Novel clinical trial design and analytic methods to tackle challenges in therapeutic development in rare diseases

Yimei Li[1], **Rima Izem**[2]

[1]Department of Biostatistics, Epidemiology, and Informatics, University of Pennsylvania Perelman School of Medicine, PA, USA; [2]Statistical Methodology and Consulting, Novartis, Basel, Switzerland

Contributions: (I) Conception and design: Both authors; (II) Administrative support: None; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: Both authors; (V) Data analysis and interpretation: Both authors; (VI) Manuscript writing: Both authors; (VII) Final approval of manuscript: Both authors.

Correspondence to: Yimei Li. Department of Biostatistics, Epidemiology, and Informatics, University of Pennsylvania Perelman School of Medicine, 423 Guardian Dr, 626 Blockley Hall, Philadelphia, PA 19104, USA. Email: yimeili@pennmedicine.upenn.edu.

Abstract: While only a fraction of the worldwide population may have a particular rare disorder, millions of people worldwide are affected across the over 6,000 rare disorders and do not have a safe and effective approved therapy to help them live or manage complications from the disorder. Challenges to clinical development of new therapies in rare disorders include difficulty in powering and recruiting into a study in small and often heterogenous population, scarcity of natural history data informing critical design elements such as endpoint selection and study duration, and ethical and recruitment challenges in randomizing patients to a placebo arm. In this review, we describe some existing and novel strategies to tackle these challenges, by efficient utilization of available resources. We discuss the role of natural history studies and endpoint selection as they remain critical features that apply across designs and disorders. We also review some novel clinical trial designs including incorporating external control and/or longitudinal measures, master protocol designs, and adaptive designs. Additionally, we review some analytic strategies that are often associated with these designs, such as the use of causal inference methods, and Bayesian methods. We hope this review will raise awareness of these novel approaches and encourage their use in studies of rare diseases.

Keywords: Rare disease; clinical trial; external or self-control; master protocol; adaptive design

Submitted Nov 05, 2021. Accepted for publication Feb 28, 2022.
doi: 10.21037/atm-21-5496
View this article at: https://dx.doi.org/10.21037/atm-21-5496

Introduction

Developing novel therapies in rare disorders is working to serve millions of lives and giving them access to therapies they need to live or thrive. While only a fraction of the worldwide population may have a particular rare disorder, at any time 260 million lives worldwide are affected across the over 6,000 identified rare disorders (1,2). This is an area of high unmet need as only 20% of rare disorders have approved therapies.

There are many obstacles to the endeavor of developing novel therapies in rare disorders (3). Compared to therapeutic development in more common indications, the small and heterogenous population in a rare disorder increases the logistical burden of recruitment and limits the statistical power of any study. Some of these disorders have poorly understood natural history with yet to be characterized fit-for-purpose biomarkers or endpoints measuring benefits or risks, and poorly understood timing for measurement of the impact of novel interventions. Finally, when therapies are limited or timing to assess benefit is long, randomizing subjects to a placebo arm may not be feasible or ethical.

This paper reviews some existing and novel strategies to efficiently use available resources in therapeutic development in rare disorders. While several review of design and analyses strategies already exist in the rare disease setting, including for example (4-6), this review updates these strategies with more recent innovations and guidelines. Section 'General considerations' discusses the role of natural history studies and endpoint selection as they remain critical features that apply across designs and disorders. Then, Section 'Examples of novel study designs' reviews recent design features including use of novel sources, external control, longitudinal designs, master protocol designs, and adaptive design. Additionally, Section 'Examples of novel analytical strategies' reviews recent developments in analysis, such as the use of causal inference methods, and Bayesian methods. While we discuss each of the designs and the analyses separately, in practice some of these strategies are combined for added efficiencies, and this will be summarized in our conclusions.

General considerations

The role of natural history studies

Understanding the experience of subjects living with a rare disorder through natural history studies is critical to supporting development of new therapies (7). These data can be collected by different stakeholders including academic, government, or industry organizations. They vary in scope from a small set of volunteers with a rare disease diagnosis to a quasi-census of everyone diagnosed with the disorder. For example, the cystic fibrosis foundation patient registry (8) collects natural history data from volunteers. This registry is a poster-child of what these studies can accomplish in assessing the patient's needs and the disease burden, and supporting and partnering in therapy development (9). Yearly reports from this longitudinal study present a snapshot of the incidence and prevalence of this disorder, and the population geographic and medical diversity. The report also includes population-level longitudinal information on progression of signs, symptoms, and associated treatments, as a subject ages, and as new diagnostics or therapies enter the market. Mining this natural history study informs identification of risk factors, and development of guidelines for treatment and management of the disorder. Finally, a rare disorder natural history study itself can contribute data to comparative effectiveness and safety studies, as illustrated by the use of the urea cycle disorder natural history study to assess the effectiveness of liver transplantation (10) using causal inference methods discussed in Section 'Causal inference and pharmacoepidemiology method'.

Beyond being a data depository of medical history or current medical practice, natural history studies can help development of new therapies in rare disorders in multiple ways. For example, understanding the existing information can help identify the gaps in needs that a new therapy can fill and justify prospectively planned study design attributes such as population, endpoints, treatment comparators, and ideal duration of a study to characterize benefits and risks. Also, disorder-specific patient registries represent an existing infrastructure reuniting a community of patients, their families, and their physicians that can be leveraged to invite participants in new clinical trials. This infrastructure facilitates recruitment, access to data collected in the registry before and after the clinical trials, and dissemination of clinical trial results. Such infrastructure is exploited by master protocol studies discussed in Section 'Master protocols'. Lastly, summary-level or subject-level data from natural history studies can themselves serve as an external control to prospectively planned single arm clinical trials as will be further discussed in Section 'External or historical controls'.

Despite the multiple benefits of these data, starting, growing, and maintaining natural history studies is resource intensive when the benefits may only materialize in the long run when sufficient information has accrued. Thus, leveraging existing networks of professional medical societies, patient interest groups, or healthcare organizations can facilitate the initiation, the success, and the sustainability of these studies. Such networks can help refine the questions of interest, the metrics of interest, and lower the burden to patients and providers of new measures in epidemiological studies or clinical trials. Examples of these networks include the US-National Institute of Health Rare Diseases Clinical Research Network (11), and the US-National Organization of Rare Diseases (12). Additionally, retrospective chart reviews or mining data sources from electronic healthcare records (EHR) and claims, for example Patient-Centered Outcome Research Network (13), can be a starting point for a natural history study or can augment prospective data collection within the study with concurrently collected data in routine care.

Endpoint

In therapeutic development in new disorders, the selection and justification of endpoints measuring how a patient feels or functions contribute to the success of clinical studies.

This selection is ideally informed by the natural history of the disease but also by the mechanism of action of the new therapy. The choice has also consequences on study power, duration, and ability to measure change (3).

In cases where existing clinical measurements used in routine care are not sufficient to measure benefits or risks, a clinical study may serve both to prove benefit but also to qualify new biomarkers, new digital endpoint, new patient reported outcomes, or new clinical outcome assessments. This typically entails proving that this new endpoint is fit-for-purpose using standalone or embedded validation studies in the clinical program (14-16). For example, the clinical development for a gene therapy voretigene neparvovec to treat inherited retinal atrophy included investigation and validation of a novel endpoint measuring improvement in vision by rating each subject's success navigating an obstacle course (17,18).

The increased use of wearables, digital technology, and tele-health in clinical care has impacted the design and the conduct of clinical trials, including the choice and assessment of endpoints. Novel digital endpoints have the potential to decrease the burden of participation of patients in clinical trials while increasing the frequency of assessments and generalizability of results (19). Also, high-frequency data collection can potentially enable building disease progression models and predict individual trajectories.

