *Pharm. Res.* **37**, 252 (2020).
- Spillane, S. *et al.* Organ dysfunction in patients with advanced melanoma treated with immune checkpoint inhibitors. *Oncologist* **25**, e1753–e1762 (2020).
- Sybing, R. *et al.* American Society for Clinical Pharmacology and Therapeutics: CHARACTERIZATION OF HYPERFILTRATION IN SICKLE CELL DISEASE PATIENTS USING REAL WORLD DATA. *Clin Pharmacol Ther* **111**(Suppl 1), S74 (2022).

31. European Medicines Agency. ICH guideline E11A on pediatric extrapolation, Step 2b (2022).
32. Chanu, P. *et al.* Model-based approach for methoxy polyethylene glycol-epoetin beta drug development in paediatric patients with anaemia of chronic kidney disease. *Br. J. Clin. Pharmacol.* **86**, 801–811 (2020).
33. Zhang, W. Application of real world data (RWD) in clinical pharmacology. AAPS National Biotechnology Conference (2021). p. Rapid Fire Oral Presentation Session.
34. Lukka, P.B., Woods, M., Chhim, R., Phelps, S.J., Wheless, J.W. & Meibohm, B. Use of real-world data and Pharmacometric modeling in support of Lacosamide dosing in pediatric patients under 4 years of age. *J. Clin. Pharmacol.* **61**, 881–888 (2021).
35. Madabushi, R., Seo, P., Zhao, L., Tegenge, M. & Zhu, H. Review: role of model-informed drug development approaches in the lifecycle of drug development and regulatory decision-making. *Pharm. Res.* **39**, 1669–1680 (2022).
36. Delor, I., Charoin, J.E., Gieschke, R., Retout, S. & Jacqmin, P. Modeling Alzheimer's disease progression using disease onset time and disease trajectory concepts applied to CDR-SOB scores from ADNI. *CPT Pharmacometrics Syst. Pharmacol.* **2**, e78 (2013).
37. Jamalian, S. *et al.* Development of a disease progression model for Alzheimer's disease informed by multiple clinical trials and ADNI to predict longitudinal trajectory of CDR-SB score. *Clinical Trials on Alzheimer's Disease (CTAD)* (2020).
38. Boucher, M., Harnisch, L. & Riley, S. Longitudinal modeling of disease-progression in transthyretin familial polyneuropathy with tafamidis. International Society of Amyloidosis Conference. Kumamoto, Japan (2018). p. Poster PA026.
39. Wang, X. *et al.* Disease progression modelling for Duchenne muscular dystrophy using north star ambulatory assessment and forced vital capacity. 10th American Conference on Pharmacometrics. Orlando, FL (2019).
40. Abrams, R. *et al.* A quantitative systems pharmacology model of Gaucher disease type 1 provides mechanistic insight into the response to substrate reduction therapy with eliglustat. *CPT Pharmacometrics Syst. Pharmacol.* **9**, 374–383 (2020).
41. Bruno, R. *et al.* Progress and opportunities to advance clinical cancer therapeutics using tumor dynamic models. *Clin. Cancer Res.* **26**, 1787–1795 (2020).
42. Chanu, P. *et al.* A disease model for multiple myeloma developed using real world data. Population Approach Group Europe (PAGE) (2021) PAGE 29 (2021) Abstract 9878.
43. Kotani, N., Nina, W.N., Dhillon, P.K., Arndorfer, S. & Chanu, P. A model based approach to infer overall survival (OS) in hormone receptor positive (HR+) human epidermal growth factor receptor 2 negative (HER2-) metastatic breast cancer (HR+/HER2- mBC). Population Approach Group Europe (PAGE) (2021). PAGE 29 (2021) Abstract 9655.
44. Julian, C. *et al.* Real-world data prognostic model of overall survival in patients with advanced NSCLC receiving anti-PD-1/PD-L1 immune checkpoint inhibitors as second-line monotherapy. *Cancer Rep. (Hoboken)* **5**, e1578 (2022).
45. Yun, H. *et al.* Real World Validation of a Rule to Predict Response to Sarilumab in Patients with Rheumatoid Arthritis: Analysis from the RISE Registry. American College of Rheumatology 2021. ACR (2021) Abstract 1243.
46. Eichler, H.G. *et al.* Randomized controlled trials versus real world evidence: neither magic nor myth. *Clin. Pharmacol. Ther.* **109**, 1212–1218 (2021).
47. Slattey, J. & Kurz, X. Assessing strength of evidence for regulatory decision making in licensing: what proof do we need for observational studies of effectiveness? *Pharmacoepidemiol. Drug Saf.* **29**, 1336–1340 (2020).
48. Corrigan-Curay, J., Sacks, L. & Woodcock, J. Real-world evidence and real-world data for evaluating drug safety and effectiveness. *JAMA* **320**, 867–868 (2018).
