Planning and Monitoring Equitable Clinical Trial Enrollment Using Goal Programming

Miao Qi ⦿, Amar K. Das, and Kristin P. Bennett ⦿

Abstract—Randomized clinical trial (RCT) studies are the gold standard for scientific evidence on treatment benefits to patients. RCT outcomes may not be generalizable to clinical practice if the trial population is not representative of the patients for which the treatment is intended. Specifically, enrollment plans may not adequately include groups of patients with protected attributes, such as gender, race, or ethnicity. Inequities in RCTs are a major concern for funding agencies such as the National Institutes of Health (NIH) and for policy makers. We address this challenge by proposing a goal-programming approach, explicitly integrating measurable enrollment goals, to design equitable enrollment plans for RCTs. We evaluate our model in both single and multisite settings using the enrollment criteria and study population from the Systolic Blood Pressure Intervention Trial (SPRINT) study. Our model can successfully generate equitable enrollment plans that satisfy multiple goals such as sample representativeness and minimum total financial cost. Our model can detect deviations from a target plan during the enrollment process and update the plan to reduce deviations in the remaining process. Finally, through appropriate site selection in the planning stage, the model can demonstrate the possibility of enrolling a nationally representative study population if geographic constraints exist in multisite recruitment (e.g., clinical centers in a particular region). Our model can be used to prospectively produce and retrospectively evaluate how equitable enrollment plans are based on subjects' protected attributes, and it allows researchers to provide justifications on validity of scientific analysis and evaluation of subgroup disparities.

Index Terms—Biomedical informatics, computer aided analysis, data analysis, health equity, integer linear programming, optimization, public healthcare, randomized clinical trials, enrollment.

I. Introduction

EQUITABLE enrollment of people from diverse backgrounds for randomized clinical trials (RCTs) has been identified as a high priority to build a healthier nation [1], [2], [3]. RCTs are believed to bring significant public health benefits. However, these benefits may not be established or recognized when enrolled participants are insufficient or inappropriate to the scientific question under study and fail to represent broader populations as planned. Therefore, representation assessment should be part of enrollment planning and monitoring to ensure that the knowledge gained from research can be generalized to all individuals who have the disease or health condition that is the focus of the trial.

However, few trial planning or trial monitoring tools support representation evaluation by investigators, Institutional Review Boards (IRBs), or Data and Safety Monitoring Boards (DSMBs). If inequitable representations of one or more disadvantaged populations exist, modification of design can be recommended and the enrollment plan can be updated dynamically through the recruitment process. Additionally, for multisite studies, such decisions can determine site selection given each site's heterogeneity among the pool of potential study subjects.

To promote equitable RCTs, National Institutes of Health's (NIH) inclusion policy requires investigators to propose and justify an expected distribution of study participants, by sex/gender, race, ethnicity, and age, that reflects the broader population to accomplish the study goal (i.e., Planned Enrollment Report) [4]. For a study including an existing cohort/dataset, the Cumulative Inclusion Enrollment Report is used and justified by investigators. Additionally, NIH and IRB evaluate the plan on whether it will yield valid analyses including analysis of potential subgroup differences [5], [6], [7], [8]. For trial designers, it is challenging to create such a detailed forecast on planned subgroup enrollment regarding to the target population due to the lack of a representation evaluation layer on current enrollment planning strategies [9], [10], [11].

Also, equitable clinical trial enrollments are often complicated by many factors including poor site selection and inappropriate or non-optimal recruitment planning/monitoring. For example, given a set of available trial sites, how should we efficiently identify the sites to ensure access to an adequate number of diverse participants who meet the study requirements? What are the overall and site-level expectations in planning to enroll the right and enough subjects for the study?

Additionally, goal monitoring of enrollment targets during the stochastic recruitment process is a challenge especially for multicenter RCTs. For instance, given the interim recruitment data at a time point, how should we evaluate disparities between planned and actual numbers of patients and how should we

adjust the recruitment plan going forward to reduce any detected disparities [12], [13]?

To fulfill the NIH Inclusion Policy and address the concerns of enrollment process mentioned before, we propose a multi-objective goal-programming (GP) model for equitable enrollment planning and monitoring based on our previous work about RCT representativeness metrics [14].

Guided by the scientific aims of the study, our GP model explicitly models equitable representation as a goal in enrollment and allows researchers to set goal preferences/priorities that best fit the applied scenario and learn from intermediate analysis to obtain a more satisfying solution eventually. Enrollments goals for subgroups defined on the basis of sex/gender, race/ethnicity, age, or other desired characteristics are defined using surveillance datasets. Thus, the GP modeling process provides justification in terms of a representative sample required by NIH and IRB evaluation to perform valid analysis. It also incorporates lower bounds on subgroup sizes to ensure that the study has sufficient power to determine subgroup differences.

The GP model is easily understood by users once targets of different goals are set and deviations to targets in different directions are not necessary weighted the same. Goals can go beyond equitable representation; we include goals to improve the RCT effectiveness and efficiency of RCT recruiting in the GP model. Furthermore, previous knowledge and experience can be an input of the model to generate better estimations in planning. Finally, GP guarantees a Pareto-optimal solution and is computationally efficient, especially when multiple conflicting goals exist. This means that no solutions that improve lower-priority goal achievements will degrade the high-priority goal achievements. Due to its flexibility, our model could be a complementary part which could be added into other existing enrollment models.

A. Enrollment Planning and Monitoring

To understand the proposed approach, we first introduce some terms from population representativeness.

- **Target population**: The entire group of individuals potentially affected by the researched diseases or conditions. It can be local, regional, national, or global based on study goals and other trial-specific conditions such as eligibility criteria.
- **Protected attributes**: Any baseline subject attributes that can classify the target population into different subgroups with desired parity in terms of health outcomes received. Protected attributes can include demographic, clinical, laboratory, or risk factors.
- **Subgroups**: Subsets of target population that share common subject attribute values and thus can be distinguished from the rest.
- **Representativeness metrics**: Quantitative measures for disparities between a target population and an observed sample [14].