Another strategy of endpoint selection is to develop or use a composite. Many rare disorders are multi-symptomatic and using a composite endpoint grouping clinically relevant signs and symptoms into one summary measure may be a useful strategy to boost study power and measure clinical benefit. For example, the development program for cerliponase alfa for CLN2 disease used the motor and language scales of a multi-item questionnaire (20). Important considerations when constructing the composite are to group elements that are expected to be similarly impacted by the therapy (go in the same direction), to explore the contribution of each element of the composite in the findings, and to address concerns of multiple testing in the analysis (21,22).

Efficient use of information and duration of follow-up is also critical in studies with dichotomous endpoints, such as toxicity in phase I and complete response in phase II oncology trials. In those situations, patients may not all have completed follow-up for endpoint evaluation at the time of a planned interim analysis. A traditional approach is to suspend the trial to wait for outcomes becoming available on the already enrolled patients. For example, in phase I trial the enrollment is suspended after each dose cohort, and in two-stage phase II trial the enrollment is suspended after reaching the sample size for the first stage. This may lengthen the trial, dampen the enthusiasm of participating investigators, and turn away eligible patients as the accrual is halted (23). An alternative strategy is to re-define the outcome as a time-to-event endpoint such as time to toxicity or time to response in earlier examples. In this way, the partial information observed from patients with incomplete follow up would contribute to the interim analysis. Examples of such designs include time-to-event continual reassessment method (TITE-CRM) (24) and time-to-event Bayesian optimal interval design (TITE-BOIN) (25) in phase I settings, and time-to-event Bayesian optimal phase II (TOP) design in phase II settings (26).

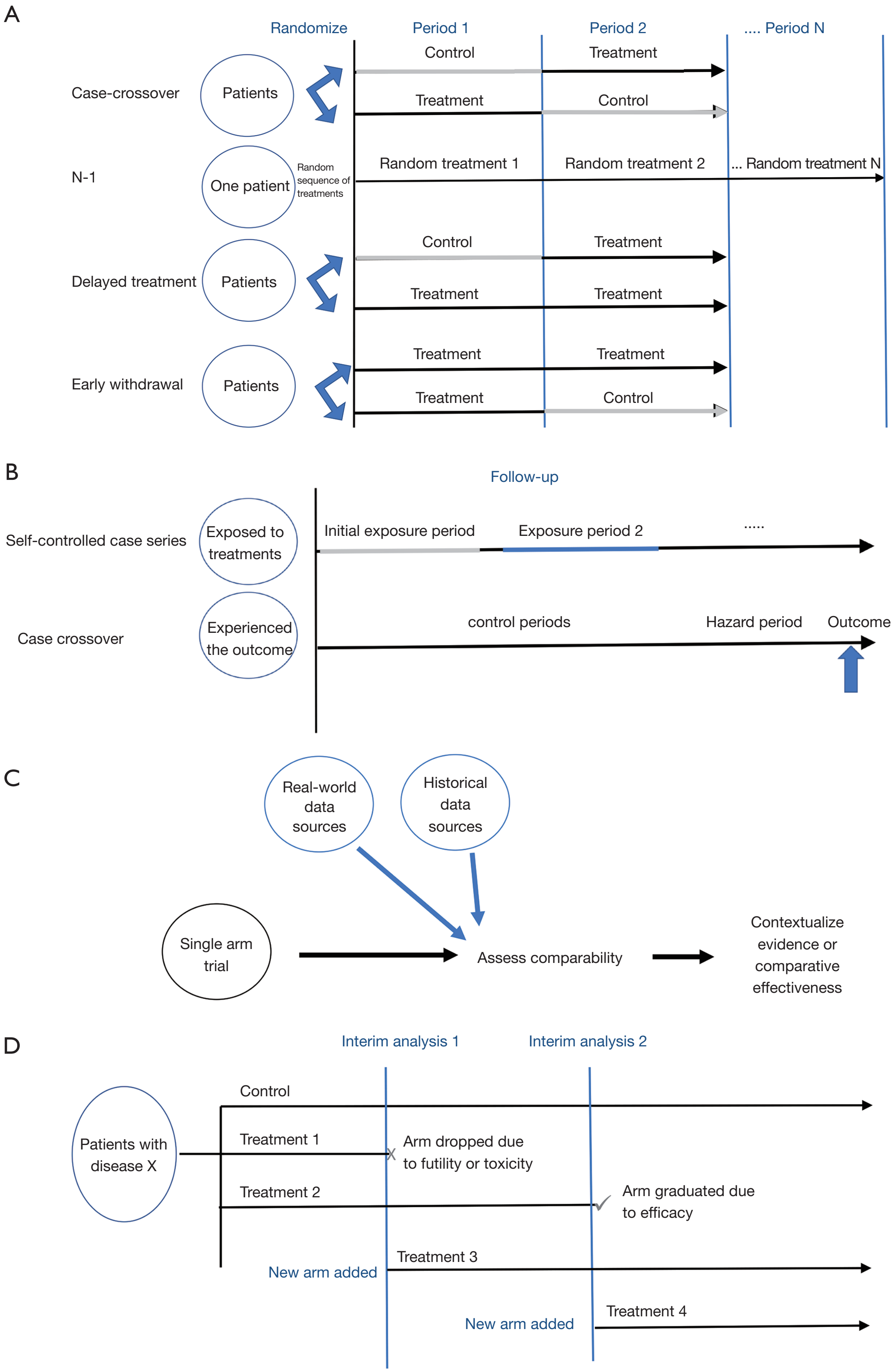
Examples of novel study designs

This section reviews some novel design strategies or new use of these strategies in rare diseases. Graphic representations of these designs are shown in *Figure 1*. Considerations for use and benefit of each strategy are summarized in *Table 1*.

External or historical controls

In settings where randomizing to a control group is unfeasible or unethical, clinical development programs of new therapies in rare diseases have sometimes relied on single arm trials to evaluate benefits and risks. Interpreting the findings typically requires contextualization of the evidence from these trials compared to existing standard of care. Traditionally, this contextualization involved expert opinion, literature review, or meta-analyses of previous studies on standard of care.

More recently, when subject-level data is available from the external control, a side-by-side comparison to a concurrent or historical external control arm is derived and presented (27-29). A recent application is the use of natural history data as an external control to the pivotal study of cerliponase alfa (30). Availability of subject-level data opens the possibilities for incorporating prior expertise as well as adjusting for known differences between the external control and the clinical study. This information can be used to derive a threshold or prior belief on the distribution of outcomes as in Bayesian methods for meta-analysis described in Section 'Borrowing evidence'. Subject-level data can also be used for a direct side-by-side comparison to


A
Randomize
Period 1
Period 2
.... Period N
Case-crossover
Patients
Control
Treatment
Treatment
Control
N-1
One patient
Random sequence of treatments
Random treatment 1
Random treatment 2
... Random treatment N
Delayed treatment
Patients
Control
Treatment
Treatment
Treatment
Early withdrawal
Patients
Treatment
Treatment
Treatment
Control
B
Follow-up
Self-controlled case series
Exposed to treatments
Initial exposure period
Exposure period 2
.....
Case crossover
Experienced the outcome
control periods
Hazard period
Outcome
C
Real-world data sources
Historical data sources
Single arm trial
Assess comparability
Contextualize evidence or comparative effectiveness
D
Interim analysis 1
Interim analysis 2
Patients with disease X
Control
Treatment 1
Arm dropped due to futility or toxicity
Treatment 2
Arm graduated due to efficacy
New arm added
Treatment 3
New arm added
Treatment 4