49. Chodankar, D. Introduction to real-world evidence studies. *Perspect. Clin. Res.* **12**, 171–174 (2021).
50. Erdman, J. *et al.* Lung transplant outcomes in adults in the United States: retrospective cohort study using real-world evidence from the SRTR. *Transplantation* **106**, 1233–1242 (2022).
51. Wedam, S. *et al.* FDA approval summary: palbociclib for male patients with metastatic breast cancer. *Clin. Cancer Res.* **26**, 1208–1212 (2020).
52. Kraus, A.L. *et al.* Real-world data of Palbociclib in combination with endocrine therapy for the treatment of metastatic breast cancer in men. *Clin. Pharmacol. Ther.* **111**, 302–309 (2022).
53. Pfizer: IBRANCE- palbociclib capsule (2022).
54. Lamba, M. *et al.* Model-informed development and registration of a once-daily regimen of extended-release tofacitinib. *Clin. Pharmacol. Ther.* **101**, 745–753 (2017).
55. Cohen, S.B. *et al.* Real-world evidence to contextualize clinical trial results and inform regulatory decisions: tofacitinib modified-release once-daily vs immediate-release twice-daily for rheumatoid arthritis. *Adv. Ther.* **38**, 226–248 (2021).
56. Hey, S.P. & Truog, R.D. The question of clinical equipoise and Patients' best interests. *AMA J. Ethics* **17**, 1108–1115 (2015).
57. ICH E10. Choice of control group and related issues in clinical trials <https://database.ich.org/sites/default/files/E10_Guide_line.pdf> (2000).
58. Thorlund, K., Dron, L., Park, J.J.H. & Mills, E.J. Synthetic and external controls in clinical trials – a primer for researchers. *Clin. Epidemiol.* **12**, 457–467 (2020).
59. Macaulay, R., Cable, G., Siddiqui, M.K. & Lucas, J. RW2 synthetic control arms: hype versus reality, the payer perspective. *Value Health* **23**, S758 (2020).
60. The US Food and Drug Administration. Center for Drug Evaluation and Research: Avelumab BLA 761049 Multidisciplinary Review and Evaluation <https://www.accessdata.fda.gov/drugsatfda_docs/nda/2017/761049Orig1s000MultidisciplineR.pdf> (2017). Accessed August 15, 2022.
61. The US Food and Drug Administration. Center for Drug Evaluation and Research: Blinatumomab BLA 125557 S-013 Multidisciplinary Review and Evaluation <https://www.accessdata.fda.gov/drugsatfda_docs/nda/2018/125557Orig1s013.pdf> (2018). Accessed August 15, 2022.
62. Popat, S. *et al.* Addressing challenges with real-world synthetic control arms to demonstrate the comparative effectiveness of Pralsetinib in non-small cell lung cancer. *Nature Commun.* **13**, 1–10 (2022).
63. Ayodele, O. *et al.* Change in estimated glomerular filtration rate in adult patients with chronic hypoparathyroidism treated with rhPTH (1–84) compared with a historical control cohort. *J. Endocr. Soc.* **5**, A257–A258 (2021).
64. Chen, K.S. *et al.* Five-year estimated glomerular filtration rate in patients with hypoparathyroidism treated with and without rhPTH(1–84). *J. Clin. Endocrinol. Metab.* **105**, e3557–e3565 (2020).
65. Gosmanova, E. *et al.* Risk of chronic kidney disease in adult patients with chronic hypoparathyroidism treated with rhPTH (1–84) compared with a historical control cohort. *J. Endocr. Soc.* **5**, A267 (2021).
66. Miksad, 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).
67. The US Food and Drug Administration. Real-World Data: Assessing Electronic Health Records and Medical Claims Data to Support Regulatory Decision Making for Drug and Biological Products (2021).
68. Rea, S., Bailey, K.R., Pathak, J. & Haug, P.J. Bias in recording of body mass index data in the electronic health record. *AMIA Jt Summits Transl. Sci. Proc.* **2013**, 214–218 (2013).
69. Goodloe, R., Farber-Eger, E., Boston, J., Crawford, D.C. & Bush, W.S. Reducing clinical noise for body mass index measures due to unit and transcription errors in the electronic health record. *AMIA Jt Summits Transl. Sci. Proc.* **2017**, 102–111 (2017).

70. The US Food and Drug Administration. Real-World Data: Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products Guidance for Industry (2021).
71. Muntoni, F. *et al.* Real-world and natural history data for drug evaluation in Duchenne muscular dystrophy: suitability of the north star ambulatory assessment for comparisons with external controls. *Neuromuscul. Disord.* **32**, 271–283 (2022).
72. Villarreal-Garza, C. *et al.* Improving collection of real-world data: the experience of the Joven & Fuerte Prospective Cohort for Mexican young women with breast cancer. *Clin. Breast Cancer* **21**, e675–e680 (2021).