Ideally, subgroup sizes are equal to their proportions in target population. These expected number of subgroup participants

defined over a set of protected attributes are called enrollment targets, which are calculated based on trial size and target proportion. Target populations are usually estimated from another "benchmark" dataset such as National Health and Nutrition Examination Survey (NHANES) [15] or from electronic health records (EHRs).

Contrasting with our representativeness focus, existing enrollment planning and monitoring models for clinical research focus on the prediction and evaluation of enrollment feasibility and duration to meet a target sample size. Different types of models including real-time and simulation-based prediction methods are used to estimate the real-world recruitment process and create improved predictions. For example, groups of researchers [16], [17] reviewed commonly-used models to predict accrual subject enrollment and event times and suggested to use flexible stochastic models and center-specific information in multisite studies. For multicenter studies, some enrollment models [18], [19], [20], [21] add more variables to account for heterogeneity in recruitment centers, such as various center size and temporal change in enrollment, and apply more advanced techniques, such as inhomogeneous Poisson process [22], metaheuristics [23], and Monte Carlo simulation [24], to offer improved accuracy in the prediction. Time- and capacity-relevant considerations are carefully modeled while the patient heterogeneity among centers are not present. Additionally, the monitoring committees and review boards pay more attention on patient safety and efficacy, and interim analyses are performed to mainly ensure data integrity and remove potential data errors through detecting missing, invalid data and unusual data patterns through statistical algorithms [25], [26], instead of monitoring the population representation.

In our study, we use representativeness metrics with statistical methods, embedding in a goal-programming-based optimization model, to decide a priori target enrollment range, which is a more realistic goal in complex clinical settings compared to a single number, for each subgroup in the trial planning, and monitor and mitigate the deviation of actual enrollment from these target ranges through the process.

B. Clinical Trial Site Selection

For multisite trials, we expand the single-site model to solve site selection task with multiple conflicting goals. The disconnection between clinical center selection and distribution of target subgroups makes the accrual of representative study population hard to be achieved, especially for racial and ethnic communities who are highly impacted by the geographical factors of trial recruitment [27].

Most existing approaches of site selection depend on the site performance, which includes estimated enrollment rate, site experience such as facility quality and historical enrollment rate, and other site-specific traits, to make a weighted score that helps ranking the sites that best fit the study [28], [29], [30]. The other strategies recommend sites based on feasibility assessments to fulfil the overall recruitment goal [31], [32]. For this work, we focus on DEI (diversity, equity, and inclusion)-considerations,
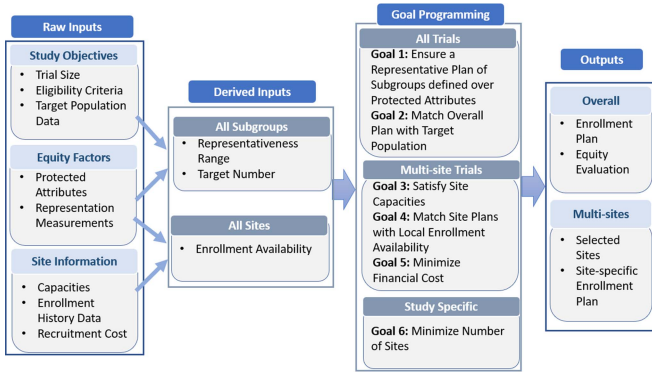
Fig. 1.   Equitable Clinical Trial Enrollment Model Framework.

and we observe that these other important practical site considerations could be incorporated as additional goals in the GP model as future work.

Our multisite model integrates patient heterogeneity of different disease, per-site population distributions, and other site-specific information used in existing models, such as patient recruitment cost and site capacity, into model objectives and constraints to design equitable enrollment plans, including the overall and site-level ones. Our models can be additionally applied with the existing predictive modeling techniques which are better for event counts and time estimation to generate clinically useful enrollment plans.

C. Contributions

Fig. 1 describes the framework of our proposed model. The data inputs provided by users include three different types. First, the inputs related to trial design are the planning trial size, study eligibility criteria, and target population data from surveillance datasets (e.g., NHANES) or electronic health records. Second, to evaluate the equitable representation of an RCT, protected attributes and representativeness measurements should be defined. Third, for multisites RCTs, site-level information such as site capacities, enrollment history data, and recruitment cost are needed. The first two types of raw inputs will be processed to derive a target enrollment size and an acceptable enrollment range based on the statistical tests and representation measurements for each subgroup defined over the protected attributes. When more than one site will be used, the last two types of inputs are processed to generate enrollment availability of protected subgroups from each site.

With sufficient data inputs, our model is able to achieve three different groups of goals. First, for any trial, the model ensures optimized equity in subgroup representation. Goal 1 ensures a representative plan across all subgroups of interest. Goal 2 additionally matches the overall plan with the target population. For multisite planning, example goals include satisfying site capacities (i.e., Goal 3) and matching site plans with local available population compositions (i.e., Goal 4). Additional study-specific goals such as Goal 5 which minimizes overall enrollment cost and Goal 6 which minimizes site numbers could also be considered. The model output will be an enrollment plan

with equity evaluation that balances a trade-off in conflicting goals. Site selection result and specific planning for each selected site are available for multisite trials.

This paper introduces a novel and flexible multi-objective GP approach, explicitly integrating measurable enrollment goals with other (conflicting) recruitment goals, to design equitable enrollment plans and to monitor recruitment process for single- and multi-site RCTs. We first describe two GP models, one for single site RCTs and one for multisite RCTs based on single-site model with additional site-level goals. Then we present experimental settings and results of four use cases of the approach: (1) design equitable enrollment for new single-site RCTs, (2) provide remedial re-planning if any inequitable representation is identified during interim analysis of single-site RCTs, (3) recruit a national representative sample with optimized site selection under area constraints, and (4) make an equitable plan for multi-site RCTs. Finally, we address the advantages and limitations of the approach with future works.

II. METHODS

We construct a multi-objective model for RCTs using GP. The model regards pre-defined target population data as benchmark to calculate a target plan.