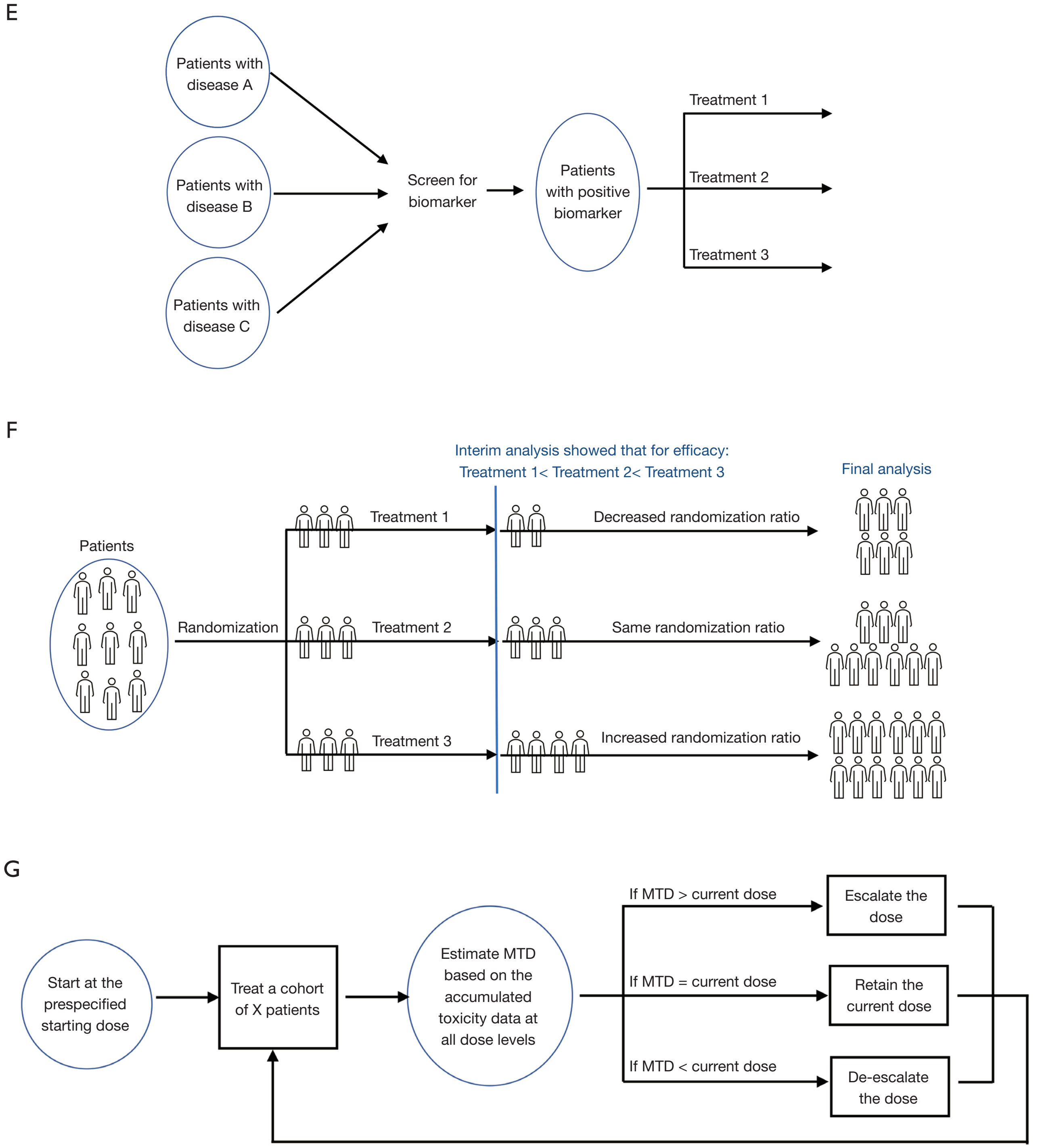

E
Patients with disease A
Patients with disease B
Patients with disease C
Screen for biomarker
Patients with positive biomarker
Treatment 1
Treatment 2
Treatment 3
F
Interim analysis showed that for efficacy: Treatment 1< Treatment 2< Treatment 3
Final analysis
Patients
Randomization
Treatment 1
Decreased randomization ratio
Treatment 2
Same randomization ratio
Treatment 3
Increased randomization ratio
G
Start at the prespecified starting dose
Treat a cohort of X patients
Estimate MTD based on the accumulated toxicity data at all dose levels
If MTD > current dose
Escalate the dose
If MTD = current dose
Retain the current dose
If MTD < current dose
De-escalate the dose

Figure 1 Graphic representations of the novel designs. (A) Randomized studies (incorporating self-control): case-crossover, N-1, delayed treatment, or early withdrawal. (B) Non-random longitudinal studies: self-controlled case-series, case-crossover, longitudinal studies. (C) External controls. (D) Platform trials. (E) Basket trials. (F) Response adaptive randomization. (G) Phase I adaptive trials.

Table 1 Considerations for use of various non-traditional study designs

Design	Specific type	Considerations for use	Benefit of the design
External control		Evaluate the data source capture of critical information to the effectiveness or safety question of interest (exposure, outcome, potential confounders) Evaluate comparability of design attributes (population selection criteria, treatment, and outcome) of the external control to the clinical trial Identify potential sources of bias (e.g., selection bias, information bias, immortal time bias) Design the study using the estimand and target trial framework to minimize bias Pre-specify protocol and plan for sensitivity analyses	A good external control can potentially augment or replace an internal control group External control data could potentially already exist saving resources for novel data collection
Longitudinal designs (units of analyses are periods within subjects)	Randomized (e.g., case-crossover, cluster randomized design, treatment withdrawal design)	Better suited for therapy and outcome pairs that can be observed within a short time period Some important considerations are the duration of each period, identifying and adjusting for any time-varying confounding (e.g., order effect, learning), and adjusting for correlation of measures within each subject using random subject effect or hierarchical models	#1-Analyses units in these designs are determined by number of subjects and time periods. Thus, even one subject can still contribute many analyses units investigating different therapies #2-Can uses randomization in different manners to answer questions about duration of therapy across subject or time periods
	Observational (e.g., self-controlled case series, case-crossover, nested controls)	All of above, and assuming exchangeability of time periods within a subject	Same as #1 above
Master protocol	General	Require upfront investment in the infrastructure Design is more complex, both logistically and statistically. Best suited for outcomes assessed within a short time frame of therapy, so that interim analyses/decisions can be updated	Once the infrastructure is set up, could easily add new experimental therapies, which expedites the discovery in the long run Could enhance patient participation due to the increased likelihood of eligibility for at least one of the treatments under study
	Platform	Suited for therapeutic areas where multiple therapies are being developed Concurrency of a treatment arm and the control arm needs to be considered	The use of the shared control can reduce the overall sample size substantially
	Basket	Suited for the scenario where the same biological pathway is shared by multiple disease population	Enable the evaluation of rare mutations that are difficult to study solely within a disease-specific context

Table 1 (*continued*)

Table 1 *(continued)*

Design	Specific type	Considerations for use	Benefit of the design
Adaptive design	General	Should be used with appropriate methods to control for type 1 error inflation from multiple testing in Phase II and Phase III studies	Allow alterations to trial procedures and/or statistical procedures of the trial to maximize the efficiency
	Response adaptive randomization (RAR)	Response needs to be short Consider potential issue of non-concurrency of the control due to shifting of randomization probabilities Cautions needed as in some extreme cases RAR design could have lower power than standard 1:1 randomization design	Could enhance patient participation due to the potential higher chance to get the new therapy Increase the available information for the most effective treatment(s)
	Phase I	Adaptive phase I designs are more efficient and flexible than the standard 3+3 designs so should always be considered Hybrid phase I/II designs should be considered when responses can be evaluated in a short time	Leads to higher probability of selecting the correct MTD and potentially could find MTD faster so yields smaller total sample size

MTD, maximum tolerable dose.

the current study using causal inference methods described in Section 'Causal inference and pharmacoepidemiology method'.

Whichever methods are used, important considerations for a meaningful comparison are that of similarity of multiple attributes: cohort characteristics, inclusion and exclusion criteria into the cohort, outcome ascertainment, anchor time for start of follow-up, duration of follow-up, similarity of measurement and handling of intercurrent events. When standard of care rapidly evolves, concurrency of the external control to the clinical trial is also important. Lack of similarity or concurrency does not always preclude use of the external control data but may call for some discounting of the evidence either qualitatively or quantitatively using some of the methods described in Section 'Borrowing evidence' (e.g., robust prior or power prior). They could also call for additional sensitivity analyses, such as quantitative bias analyses in causal inference, to determine the impact of a potential dissimilarity on the comparison.

Designs incorporating longitudinal measures

Several designs can leverage longitudinal information, or repeated measures, on the same patient at different times to increase units of analysis and possibly reduce heterogeneity relative to a parallel arm study with a unique endpoint assessment for each subject. As discussed by several other authors including (4,31,32), longitudinal designs are particularly relevant in those rare diseases and therapies where a short observation period is sufficient to measure treatment effect and no strong carry-over effect. The analytical sample size in longitudinal studies is determined by the number of periods across subjects, rather than subjects (4,31,32). Heterogeneity can decrease when the variability within a subject is smaller than between subjects.

These longitudinal designs include pure self-controlled studies where treatment responses are compared within each subject and more general repeated measures-designs augmenting between-subject comparison with within-subject comparison. Pure self-controlled studies include crossover and N-of-1 studies. In a crossover design, each subject contributes multiple treatment periods separated by a washout period with a randomly assigned sequence of therapies. Thus, when comparing a new therapy to a control in a crossover design, patients are randomized to one of two groups, either receiving the new therapy first then the control, or receiving the control first then the new therapy.

The N-of-1 design includes only one subject contributing multiple periods with randomization of treatments to each period.

With repeated measures, one can also augment between-subject comparison in parallel randomized studies with within-subject comparison. These designs can also have a period as the unit of analysis and compare treatment and controls cross-sectionally and within a subject. Thus, these designs have typically more power than a parallel control arm of the same number of subjects because each subject contributes more than one analysis unit and the within-subject variability is typically lower than the between-subject variability.

A recent example of exploiting repeated measures and their use include a randomized withdrawal designs in the Phase III study investigating the efficacy of pegvaliase in treating Phenylketonuria (33). This study randomized responders to therapy, determined after an assessment period, to either discontinue or continue therapy for eight weeks. Another example is the randomized study investigating the efficacy of N-carbamylglutamate in reducing ammonia levels in subjects with urea cycle disorders (34). In that study, each subject could contribute multiple hospitalization periods and treatments were randomized to each period. Analyses adjusted for correlation of measures using a subject random effect in the mixed effect outcome model.

While less well known or used in rare disorders, natural history studies can also leverage repeated measures to assess comparative safety or effectiveness (35). These designs include self-controlled case series (36), or case crossover designs (37,38) often used in comparative safety for rare outcomes. These designs also include causal inference methods adjusting for time-varying confounding (39). For example, a previously mentioned study leveraged a natural history study to assess comparative effectiveness of liver transplantation compared to medical management in urea cycle disorder (10).

Master protocols

In contrast to traditional clinical trials that focus on one experimental therapy in one disease population, master protocol is a new type of study design that attempts to evaluate multiple experimental therapies in one or multiple indications, under one overarching protocol (40-42). A platform trial is one type of master protocol that randomizes patients to a common control arm and many different experimental arms, where new experimental arms could be added over time. Futility and/or efficacy stopping rules are often built in the design to allow the experimental agents to enter and exit the trial based on the interim analyses results (43,44). Because a shared control arm is used for efficacy evaluations of each experimental therapy, such a design can reduce the overall sample size substantially and be much more efficient for finding effective therapies than multiple stand-alone trials (45,46). This feature is particularly useful for rare diseases where patient numbers are very limited. As the experimental agents could enter the study at different times, it is important to consider the time that the treatment arm was active for recruitment. When the control arm data are outside of this period, the validity of the comparison needs to be evaluated carefully, considering potential temporal change in the control data. Examples of platforms trials include I-SPY2 (47) in breast cancer, BATTLE (48) and LUNG-MAP (49) in lung cancer, DIAN-TU in Alzheimer's disease (50), and NTUITT-NF2 in patients with neurofibromatosis type 2 (51).

Another type of master protocol is basket trials, whose study population is defined by the presence of a particular biomarker or molecular alteration, rather than a particular disease type as in a traditional trial (40-42). A basket trial could have multiple baskets, each defined by a distinct set of biomarkers and treated with a matched targeted therapy. Because basket trials are disease agnostic, they provide great potential for patients with rare diseases to be eligible to participate in clinical trials and enable the evaluation of rare mutations that are difficult to study solely within a disease-specific context (52). Examples of basket trials include NCI MATCH (53), BRAF-vemurafenib trial (54), and DART trial for rare cancers (55). One important issue in designing basket trials is how to utilize information collected across disease types, and multiple designs have been proposed in the literature, such as Bayesian hierarchical model (BHM) (56), calibrated BHM (57), Bayesian latent subgroup trial (BLAST) design (58), and robust exchangeability-nonexchangeability (EXNEX) design (59). While BHM could lead to inflated type I error rate, the rest of the designs were able to obtain desired type I error control.

Some considerations with master protocol designs are that the potential gains in recruitment and sample size require an upfront investment in the infrastructure for setting up recruitment, randomization, and analyses. As with longitudinal designs, master protocol designs are also best suited for outcomes assessed within a short time frame of therapy, so that interim analyses and decisions rules can be updated frequently enough to be useful. Platform trials also are suited for therapeutic areas where multiple products are developed in a specific indication. Similarly, Basket trials are suited for those indications where the therapy and the disorder biological mechanism are well understood.

Adaptive designs

Adaptive designs are the type of designs that allow alterations to trial procedures and/or statistical procedures of an on-going trial after its initiation, without undermining the validity and integrity of the intended study. Such designs provide flexibility and efficiency and thus could be very useful for studies of rare diseases (60). While there are many different types of adaptive designs, the common element is to perform interim analysis to make decisions on adjustment of trial procedure(s) so that the overall trial efficiency could be improved. However, the specific decisions and adjustments need to be pre-planned and clearly described at the design stage. Those include when the interim analysis will occur, what kind of observed data will trigger a change, what aspects of the trial will change, and the impact of these modifications on statistical power and type 1 error.

One type of adaptive design is through use of adaptive randomization, to allow adjustment of randomization procedure based on accumulated information during the trial. In contrast to a traditional design that randomizes patients to treatment arms with equal and fixed probabilities, adaptive randomization could change the randomization probability to be treatment-adaptive, covariate-adaptive, or response-adaptive. Specifically, response adaptive randomization (RAR) modifies the randomization probability to allocate more future patients to the drug that are empirically superior to others, based on the observed outcome data from the patients already enrolled in the trial. It improves the collective outcomes of trial participants and increases the available information for the most effective treatment(s) (61,62). RAR has been used in various contexts and is also recently used in platform trials (44,48). The feature of RAR that adaptively assign less patients to the inefficacious treatment(s) is naturally aligned with the objective of platform trials to adaptively drop inefficacious agent(s) from the study platform. Limitations of RAR include the requirement of the response being evaluable in a short time, potential non-concurrency of the controls due to the shifting of randomization probabilities, and in some cases the issue of lower power than the standard 1:1 randomization design.

Another type of adaptive design is adaptive dose-finding (phase I) trials in early phase drug development to identify maximum tolerable dose (MTD) of the new drug. The traditional 3+3 dose-finding design enrolls three patients to a dose cohort and determine the dose for the next cohort according to a pre-specified algorithm. Despite its simplicity, this design may lead to wrong selection of MTD, require a large number of patients and take a long time to complete, especially when many doses are being explored (63). In rare diseases such as pediatric cancer, it could take several years to finish one dose-finding trial. Many adaptive designs have been proposed to improve the efficiency of phase I trials, such as continuous reassessment method (CRM) (64), Bayesian optimal interval (BOIN) design (65) and modified toxicity probability interval (mTPI) design (66). For example, CRM assumes a statistical model for the dose-toxicity curve and uses all the observed toxicity data across dose cohorts to update the model parameter and the estimate of MTD. In addition, traditional phase I trials only focus on the toxicity outcome and wait until phase II trials to evaluate the treatment response. However, a better strategy could be characterizing patient outcome in terms of both toxicity and response, in a hybrid phase I/II design. This produces both dose-finding rules and early stopping rules with respect to toxicity and response, combines elements of typical phase I and phase II designs, and therefore could shorten the timeline of drug development. We refer to Thall 2004 (67), Yin 2006 (68), Houede 2010 (69), and Guo 2015 (70) for examples of hybrid phase I/II designs.