As shown in Fig. 1, this weighted GP model sets numeric goals on representativeness of *all* subgroups for both the overall enrollment planning/monitoring and each site if applicable. Other goals with different priority levels are added as needed. The model generates a feasible solution composed of an optimal combination of enrollments within each selected recruitment site that maximally contribute to the balance between representativeness and other (conflicting) goals such as recruitment costs.

The primary goal of the model is to minimize the total weighted sum of subgroup deviations from their target/planned enrollment, where the penalty weights for underachievement and overachievement of the desired goals can be set by investigators based on study design. The weights allow deviations from each goal to have different importance.

The mathematical models for enrollment planning and monitoring are described in the following subsections. The notations of models are introduced in Table I.

In our models, we optimize enrollment for all possible subgroups $\hat{S}$ defined over the projected attributes. In this work, we use two protected attributes gender x_1 and race/ethnicity x_2 in our example experiment, where x_1 can take values {*female, male*} and x_2 can take values {*Hispanic, non-Hispanic White, non-Hispanic Black, non-Hispanic Asian, Other*}. Thus, there are 17 possible subgroups. We denote $\hat{S}$ to be the set of 10 subgroups in which x_1 and x_2 take on unique values, e.g. *female Hispanic, male non-Hispanic white*, etc. The superset S also contains subsets defined using a single attribute, i.e., *female, male, Hispanic, non-Hispanic White, non-Hispanic Black, non-Hispanic Asian*, and *Other*. For $j \in \hat{S} \setminus S$, the function $child(j)$ returns the set of lower levels subgroups that make up j. In this example, if $j = Hispanic$, then $child(j)$ returns the subgroups *female Hispanic* and *male Hispanic*. These ideas can easily be generalized for any number of protected attributes.

TABLE I
MATHEMATICAL NOTATIONS IN MODEL

Notation	Definition
Sets	
S	Set of lowest-level subgroups
$\widehat{S}$	Set of subgroups
C	Set of recruitment centers
Indices	
j	Subgroup $\in \widehat{S}$
c	Recruitment center $\in C$
Parameters	
N	Trial size
$n_{j,target}$	Target enrollment size of subgroup j
l_j	Representativeness lower boundary for subgroup j
u_j	Representativeness upper boundary for subgroup j
rw^+	Overrepresentation penalty weight of subgroups
rw^-	Underrepresentation penalty weight of subgroups
dw	Target deviation penalty weight of subgroups
pw	Population availability deviation penalty weight
o_c	Recruitment cost of center c
$ow_{o,c}$	Recruitment cost penalty weight of center c
$p_{c,j}$	Proportion of subgroup j in center c
$n_{c,min}$	Site capacity lower bound of center c
$n_{c,max}$	Site capacity upper bound of center c
Variables	
n_j	Planned enrollment size of subgroup j
$n_{j,min}$	Previously enrolled number of subgroup j
rd_j^+	Positive deviation from representativeness upper bound u_j of subgroup j (i.e., overrepresentation)
rd_j^-	Negative deviation from representativeness lower bound l_j of subgroup j (i.e., underrepresentation)
dd_j^+	Positive deviation from target value of subgroup j
dd_j^-	Negative deviation from target value of subgroup j
e_c	Selection decision of center c
	$e_c = \begin{cases} 1, & \text{if center c is selected} \\ 0, & \text{otherwise} \end{cases}$
$n_{c,j}$	Planned enrollment size of subgroup j from center c
$p_{c,j}$	Enrollment proportion of subgroup j in center c
$pd_{c,j}^+$	Positive deviation from target proportion of subgroup j from site c
$pd_{c,j}^-$	Negative deviation from target proportion of subgroup j from site c

For subgroup $j \in \widehat{S}$, the ideal enrollment target $n_{j,target}$ is estimated by multiplying the trial size by the subgroup proportion from the target population. The actual subgroup size n_j should lie within either the confidence interval (CI) of the enrollment target (CI_j^{min}, CI_j^{max}) estimated using the target population data, or the limit range (n_j^{min}, n_j^{max}). Here, n_j^{min} is the smallest and n_j^{max} is the largest enrollment size for subgroup j, allowed by representativeness constraints.

Our previously work on RCT representativeness metrics [14], derived from machine learning fairness metrics, is used to evaluate enrollment representativeness. These metrics have a lower threshold τ_l and an upper threshold τ_u. The expected enrolled subgroup size should lead to a metric score within $[-\tau_l, \tau_l]$ and thus determine the limit range $[l_j, u_j]$. τ_u is used to further categorize representativeness levels (i.e., highly under-/over-representations). In our experiment, the metric

$$\text{Log Disparity} = \text{subgroup } log\frac{planned\ enrollment\ odds}{target\ enrollment\ odds}$$

is used to evaluate representation of a subgroup in enrollment planning compared to the target value estimated from target population. Other metrics such as enrollment fraction [34],

[35] and GIST 2.0 [36], [37], [38] could be used as alternative representativeness scores.

According to the CI and equitable range estimated by the metric, a subgroup j is equitably represented if n_j falls into a target range $[l_j, u_j] = [min(CI_j^{min}, n_j^{min}), max(CI_j^{max}, n_j^{max})]$.

Since multisite enrollment has goals that are applicable only when more than one recruitment center exist, we extend the single-site model by adding additional goals in the model objective and constraints to evaluate our approach under different enrollment scenarios.

A. Single-Site Enrollment Planning/Monitoring

Requirements:
1) Create an equitable enrollment plan with respect to protected attributes x for an RCT with trial size N.
2) Evaluate the representiveness of $N_{enrolled}$ enrolled participants based on interim data summaries with respect to protected attributes x and provide a remedial re-planning to reduce the deviations from original plan if subgroup underrepresentations are identified.

The GP model represents the two goals of single-site planning through objectives and constraints.

Objective function:

$$Min \sum_{j \in \widehat{S}} (rw^+ rd_j^+ + rw^- rd_j^- + dw \cdot dd_j^+ + dw \cdot dd_j^-) \quad (1)$$

In general, the model minimizes the total weighted deviations between subject distributions in enrollment plan and target population. For Goal 1, it minimizes the total representativeness range deviation ($\sum_{j \in \widehat{S}} rd_j^\pm$) for every subgroups j. For Goal 2, it minimizes the total target enrollment deviation ($\sum_{j \in \widehat{S}} dd_j^\pm$) for every subgroups j. The overrepresentation penalty weight (rw^+) has a default value 1 while the underrepresentation(rw^-) penalty weight has a default value 3. This helps the model to penalize more on underrepresentation, which may lead to more severe results than overrepresentation. These penalties can be customized to prioritize different subgroups.