Other examples of adaptive designs include group sequential designs that allow for stopping a trial early due to futility and/or efficacy (71,72), sample size re-estimation at interim analysis to achieve the desired statistical power (73,74), seamless phase II/III design that consists of a learning stage (phase II) and a confirmatory stage (phase III) with an interim analysis following the learning stage to make a selection decision (on treatment arm or study population) (75), and small n sequential multiple assignment randomized trial (snSMART) design where individuals are randomized to a set of treatment options and may be re-randomized at an interim timepoint in the trial (76,77). Lastly, a design could be multiple-adaptive, combining the elements of various adaptive designs. For example, Li *et al.* proposed a platform trial design combining the concepts of hybrid phase I/II trial and the RAR for treatment and dose prioritization in the context of platform trials (78). Although careful planning (both logistical and statistical) is required and implementation challenges exist, appropriate use of

Table 2 Examples of combinations of non-traditional designs and analytical strategies

Design	Causal inference	Borrowing evidence	Adaptive analytical strategy
External control	Propensity score methods (29) or exact matching (20)	Meta-analytic-predictive approach (29,79,80) Bayesian methods (81)	
Longitudinal design randomized		Hierarchical models (34)	
Longitudinal design observational	Propensity score methods (10,82,83)		
Master protocol platform			BATTLE (48)
Master protocol basket		NCI MATCH (53)	BRAF-vemurafenib trial (54)

adaptive designs has great potential to reduce costs and increase efficiency in clinical development of rare diseases.

Examples of novel analytical strategies

The analytical strategies that we discuss in this section are often used with the design strategies discussed in the earlier section. The overlap is illustrated in *Table 2*.

Causal inference and pharmacoepidemiology method

In the absence of randomization, causal inference methods can control for biases and result in valid inference when comparing a new therapy to a control group (84,85). These methods are playing an increasingly important role as real-world data is incorporated in evaluation of benefit-risk evidence throughout the therapeutic development process (86,87). They are particularly relevant in the rare disease setting to evaluate comparative safety or effectiveness in an epidemiological study. They are also relevant when using an external or historical control to contextualize evidence of a single arm study.

Conceptually, causal inference estimates the treatment effect by comparing a subject's observed outcome under treatment received to the counterfactual outcome if the subject had received an alternate treatment. Potential threats to internal validity of this comparison include selection bias, outcome ascertainment bias, and confounding. These result in lack of comparability of the two treatment groups in observed baseline characteristics that impact the outcome independently of the treatment.

A useful framework in designing and analyzing observational study to control or minimize some of these biases is that of the target trial framework (88). Under this framework, one designs and analyze an epidemiological study to mimic the design and analysis of a randomized trial, if that trial were possible. Thus, by putting the emphasis on incident cohort designs, similarity of inclusion and exclusion criteria, and suitable choice of index date, one can control or minimize multiple biases. Additional confounding can be further controlled by matching or weighting such as use of propensity score matching or inverse probability treatment weighting (89). Several examples exist of these method's use in rare disorders. For example, the cerliponase alfa pivotal study (20) mentioned above used exact matching as a secondary analysis to compare clinical trial subjects to historical control subjects, then a logistic regression outcome model estimated the treatment effect. Additionally, the study evaluating the effectiveness of warfarin in pulmonary arterial hypertension used propensity score matching to minimize confounding in the analytical cohort before using Bayesian methods to measure treatment effect (82). Also, the study evaluating effectiveness of aggressive corticosteroid therapy versus standard therapy in treating juvenile dermatomyositis used propensity score matching to reduce confounding by indication (83).

Causal inference methods can mitigate some observational study biases by design or analysis. In rare disorders, these methods were especially successful in mitigating bias in observational data when the treatment effect size is large. Application of these methods may not be feasible when effect sizes and sample sizes are small. While the comparative performance of different causal inference methods is well understood in studies of at least a few hundred patients, for example using simulation results from Austin under different settings (90-93), to our knowledge, no simulation studies explored their feasibility and performance of causal inference with small samples.

Current best practices in pharmacoepidemiology recommend for pre-specification of design and analysis methods prior to "unblinding" of the potential association between treatment and outcome (94). This poses a unique challenge in rare disease because while one may want to learn about natural history under the control arm before refining the endpoints, this may introduce investigator bias in using this same natural history study in analysis as an external control.

Borrowing evidence

Borrowing evidence is an important analytical strategy to overcome the issue of limited sample size in rare disease setting and could occur in different ways depending on the context. For a basket trial (discussed in Section 'Master protocols') that enrolls patients with different disease types but the same molecular alteration, information borrowing mainly occurs across the disease types and within the given trial. For a trial designed with external controls (discussed in Section 'External or historical controls'), information borrowing occurs between the external control data and current trial data. For pediatric trials that are typically initiated after some data have been accumulated in the adult population, the natural information borrowing is from adult to pediatric, and there is an increasing interest recently to use innovative method to integrate evidence from the adult to pediatric decision making (95). Data sources for borrowing could also vary, with either summary-level data or patient-level data, from either clinical trials or observational databases (EHR/claims), and could be either historical or contemporary. Moreover, justifications for information borrowing should be assessed, such as if the study populations from various data sources share any similarities.

To facilitate information borrowing, Bayesian analysis framework is typically used, due to its ability of incorporating a data structure with multiple levels of hierarchy and accounting for external (to the current trial) information through the use of priors. For information borrowing across multiple cohorts with a given trial (e.g., in a basket trial), hierarchical Bayesian modeling (HBM) or its modifications are the most common approaches, as discussed in Section 'Master protocols'. For information borrowing between trials, informative priors of various forms could be used. The simplest informative prior is elicited priors, which elicit and summarize the experts' opinions into a prior distribution for the model parameter (96). A more complicated approach is power prior, where the historical data is used to calculate the posterior distribution of the model parameter and the resulting posterior becomes the new prior for the current trial (97). A power parameter was included in this approach to control how much information to borrow from the historical data, and the value of this parameter could be determined based on expert opinion or some model fit criteria (97,98). Another method is the meta-analytic-predictive (MAP) approach, which summarizes the historical information to make a prediction for the distribution of the model parameter in a new trial, and this predictive distribution serves as the informative prior for the current trial (79). The information contained in this informative prior distribution could also be expressed in terms of number of patients, as prior effective sample size, which provides an intuitive summary on the amount of information borrowed from the historical data. Moreover, for between trial information borrowing, although historical data is the most common type, concurrent trial data may be available in some cases and can be incorporated as well; see an example design that integrates a concurrent adult phase I trial data into the design of the pediatric phase I trial (99).

One key assumption in the above methods is exchangeability (or similarity) among multiple cohorts or trials, while this assumption may not always hold. Therefore, extensions to these methods have been proposed, to allow adaptive adjustment on the degree of borrowing according to the level of evidence on similarities between cohorts/trials accumulated during the current trial. Such methods allow more borrowing when there is strong evidence of similarity and reduces the amount of borrowing as the evidence decreases, and sometimes are referred to as 'dynamic' (as opposed to 'static') borrowing methods. For example for basket trials, the BLAST design allows the treatment response vary across disease cohorts (beyond the randomness defined by HBM) by assuming two underlying latent subgroups ('sensitive' *vs.* 'insensitive' to treatment), to achieve strong information borrowing within each latent subgroup but little borrowing between the two subgroups (58). For external evidence borrowing methods, extensions of power prior [such as calibrated power prior (100), joint power prior (101), and commensurate power prior (102)] and extensions of MAP [such as robust-MAP (80)] were developed to accommodate potential differences between the external data and the current trial data.