Constraints:
1) Group Hierarchies: $\forall j \in \widehat{S} \setminus S$.

$$\sum_{k \in child(j)} n_k = n_j \quad (2)$$

2) Goal 1: Representative Plan Goal $\forall j \in \widehat{S}$

$$n_j + rd_j^- \geq l_j \quad (3)$$

$$n_j - rd_j^+ \leq u_j \quad (4)$$

$$rd_j^+, rd_j^- \geq 0 \quad (5)$$

3) Goal 2: Target Enrollment Goal $\forall j \in \widehat{S}$

$$n_j - dd_j^+ + dd_j^- = n_{j,target} \quad (6)$$

$$dd_j^+, dd_j^- \geq 0 \quad (7)$$

4) Restrictions on Study Size

$$\sum_{j \in S} n_j = N \tag{8}$$

$$\forall j \in S, n_j \geq n_{j,min} \geq 0 \text{ are integers} \tag{9}$$

Description of the constraints: Equation (2) ensure that each subgroup is the sum of its child subgroups. Equation (3) and (4) are goal constraints to minimize deviations from representativeness ranges. For n_j within the range, the model will match the enrollment sizes to target values, as shown in (6). Equation (8) ensures that the total enrollment size is within the planned trial size and (9) are hard constraints to ensure that the new subgroup sizes are integers that are greater than or equal to the current sizes. When design a new trial, $n_{j,min}$ equals 0; when monitor a trial during the process, $n_{j,min}$ is the number of subjects already enrolled that belongs to subgroup j. Non-negativity restrictions such as (5) and (7) are required for the implementation.

We use the R package "lpSolve" [33], which is designed to solve general linear/integer programs, to implement our models. To improve computational performance, we further simplify each GP by eliminating the linear constraints in the single site model and in the multi-site model before solving the GP.

B. Multisite Enrollment Planning

Requirements:
1) Create an overall enrollment plan, along with site-specific plans for a subset of selected recruitment centers selected from all available centers C, with respect to protected attributes x for an RCT with trial size N. For each site, there is an operation cost and capacity limit.

Objective function:

$$Min \sum_{j \in \widehat{S}} (rw^+ rd_j^+ + rw^- rd_j^- + dw \cdot dd_j^+ + dw \cdot dd_j^-)$$

$$+ \sum_{c \in C, j \in \widehat{S}} pw \cdot pd_{c,j}^+ + pw \cdot pd_{c,j}^-$$

$$+ \sum_{c \in C} e_c ow_{o,c} o_c \tag{10}$$

For multisite studies, additional goals on sites qualities, business rules, and resources are required. Here, we expect to satisfy site capacities (Goal 3), match previous site-specific enrollment availability (Goal 4), minimize total cost (Goal 5), and minimize the number of selected sites (Goal 6). These goals are examples to demonstrate the flexibility of our GP model to be adapt to other complex study conditions. Other goals can be added as needed. Here, the variables $pd^{\pm}$ help the model to match the previous enrollment distribution of participants and o_c helps reduce the total financial cost from select sites (i.e., all sites with $e_c = 1$) and minimize the number of sites. Since site capacities are "hard" goals which are not allowed to be violated, they are modeled as constraints.

Constraints: The constraints of single-site model are applicable to the multisite model. Additionally, new site-level constraints are added to achieve site-level goals.

1) Group Hierarchies $\forall j \in \widehat{S} \setminus S$

$$\sum_{c \in C} \sum_{k \in child(j)} (n_{c,k}) = n_{c,j} \tag{11}$$

2) Goal 1: Representative Plan Goal $\forall j \in \widehat{S}$

$$\sum_{c \in C} n_{c,j} + rd_j^- \geq l_j \tag{12}$$

$$\sum_{c \in C} n_{c,j} - rd_j^+ \leq u_j \tag{13}$$

$$rd_j^+, rd_j^- \geq 0 \tag{14}$$

3) Goal 2: Target Enrollment Goal $\forall j \in \widehat{S}$

$$\sum_{c \in C} n_{c,j} - dd_j^+ + dd_j^- = n_{j,target} \tag{15}$$

$$dd_j^+, dd_j^- \geq 0 \tag{16}$$

4) Restrictions on Study Size

$$\sum_{c \in C} \sum_{j \in S} n_{c,j} = N \tag{17}$$

$$\forall j \in S, \sum_{c \in C} n_{c,j} \geq n_{j,min} \geq 0 \text{ are integers} \tag{18}$$

5) Goal 3: Restriction on Site Capacity $\forall c \in C$

$$e_c \in \{0, 1\} \tag{19}$$

$$e_c n_{c,max} \geq \sum_{j \in S} n_{c,j} \geq e_c n_{c,min} \tag{20}$$

6) Goal 4: Enrollment Availability Goal $\forall j \in S, \forall c \in C$

$$n_{c,j} - pd_{c,j}^+ + pd_{c,j}^- = p_{c,j} \sum_{j \in S} n_{c,j} \tag{21}$$

$$pd_j^+, pd_j^- \geq 0 \tag{22}$$

Description of the constraints:

Similar to the single-site model, (12), (13), and (15) force the model to generate a plan with maximal overall representativeness (Goal 1 and Goal 2). Equation (17) restricts the total enrollment size and (18) ensure the new subgroup size is reasonable. Some new goals for multisite model are added. Since we want to reduce the total number of sites when possible, we introduced a new decision variable e_c to decide whether the site $c \in C$ is selected into the plan, displayed in (19). Then, for each selected site, the model ensures the number of enrollees within site's maximum capacity and a practical minimum number of the site, as shown in (20). Finally, (21) make the proportion of the planned enrollment from a site for a specific subgroup match the actual enrollment of the site. Patient availability distributions are estimated by previous participant data of each site.

Besides the constraints mentioned above, constraints relevant to particular clinical trial design, such as $\sum_{c \in C} e_c \leq 10$ which requires a maximum of 10 sites to be selected and $e_3 + e_9 = 1$ which requires to select one center from site 3 and site 9, can be added.

C. Model Input Sources

To create an enrollment plan, our model requires the following three groups of raw inputs: trial size, eligibility criteria, and target population source data based on trial design; protected attributes and representativeness measurements for equity evaluation; and site-level information such as capacity, enrollment history data, and financial cost for multisite planning.

To illustrate the single-site enrollment modeling process, we imagine redoing a real-world RCT, the Systolic Blood Pressure Intervention Trial (SPRINT, NCT01206062) trial, assuming two different scenarios described in Section II-A: designing an equitable plan to enroll N subjects (A1) and making a remedial re-planning if inequities are identified in the interim analysis (A2). For A1, from the SPRINT participant data, we get the trial size of 9361 and trial specific eligibility criteria. Then, we use the NHANES, which is designed to provide nationally representative estimates for the whole U.S. population, as our source of target population estimation. NHANES participants that satisfy the SPRINT eligibility criteria are treated as target population. Details of target estimation is available in Appendix I. All experiments use the two protected attributes race/ethnicity and gender to demonstrate the model's ability on multivariate cases. The representativeness measurement used is Log Disparity [14], which is introduced at the beginning of Section II. For A2, additional information of enrolled SPRINT participants is needed since we want to improve the representativeness of protected subgroups through the rest of recruitment process based on current enrollees.

For multisite trials, we experiment under two scenarios. The first scenario assumes that we are trying to redo the SPRINT trial using only clinical centers in the New York State (NYS). Thus, besides inputs of single-site shown in previous case A1, we need clinical sites information in NYS. We collected clinical studies under the following conditions from ClinicalTrials.gov: (1) Condition or disease: Hypertension; (2) Study type: Interventional Studies (Clinical Trials); (3) Recruitment: Completed; (4) Age groups: Adult (18–64) & Older Adult (65+); and (5) Location: New York State, US. The prevalence of subgroups defined over race/ethnicity and gender are estimated from the data provided by the County Health Indicators by Race/Ethnicity (CHIRE) [39] and county-level cardiovascular disease hospitalization data [40]. We assume that gender and race/ethnicity of the population who have heart disease are independent to obtain multivariate subgroup summaries. For each county, we collected the following data: (1) county size, (2) heart disease hospitalizations rates, (3) heart disease hospitalizations rates for each racial/ethnic groups, (4) percentage of each racial/ethnic in county, (5) heart disease hospitalizations rates for each gender, and (6) percentage of each gender in the county. Then, according to the Bayes' Theorem, we calculated the proportion of each race/ethnicity and gender in the total population with heart disease. The financial cost for each site is estimated using a normal distribution with mean = 805,785 according to the United States Department of Health and Human Services (HHS) [41] and sd = mean/3. For scenario 2, we redo the SPRINT trial with fewer participants and sites to show that an appropriate planning could

TABLE II
COLOR DESCRIPTION FOR VISUALIZATIONS

Color	Description	Representativeness
Red	Absent	-Inf
Orange	Highly under-represented	$< -\tau_u$
Light orange	Under-represented	$[-\tau_u, -\tau_l)$
Teal	Adequately represented	$[-\tau_l, \tau_l)$ or $p > 0.05$
Light blue	Over-represented	(τ_l, τ_u)
Blue	Highly over-represented	$\geq \tau_u$

reduce the use of resources to achieve the same goals. This requires inputs of single-site described in case A1 and with site-level information from the SPRINT.

D. Statistical Analysis and Visualizations

To monitor clinical trial enrollment performance regarding to representativeness, metric values are categorized and represented in different color codes as shown in Table II. The colored visualizations facilitate underrepresentation identification/monitoring and comparison before and after refinements in enrollment planning.

III. RESULTS

To illustrate our models and evaluate their performance, we apply them to the real-world multicenter RCT of hypertension SPRINT, which enrolled 9,361 participants from 102 sites across the United States [42]. The SPRINT target population, which is the U.S. population with known hypertension who satisfied SPRINT eligibility criteria, is estimated from the NHANES 2015-2016. Equity evaluations are based on the example equity metric Log Disparity. To demonstrate the flexibility of the GP approach, we explore the results for different scenarios defined by example requirements.

A. Single-Site Planning for New RCTs

Example Requirement Make an equitable enrollment plan for a new hypothetical single-site SPRINT study based on the Log Disparity metric which follows the 80% rule. (i.e., design a single-site SPRINT study that satisfies Goals 1 and 2.) Here, the 80% rule, previously used in [14], means that enrollment of protected groups should be at least 80% of that of unprotected groups.

Data Sources NHANES 2015-2016 and SPRINT study design.

Enrollment Planning Based on the target population estimated from NHANES, we calculate a target enrollment range for which the representativeness scores of subgroups from Log Disparity metric fall into the adequate representation range described in Table II. The new plan is shown in Table III under "Target Enrollment Plan." The "Score" presents equity evaluation results based on the selected metric, which proves that all subgroups of interest are equitably represented in the suggested plan.

In Table IV, we demonstrate how the proposed plan can be easily adapted to the NIH enrollment reporting form.

<div align="center">

TABLE III

RACE/ETHNICITY AND GENDER. NEW SINGLE-SITE SPRINT TRIAL SIZE = 9360; 2 × SPRINT TRIAL SIZE = 18722. NH = NON-HISPANIC

</div>

Subject Characteristics		Target Enrollment Plan				SPRINT Trial		2 × SPRINT Trial		
Race/ Ethnicity	Gender	Enrollment Target	Lowest Enrollment	Highest Enrollment	Score	Actual Enrollment	Score	New Enrollment	Cumulative Enrollment	Score
All	Female	5762	5256	6241	0.000	3331	-1.064	8190	11521	0.000
NH Asian	Female	367	296	453	0.001	25	-2.723	709	734	0.001
NH Asian	Male	140	112	173	0.004	50	-1.035	174	224	-0.222
NH Asian	All	507	410	624	0.002	75	-1.956	883	958	-0.058
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
Hispanic	All	1009	825	1227	0.001	984	-0.027	892	1876	-0.080

<div align="center">

TABLE IV

SUGGESTED NIH ENROLLMENT PLAN FOR SPRINT

</div>

	Ethnic Categories				
	Not Hispanic or Latino		Hispanic or Latino		Total
Racial Categories	Female	Male	Female	Male	
Asian	366	140	0	0	506
Black	586	336	0	0	922
White	3963	2610	0	0	6573
Other	202	149	644	364	1359
Total	5117	3235	644	364	9360

We harmonized the race/ethnic categories from the target population data to make them match those required by NIH.