Other analytical considerations

For randomized clinical trials (RCTs), the most common approach to inference on the treatment effect is the likelihood-based inference, such as a two-sample *t*-test for normal outcome data and chi-square test for binary outcome data. Randomization-based inference (RBI) is an alternative approach, where the implanted randomization forms the basis for statistical inference (103). In this approach, treatment assignments are permuted in all possible ways that are consistent with the randomization procedure used in the trial, and a P value is calculated based on these permutations. RBI is robust against potential biases that may be caused by time trends (104,105), which is one of the issues in rare disease trials that take long times to complete. RBI could also yield valid tests even when the sample size is very small (103), so is particularly useful for rare disease studies.

Moreover, adaptive analytical strategy needs to be utilized for all the adaptive designs we described in Section 'Adaptive designs'. Specifics of the analytical strategies may vary but should match with the features of the design. For example, for the designs with interim decisions to potentially graduate an efficacious treatment early, the interim analysis should provide criteria to facility such a change and the final analysis should be adjusted to reflect this change (e.g., a revised significance level other than the standard 0.05).

Assessing the sample size needed to power a given study is an important feasibility consideration. In complex parallel-controlled randomized designs such as master protocol designs or with complex analytical strategies borrowing evidence, simulations can provide a flexible way for evaluating performance of design and analytical methods for different operating characteristics.

While we focused this paper on statistical considerations for study design and analysis, we acknowledge there are other strategies such as pharmacometrics approaches. As described by Ryeznik *et al.* (106), pharmacometrics models allow extrapolations of relevant information from one population to another (such as from adult to pediatric) and may be attractive from various perspectives including making the trial more ethical and reducing development costs.

Conclusions

In this paper we have focused on reviewing some novel designs and analytic methods to tackle challenges in clinical trials for rare diseases. We refer to other review papers for general overview of issues in developing and conducting trials for rare disease (3,31,32,107). For the approaches we described in this paper, some have already been used in rare disease research (such as leveraging longitudinal measurements), some are starting to be applied in certain type of rare diseases (such as basket trials in pediatric oncology and use of external controls in non-oncology rare disorders), and some have been rarely used in rare disease research but has great potentials and should be advocated (such as use of longitudinal designs and causal inference analyses in natural history studies).

It is interesting to note that although these novel approaches appear to have unique characteristics, they also share a few common themes and principals to increase efficiency and feasibility. First, they all attempt to augment sample size and increase statistical power, by either borrowing the existing information from external data sources, or leveraging the longitudinal information collected on the same patient. Second, they aim to use the limited resources more efficiently, such as sharing the control data in platform trials and using the same patient for testing multiple treatments in studies with self-control. Third, flexibilities are built into these designs/methods, such as allowing additions of experimental agents over time in platform trials and modifying trial procedures as more information becomes available in adaptive designs. Lastly, these approaches do require more upfront planning, appropriate infrastructures, availabilities of computing software, and effective communications between study personnel (physicians, statisticians, trial coordinators, etc.). We hope this paper will raise awareness of these novel approaches and encourage their use in studies of rare diseases.

Acknowledgments

Funding: None.

Footnote

Provenance and Peer Review: This article was commissioned by the Guest Editors [Yanhong Deng, Qian Shi and Jun (Vivien) Yin] for the series "Challenges in Clinical Trials" published in *Annals of Translational Medicine*. The article has undergone external peer review.

Conflicts of Interest: Both authors have completed the ICMJE uniform disclosure form (available at https://atm.amegroups.com/article/view/10.21037/atm-21-5496/coif). The series "Challenges in Clinical Trials" was commissioned by the editorial office without any funding or sponsorship. RI is an employee of Novartis and receive salary and benefits from this company since February 2021. The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Nguengang Wakap S, Lambert DM, Olry A, et al. Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. Eur J Hum Genet 2020;28:165-73.
2. "Prevalence of rare diseases: Bibliographic data", Orphanet Report Series, Rare Diseases collection, January 2021, Number 2 : Diseases listed by decreasing prevalence, incidence or number of published cases. Available online: http://www.orpha.net/orphacom/cahiers/docs/GB/Prevalence_of_rare_diseases_by_decreasing_prevalence_or_cases.pdf
3. Kempf L, Goldsmith JC, Temple R. Challenges of developing and conducting clinical trials in rare disorders. Am J Med Genet A 2018;176:773-83.
4. Hilgers RD, König F, Molenberghs G, et al. Design and analysis of clinical trials for small rare disease populations. Journal of Rare Diseases Research & Treatment 2016;1:53-60.
5. Hilgers RD, Bogdan M, Burman CF, et al. Lessons learned from IDeAl - 33 recommendations from the IDeAl-net about design and analysis of small population clinical trials. Orphanet J Rare Dis 2018;13:77.
6. Abrahamyan L, Diamond IR, Johnson SR, et al. A new toolkit for conducting clinical trials in rare disorders. J Popul Ther Clin Pharmacol 2014;21:e66-78.
7. The Food and Drug Administration. Rare Diseases: Natural History Studies for Drug Development, Guidance Document. 2019. Available online: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/rare-diseases-natural-history-studies-drug-development
8. Cystic Fibrosis Foundation Patient Registry. Annual Data Report 2018. Bethesda, MD, 2019.
9. Dasenbrook EC, Sawicki GS. Cystic fibrosis patient registries: A valuable source for clinical research. J Cyst Fibros 2018;17:433-40.
10. Ah Mew N, McCarter R, Izem R, et al. Comparing Treatment Options for Urea Cycle Disorders. Patient Centered Outcome Research: Patient Centered Outcome Research, 2020. Available online: https://www.pcori.org/research-results/2015/comparing-treatment-options-urea-cycle-disorders
11. Krischer JP, Gopal-Srivastava R, Groft SC, et al. The Rare Diseases Clinical Research Network's organization and approach to observational research and health outcomes research. J Gen Intern Med 2014;29 Suppl 3:S739-44.
12. Putkowski S. National Organization for Rare Disorders (NORD): providing advocacy for people with rare disorders. NASN Sch Nurse 2010;25:38-41.
13. Canterberry M, Kaul AF, Goel S, et al. The Patient-Centered Outcomes Research Network Antibiotics and Childhood Growth Study: Implementing Patient Data Linkage. Popul Health Manag 2020;23:438-44.
14. The US Food and Drug Administration. Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims, Guidance for Industry. 2009.
15. The US Food and Drug Administration. Biomarker qualification evidentiary framework guidance for industry. 2018. Available online: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/biomarker-qualification-evidentiary-framework
16. Administration. TUSFAD. Digital Health Technologies for Remote Data Acquisition in Clinical Investigations. Available online: https://www.fda.gov/media/155022/download
17. Maguire AM, Russell S, Wellman JA, et al. Efficacy, Safety, and Durability of Voretigene Neparvovec-rzyl in RPE65 Mutation-Associated Inherited Retinal Dystrophy: Results of Phase 1 and 3 Trials. Ophthalmology 2019;126:1273-85.
18. Chung DC, McCague S, Yu ZF, et al. Novel mobility test to assess functional vision in patients with inherited retinal dystrophies. Clin Exp Ophthalmol 2018;46:247-59.
19. Mantua V, Arango C, Balabanov P, et al. Digital health technologies in clinical trials for central nervous system drugs: an EU regulatory perspective. Nat Rev Drug Discov 2021;20:83-4.
20. Schulz A, Ajayi T, Specchio N, et al. Study of Intraventricular Cerliponase Alfa for CLN2 Disease. N Engl J Med 2018;378:1898-907.