B. Single-Site Monitoring for Conducted/Ongoing RCT

Example Requirement The data monitoring committee needs to perform the planned interim analysis of the data at the halfway point to consider limiting enrollment of patients from subgroups of interest in the continuing trial. How should the enrollment plan be adjusted to make the cumulative sample more equitably represent the target population in order to achieve Goals 1 and 2?

Data Source NHANES 2015-2016 and SPRINT enrollment summary data.

Enrollment Planning The actual SPRINT enrollment for subgroups are shown as "SPRINT Trial" in Table III. Subgroups such as female and non-Hispanic Asian are highly underrepresented in the SPRINT compared to the target population. Since the interim analysis is performed at the halfway point, we assume that another 9,361 participants will be enrollment through the rest of the enrollment process. To maximally optimize the equity across all subgroups of interest defined over race/ethnicity and gender, we, for example, suggest to recruit 709 non-Hispanic Asian female subjects. The final equity evaluation shows that most subgroups' equity levels are improved. Additionally, we estimate that 21,469 new subjects are needed to achieve an equitable cumulative RCT based on the performance of the first half of enrollment process. This indicates the importance of equitable planning at the design stage if clinical researchers do not expect to compensate with greater additional efforts in later stages.

Fig. 2 shows that the suggested plan improves the representativeness of subgroups which were identified as inequitably represented during the interim analysis. Each point represents

the target versus the actual/panned enrollment rates of a subgroup, green for gender subgroups, purple for racial/ethnic subgroups, and orange for multivariate subgroups defined over gender and race/ethnicity. The black line is the target which means that subgroups are perfectly represented; the grey region is the area for adequate subgroup representation calculated based on statistical tests and representativeness metrics (i.e., equity score $\in [-\tau_l, \tau_l]$ or p > 0.05). The slope of the red line is created by linearly regressing the subgroup enrollment and target rates indicates the representativeness of the enrollment plan for all subgroups. The figure of interim analysis (i.e., SPRINT results) shows that 16 of 17 subgroups, except Hispanic subjects, are out of the equitable region and some subgroups are underrepresented (e.g., female, especially the non-Hispanic White females); if we treated the SPRINT study as a midpoint of recruitment, a plan that is very close to the targets could be achieved in the following recruitment process. Only 2 out of 17 (i.e., non-Hispanic Black and non-Hispanic Black male) subgroups are slightly out of the boundaries.

C. Multisite Planning With Site Selection

Example Requirement Solve an area-constrained clinical trial site selection problem. Specifically, create an enrollment plant representative of the entire U.S. using sites only located in New York State. The enrollment plan should be equitable while minimizing the recruitment cost (i.e., achieve Goal 1 to Goal 6).

Data Sources Data are collected and aggregated from several heterogeneous data sources: the NHANES data, completed RCTs of hypertension in NYS from clinicaltrials.gov, and population distribution of NYS counties from health.data.ny .gov.

Enrollment Planning Based on the completed RCTs of hypertension in NYS, we estimate the available recruitment centers distributed in NYS. For each RCT, we estimate the single-site enrollment as the average of the total study size divided by the number of sites involved; then we removed the sites outside NYS (some sites in multicenter RCTs are outside the NYS) and treat each of the remaining site as a candidate for recruitment. We eventually obtain 80 sites. The distribution of subjects is approximated by the hypertension population of each county that the site belongs to. For each site, the lower capacity is 11 and the upper capacity is 20% over the site enrollment size. The site financial cost is estimated by the normal distribution with mean= \$805,785 and a standard deviation = mean/3. Since reducing underrepresentation of marginalized groups is
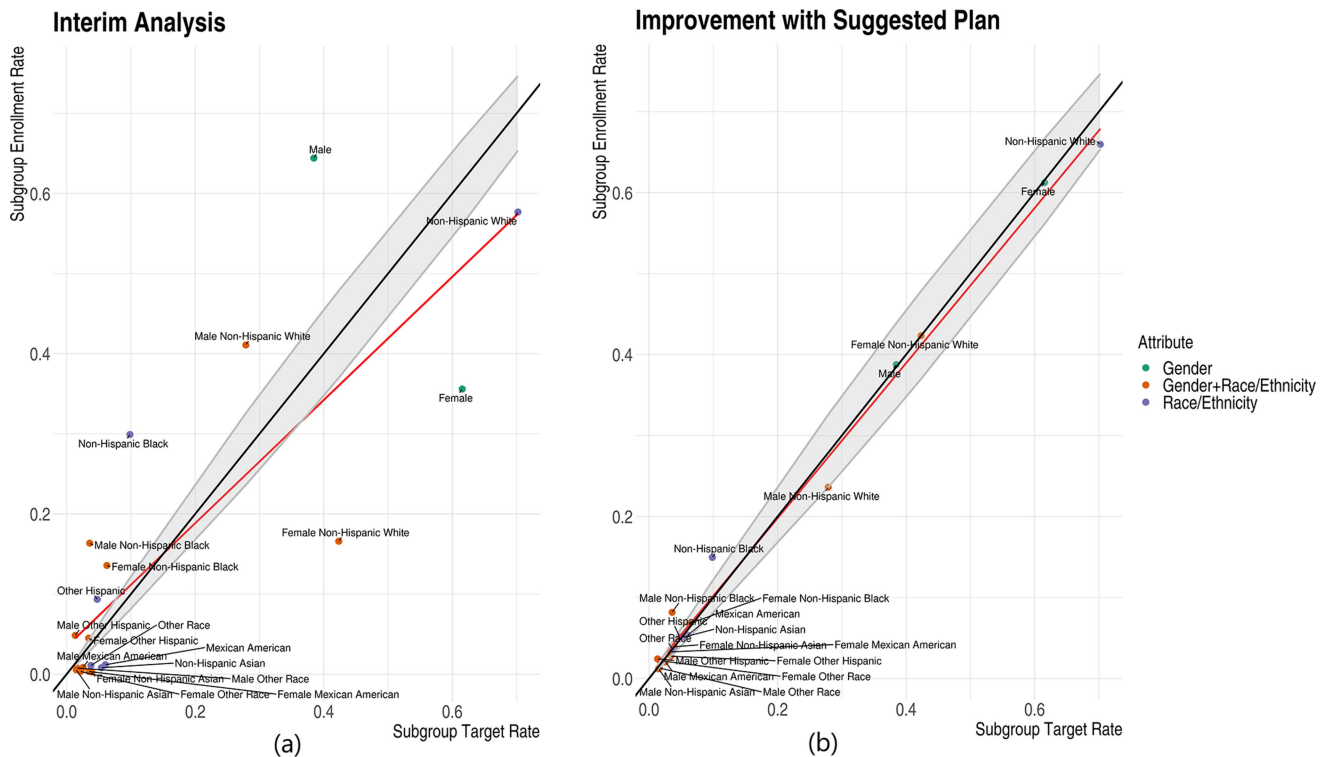
Fig. 2. Evaluation of representativeness of the actual SPRINT enrollment, proposed enrollment planning using 2 × SPRINT trial size, and ideal target enrollment. (a) Actual SPRINT enrollment (treated as interim analysis point) vs ideal target enrollment; (b) Proposed planning with 2 × SPRINT trial size vs ideal target enrollment. Green points are subgroups defined over gender (female and male); purple points are subgroups defined over race/ethnicity (Hispanic, non-Hispanic White, non-Hispanic Black, and non-Hispanic Asian); orange points are multivariate subgroups defined over both gender and race/ethnicity (all gender-race/ethnicity combinations). The black line is the ideal relationship between the real/planned enrollment and target enrollment rates; the grey region is the equitable representation region; the red line is the linear regression between the real/planned enrollment and target enrollment rates.

TABLE V
RESULTS WITH DIFFERENT GOAL WEIGHTS

Underrepresentation Penalty	Overrepresentation Penalty	Enrollees Availability Penalty	Financial Cost Penalty	Number of Sites	Underrepresentation Violation	Distribution Violation	Financial Cost
3	1	1	1000	3	0	0.469	$2540851
3	1	1	10000	12	115	1.155	$8480263
3	1	3	1000	6	0	0.603	$4867988
3	3	1	1000	3	0	0.487	$2540851

more critical than preventing than overrepresentation, we use the default underrepresentation penalty of 3; the site financial cost is in thousands; all other penalties are default with 1. Different penalties could determine the priorities/preferences of decision makers for different goals. For instance, the financial cost penalty determines the number of selected sites. As shown in Table V, if we scale the financial cost by 1/1,000, a plan based on 3 sites is designed; if the penalty becomes 1/10,000, 12 sites compose the final plan. If researchers want more sites to recruit participants simultaneously based on the time constraints, the estimated total financial cost would be tripled, some subgroups will be underrepresented, and the enrollment availabilities of sites are harder to satisfy. In Table V, we display how the change in goal priorities/weights could influence the achievement of different goals in the plan. The first row of Table V displays the default penalty values and can be treated as the baseline for comparisons with the other rows.

In this experiment, we aim to recruit 5,000 participants from the 80 available NYS sites. For each site, add-k smoothing method is available if some subgroups are missing from the prior enrollment availability. The overall plan is displayed in the "Overall Plan" in Table VI.

D. Multisite Planning Using SPRINT Sites

Example Requirement How to make an equitable enrollment planning using about half of the SPRINT size (i.e., 5,000) and fewer sites from the previous SPRINT study?

Data Source NHANES 2015-2016 and the per-site SPRINT enrollment summary.

Enrollment Planning We select the 51 (i.e., half of the sites from SPRINT) largest recruitment centers from SPRINT and design a plan based on the previous participant distributions. The lower capacity of each site is equal to the smallest site (except

TABLE VI
MULTISITE PLANNING OF SPRINT

Race/ Ethnicity	Gender	Overall Plan	Enrollment Range	Site A	Site B	Site C	Score	SPRINT Score
Hispanic	Female	344	280-422	0	272	72	-0.000	-0.376
Hispanic	Male	195	157-240	0	195	0	0.003	0.399
Hispanic	All	539	441-655	0	467	72	0.001	-0.027
All	Female	3077	2808-3333	554	1490	1033	-0.000	-1.064
All	Male	1923	1667-2192	251	821	851	0.000	1.064
NH Asian	Female	195	158-242	7	47	141	-0.004	-2.723
NH Asian	Male	75	60-92	0	48	27	0.007	-1.035
NH Asian	All	270	219-333	7	95	168	-0.001	-1.956
NH Black	Female	313	254-385	0	238	75	0.000	0.855
NH Black	Male	180	145-222	0	180	0	0.003	1.658
NH Black	All	493	402-600	0	418	75	0.001	1.363
NH White	Female	2117	1850-2392	547	825	745	0.000	-1.306
NH White	Male	1392	1181-1628	251	317	824	-0.002	0.590
NH White	All	3509	3268-3732	798	1142	1569	-0.001	-0.548
Other Race	Female	108	87-134	0	108	0	0.001	-1.892
Other Race	Male	81	65-100	0	81	0	0.006	-0.762
Other Race	All	189	152-233	0	189	0	0.003	-1.268

the site with 1 subject); the upper capacity bound is 20% over the enrollment size of site provided in SPRINT. This experiment aims to show if researchers could use fewer subjects to achieve a more representative plan based on fewer sites.