21. The US Food and Drug Administration. Multiple Endpoints in Clinical Trials Guidance for Industry. 2017. Available online: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/multiple-endpoints-clinical-trials-guidance-industry
22. Finkelstein DM, Schoenfeld DA. Combining mortality and longitudinal measures in clinical trials. Stat Med 1999;18:1341-54.
23. Li Y, Mick R, Heitjan DF. Suspension of accrual in phase II cancer clinical trials. Clin Trials 2015;12:128-38.
24. Cheung YK, Chappell R. Sequential designs for phase I clinical trials with late-onset toxicities. Biometrics 2000;56:1177-82.
25. Yuan Y, Lin R, Li D, et al. Time-to-Event Bayesian Optimal Interval Design to Accelerate Phase I Trials. Clin Cancer Res 2018;24:4921-30.
26. Lin R, Coleman RL, Yuan Y. TOP: Time-to-Event Bayesian Optimal Phase II Trial Design for Cancer Immunotherapy. J Natl Cancer Inst 2020;112:38-45.
27. Seeger JD, Davis KJ, Iannacone MR, et al. Methods for external control groups for single arm trials or long-term uncontrolled extensions to randomized clinical trials. Pharmacoepidemiol Drug Saf 2020;29:1382-92.
28. Lim J, Walley R, Yuan J, et al. Minimizing Patient Burden Through the Use of Historical Subject-Level Data in Innovative Confirmatory Clinical Trials: Review of Methods and Opportunities. Ther Innov Regul Sci 2018;52:546-59.
29. Schmidli H, Häring DA, Thomas M, et al. Beyond Randomized Clinical Trials: Use of External Controls. Clin Pharmacol Ther 2020;107:806-16.
30. Schulz MC, Korn P, Stadlinger B, et al. Coating with artificial matrices from collagen and sulfated hyaluronan influences the osseointegration of dental implants. J Mater Sci Mater Med 2014;25:247-58.
31. Cornu C, Kassai B, Fisch R, et al. Experimental designs for small randomised clinical trials: an algorithm for choice. Orphanet J Rare Dis 2013;8:48.
32. Gagne JJ, Thompson L, O'Keefe K, et al. Innovative research methods for studying treatments for rare diseases: methodological review. BMJ 2014;349:g6802.
33. Harding CO, Amato RS, Stuy M, et al. Pegvaliase for the treatment of phenylketonuria: A pivotal, double-blind randomized discontinuation Phase 3 clinical trial. Mol Genet Metab 2018;124:20-6.
34. Ah Mew N, Cnaan A, McCarter R, et al. Conducting an investigator-initiated randomized double-blinded intervention trial in acute decompensation of inborn errors of metabolism: Lessons from the N-Carbamylglutamate Consortium. Transl Sci Rare Dis 2018;3:157-70.
35. Izem R, McCarter R. Randomized and non-randomized designs for causal inference with longitudinal data in rare disorders. Orphanet J Rare Dis 2021;16:491.
36. Whitaker HJ, Farrington CP, Spiessens B, et al. Tutorial in biostatistics: the self-controlled case series method. Stat Med 2006;25:1768-97.
37. Maclure M. The case-crossover design: a method for studying transient effects on the risk of acute events. Am J Epidemiol 1991;133:144-53.
38. Dharmarajan S, Lee JY, Izem R. Sample size estimation for case-crossover studies. Stat Med 2019;38:956-68.
39. Mansournia MA, Etminan M, Danaei G, et al. Handling time varying confounding in observational research. BMJ 2017;359:j4587.
40. Woodcock J, LaVange LM. Master Protocols to Study Multiple Therapies, Multiple Diseases, or Both. N Engl J Med 2017;377:62-70.
41. Renfro LA, Sargent DJ. Statistical controversies in clinical research: basket trials, umbrella trials, and other master protocols: a review and examples. Ann Oncol 2017;28:34-43.
42. Yee LM, McShane LM, Freidlin B, et al. Biostatistical and Logistical Considerations in the Development of Basket and Umbrella Clinical Trials. Cancer J 2019;25:254-63.
43. Hobbs BP, Chen N, Lee JJ. Controlled multi-arm platform design using predictive probability. Stat Methods Med Res 2018;27:65-78.
44. Yuan Y, Guo B, Munsell M, et al. MIDAS: a practical Bayesian design for platform trials with molecularly targeted agents. Stat Med 2016;35:3892-906.
45. Berry SM, Connor JT, Lewis RJ. The platform trial: an efficient strategy for evaluating multiple treatments. JAMA 2015;313:1619-20.
46. Saville BR, Berry SM. Efficiencies of platform clinical trials: A vision of the future. Clin Trials 2016;13:358-66.
47. Barker AD, Sigman CC, Kelloff GJ, et al. I-SPY 2: an adaptive breast cancer trial design in the setting of neoadjuvant chemotherapy. Clin Pharmacol Ther 2009;86:97-100.
48. Zhou X, Liu S, Kim ES, et al. Bayesian adaptive design for targeted therapy development in lung cancer--a step toward personalized medicine. Clin Trials 2008;5:181-93.
49. Steuer CE, Papadimitrakopoulou V, Herbst RS, et al. Innovative Clinical Trials: The LUNG-MAP Study. Clin Pharmacol Ther 2015;97:488-91.
50. Bateman RJ, Benzinger TL, Berry S, et al. The DIAN-TU Next Generation Alzheimer's prevention trial: Adaptive

design and disease progression model. Alzheimers Dement 2017;13:8-19.
51. Innovative Trial for Understanding the Impact of Targeted Therapies in NF2 (INTUITT-NF2). Available online: https://clinicaltrials.gov/ct2/show/NCT04374305
52. Redig AJ, Jänne PA. Basket trials and the evolution of clinical trial design in an era of genomic medicine. J Clin Oncol 2015;33:975-7.
53. Conley BA, Doroshow JH. Molecular analysis for therapy choice: NCI MATCH. Semin Oncol 2014;41:297-9.
54. Hyman DM, Puzanov I, Subbiah V, et al. Vemurafenib in Multiple Nonmelanoma Cancers with BRAF V600 Mutations. N Engl J Med 2015;373:726-36.
55. Patel SP, Mayerson E, Chae YK, et al. A phase II basket trial of Dual Anti-CTLA-4 and Anti-PD-1 Blockade in Rare Tumors (DART) SWOG S1609: High-grade neuroendocrine neoplasm cohort. Cancer 2021;127:3194-201.
56. Berry SM, Broglio KR, Groshen S, et al. Bayesian hierarchical modeling of patient subpopulations: efficient designs of Phase II oncology clinical trials. Clin Trials 2013;10:720-34.
57. Chu Y, Yuan Y. A Bayesian basket trial design using a calibrated Bayesian hierarchical model. Clin Trials 2018;15:149-58.
58. Chu Y, Yuan Y. BLAST: Bayesian latent subgroup design for basket trials. J R Stat Soc Ser C 2018;67:723-40.
59. Neuenschwander B, Wandel S, Roychoudhury S, et al. Robust exchangeability designs for early phase clinical trials with multiple strata. Pharm Stat 2016;15:123-34.
60. Chow SC, Chang M. Adaptive design methods in clinical trials - a review. Orphanet J Rare Dis 2008;3:11.
61. Thall PF, Wathen JK. Practical Bayesian adaptive randomisation in clinical trials. Eur J Cancer 2007;43:859-66.
62. Meurer WJ, Lewis RJ, Berry DA. Adaptive clinical trials: a partial remedy for the therapeutic misconception? JAMA 2012;307:2377-8.
63. Zhu Y, Hwang WT, Li Y. Evaluating the effects of design parameters on the performances of phase I trial designs. Contemp Clin Trials Commun 2019;15:100379.
64. O'Quigley J, Pepe M, Fisher L. Continual reassessment method: a practical design for phase 1 clinical trials in cancer. Biometrics 1990;46:33-48.
65. Yuan Y, Hess KR, Hilsenbeck SG, et al. Bayesian Optimal Interval Design: A Simple and Well-Performing Design for Phase I Oncology Trials. Clin Cancer Res 2016;22:4291-301.
66. Guo W, Wang SJ, Yang S, et al. A Bayesian interval dose-finding design addressingOckham's razor: mTPI-2. Contemp Clin Trials 2017;58:23-33.
67. Thall PF, Cook JD. Dose-finding based on efficacy-toxicity trade-offs. Biometrics 2004;60:684-93.
68. Yin G, Li Y, Ji Y. Bayesian dose-finding in phase I/II clinical trials using toxicity and efficacy odds ratios. Biometrics 2006;62:777-84.
69. Houede N, Thall PF, Nguyen H, et al. Utility-based optimization of combination therapy using ordinal toxicity and efficacy in phase I/II trials. Biometrics 2010;66:532-40.
70. Guo W, Ni Y, Ji Y. TEAMS: Toxicity- and Efficacy-based Dose Insertion Design with Adaptive Model Selection for Phase I/II Dose-Escalation Trials in Oncology. Stat Biosci 2015;7:432-59.
71. Demets DL. Group sequential procedures: calendar versus information time. Stat Med 1989;8:1191-8.
72. Jennison C, Turnbull BW. Group sequential methods with applications to clinical trials. 1st edition. London: Routledge, 2000.
73. Cui L, Hung HM, Wang SJ. Modification of sample size in group sequential clinical trials. Biometrics 1999;55:853-7.
74. Chuang-Stein C, Anderson K, Gallo P, et al. Sample Size Reestimation: A Review and Recommendations. Drug Inf J 2006;40:475-84.
75. Maca J, Bhattacharya S, Dragalin V, et al. Adaptive Seamless Phase II/III Designs—Background, Operational Aspects, and Examples. Drug Inf J 2006;40:463-73.
76. Tamura RN, Krischer JP, Pagnoux C, et al. A small n sequential multiple assignment randomized trial design for use in rare disease research. Contemp Clin Trials 2016;46:48-51.
77. Micheletti RG, Pagnoux C, Tamura RN, et al. Protocol for a randomized multicenter study for isolated skin vasculitis (ARAMIS) comparing the efficacy of three drugs: azathioprine, colchicine, and dapsone. Trials 2020;21:362.
78. Li Y, Wang M, Cheung YK. Treatment and dose prioritization in early phase platform trials of targeted cancer therapies. J R Stat Soc Ser C 2019;68:475-91.
79. Neuenschwander B, Capkun-Niggli G, Branson M, et al. Summarizing historical information on controls in clinical trials. Clin Trials 2010;7:5-18.
80. Schmidli H, Gsteiger S, Roychoudhury S, et al. Robust meta-analytic-predictive priors in clinical trials with historical control information. Biometrics 2014;70:1023-32.
81. Hampson LV, Whitehead J, Eleftheriou D, et al. Bayesian methods for the design and interpretation of clinical trials in very rare diseases. Stat Med 2014;33:4186-201.
82. Johnson SR, Granton JT, Tomlinson GA, et al.

Warfarin in systemic sclerosis-associated and idiopathic pulmonary arterial hypertension. A Bayesian approach to evaluating treatment for uncommon disease. J Rheumatol 2012;39:276-85.
83. Seshadri R, Feldman BM, Ilowite N, et al. The role of aggressive corticosteroid therapy in patients with juvenile dermatomyositis: a propensity score analysis. Arthritis Rheum 2008;59:989-95.
84. Imbens G, Rubin D. Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction. Cambridge: Cambridge University Press, 2015.
85. Hernan MA, Robins JM. Causal Inference: What If. Boca Raton: Chapman & Hall/CRC, 2020.
86. Sherman RE, Anderson SA, Dal Pan GJ, et al. Real-World Evidence - What Is It and What Can It Tell Us? N Engl J Med 2016;375:2293-7.
87. Berger ML, Sox H, Willke RJ, et al. Good practices for real-world data studies of treatment and/or comparative effectiveness: Recommendations from the joint ISPOR-ISPE Special Task Force on real-world evidence in health care decision making. Pharmacoepidemiol Drug Saf 2017;26:1033-9.
88. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. Am J Epidemiol 2016;183:758-64.
89. Stuart EA. Developing practical recommendations for the use of propensity scores: discussion of 'A critical appraisal of propensity score matching in the medical literature between 1996 and 2003' by Peter Austin, Statistics in Medicine. Stat Med 2008;27:2062-5; discussion 2066-9.
90. Austin PC. A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003. Stat Med 2008;27:2037-49.
91. Austin PC. The performance of different propensity-score methods for estimating differences in proportions (risk differences or absolute risk reductions) in observational studies. Stat Med 2010;29:2137-48.
92. Austin PC. The use of propensity score methods with survival or time-to-event outcomes: reporting measures of effect similar to those used in randomized experiments. Stat Med 2014;33:1242-58.
93. Austin PC, Stuart EA. The performance of inverse probability of treatment weighting and full matching on the propensity score in the presence of model misspecification when estimating the effect of treatment on survival outcomes. Stat Methods Med Res 2017;26:1654-70.
94. Wang SV, Schneeweiss S, Berger ML, et al. Reporting to Improve Reproducibility and Facilitate Validity Assessment for Healthcare Database Studies V1.0. Value Health 2017;20:1009-22.
95. Gamalo-Siebers M, Hampson L, Kordy K, et al. Incorporating Innovative Techniques Toward Extrapolation and Efficient Pediatric Drug Development. Ther Innov Regul Sci 2019;53:567-78.
96. O'Hagan A, Buck CE, Daneshkhah A, et al. Uncertain Judgements: Eliciting Experts' Probabilities. Hoboken, New Jersey: Wiley, 2006.
97. Ibrahim JG, Chen MH. Power Prior Distributions for Regression Models. Stat Sci 2000;15:46-60.
98. Ibrahim JG, Chen MH, Chu H. Bayesian methods in clinical trials: a Bayesian analysis of ECOG trials E1684 and E1690. BMC Med Res Methodol 2012;12:183.
99. Li Y, Yuan Y. PA-CRM: A continuous reassessment method for pediatric phase I oncology trials with concurrent adult trials. Biometrics 2020;76:1364-73.
100. Pan H, Yuan Y, Xia J. A Calibrated Power Prior Approach to Borrow Information from Historical Data with Application to Biosimilar Clinical Trials. J R Stat Soc Ser C Appl Stat 2017;66:979-96.
101. Ibrahim JG, Chen MH, Gwon Y, et al. The power prior: theory and applications. Stat Med 2015;34:3724-49.
102. Hobbs BP, Carlin BP, Mandrekar SJ, et al. Hierarchical commensurate and power prior models for adaptive incorporation of historical information in clinical trials. Biometrics 2011;67:1047-56.
103. Berger VW, Bour LJ, Carter K, et al. A roadmap to using randomization in clinical trials. BMC Med Res Methodol 2021;21:168.
104. Rosenberger WF, Uschner D, Wang Y. Randomization: The forgotten component of the randomized clinical trial. Stat Med 2019;38:1-12.
105. Proschan MA, Dodd LE. Re-randomization tests in clinical trials. Stat Med 2019;38:2292-302.
106. Ryeznik Y, Sverdlov O, Svensson EM, et al. Pharmacometrics meets statistics-A synergy for modern drug development. CPT Pharmacometrics Syst Pharmacol 2021;10:1134-49.
107. Groft SC, Posada M, Taruscio D. Progress, challenges and global approaches to rare diseases. Acta Paediatr 2021;110:2711-6.

Cite this article as: Li Y, Izem R. Novel clinical trial design and analytic methods to tackle challenges in therapeutic development in rare diseases. Ann Transl Med 2022;10(18):1034. doi: 10.21037/atm-21-5496