The overall plan based on the 51 largest SPRINT sites is provided in Table VI. It shows that 30 sites could make a more representative sample. Three site-level targets are provided in the Table as sites A, B, and C. For some sites, the target value for small subgroups may be 0. This means that recruitment of subjects from that subgroup at that site is not needed to meet the overall plan goals. Subjects in that subgroup may still be recruited at that site.

For GP, a numeric target will be estimated for each goal. The objective of GP is always to minimize the total deviations from targets but priorities of goals could be designed through weights of deviations. In our experiment, we care about the deviation from goals in both directions (i.e., negative deviations such as underrepresentation and positive deviations such as overrepresentation). We assume the positive deviation which is overrepresentation of subgroups is not considered as severe as underrepresentation. So, we assign a larger weight for the negative deviation from equitable enrollment lower bound. Investigators could decide the priorities of each goal in the framework.

For multisite RCTs, subgroup enrollment targets are set for the overall plan. For each selected site, the total target enrollment is calculated. Since site-level population distribution could be estimated from site's history enrollment data, the number of participants that could be enrolled from each subgroup of interest is determined given the site total. The equity evaluation is applied to the overall plan across all selected sites. In our experiment, since enrollment rates of sites are assumed to be the same, time-related constraints are not included in the model but could be easily added if needed.

clinical researchers of intervention effects on a diverse patient population. The more robust and applicable findings RCTs generate, the more equitable and appropriate decisions healthcare providers make for the patients; thus, health equity is promoted.

Our single- and multi-site models can generate candidate enrollment plans for RCTs based on enrollment targets, trial-specific constraints, and study goals of clinical research. The targets can be used to specify enrollment of ethnic, racial, and gender subgroups for the NIH-required Planned Inclusion Enrollment form. These modeling and visualization techniques could help researchers explicitly incorporate representativeness considerations into trial design to reduce baseline characteristic differences between trial population and intention-to-treat population in RCTs. Additionally, the monitoring and remediation of enrollment target deviations ensure diverse patient populations, including previously underrepresented subgroups, can be adequately represented in the RCTs and can indicate any subgroup access challenges in the study. Our real-time evaluation could allow an optimal representativeness of patients.

In general, our main contributions are:
- We present a novel GP-based model to design enrollment and monitoring plans with optimal representativeness for RCTs using well-defined representativeness metrics with investigator-defined target populations as benchmarks.
- We develop a centralized monitoring approach of multi-site clinical trial enrollment to enable sponsor to identify sites with deviations from the enrollment planning targets based on interim data.
- We provide a scalable multi-objective optimization approach to select sites based on multiple (conflicting) objectives by assigning different weights on individual parameters or by setting priorities for each deviation from target values.

The GP framework can potentially be adaptable to a variety of study goals with different contexts with further research. For example, it could be used to design enrollment planning and monitor enrollment processes in low-resource settings, where cost constraints impact enrollment planning. Resources could be minimized through objectives and restricted within a

IV. DISCUSSION AND CONCLUSIONS

Considering DEI in RCT enrollment planning and monitoring is critical to clinical research having broad societal impact. By creating generalizable knowledge, such considerations can promote a better understanding by healthcare providers and

range through additional constraints. Additionally, our framework could help researchers to satisfy a variety of federal laws and regulations (e.g., Public Health Service Act sec. 492B, 42 U.S.C. sec. 289a-2), policies (e.g., NIH policy for the Inclusion of Women and Minorities), and procedures (e.g., NIH Enrollment Report) that address health equity through study design, conduct, and reporting. If federal or local government policy is extended to other protected attributes, such as sexual orientation or socioeconomic status, investigators could include these attributes in the framework to ensure representativeness of these subgroups of interest. Also, our framework could be enhanced to assist study review committees (e.g., IRB or DSMB) in performing oversight of DEI in trials during both planning and monitoring phases. A DSMB could use our framework to perform evaluation of interim, cumulative data to ensure that the enrollment follows pre-established statistical guidelines, to examine performance of individual site enrollment in multisite RCTs, to confirm adequacy of compliance with enrollment goals for diverse populations, and to decide whether the trial investigators should take corrective actions such including more sites or longer time for recruitment to enable equitable representations of protected subgroups. Furthermore, the framework could be extended to address other public health problems, such as access to healthcare services by adjusting DEI goals to other contexts of patient outreach. Finally, the framework could also be modified to perform evaluation of representativeness in observational studies or to design an equitable synthetic control arm using EHR data. We leave these potential enhancement to these scenarios as future research.

We have validated our framework using target populations from NHANES, a nationally representative database that is de-identified and used widely in the biomedical research community. As we noted, the population of patients that are captured in an EHR system can also serve as the basis of the target population. In this case, clinical investigators will need to ensure that the EHR-derived data does not contain protected health information or is a limited data set as approved by an IRB.

Our study has several limitations. First, not all eligibility criteria from the SPRINT trial could be applied to the NHANES dataset since relevant data was not collected in NHANES. A more accurate target population may be estimated if researchers have access to data sources with complete clinical records. Second, county-specific multi-trait values of enrollees, such as the proportion of female subjects who are Hispanic from the site in each county, could only be estimated from other data sources instead of directly obtained from the sites. The limited access permission to patient records affects the model performance. Furthermore, several sites of SPRINT recruited so few patients that it influenced our estimation of site-level information. Finally, the empiricism of our approach needs to evaluated in real time during planned or ongoing clinical trials. We plan to engage researchers to evaluate our tool and give feedback for further improvement. We note that SPRINT has no opinion on the experiments we conducted in this work. A practical challenge of applying our framework for multisite planning is the ability

to gather data on population characteristics at different sites. Although we validated this capability in using county-level population data available for New York state, such data may not be readily available from other states. Our framework could be leveraged by investigators at large multi-center research networks, such as the NIH-funded Trial Innovation Network (https://ncats.nih.gov/ctsa/projects/network). The Trial Innovation Network was designed to support multisite trials by allowing IRB harmonization, data sharing and subject recruitment at over 50 academic medical centers. As such clinical research networks expand and become more robust, we expect that there will be an increasing ability to apply frameworks such as our GP approach to address representativeness in multisite clinical trial enrollment.

Further research should enhance the models to handle more complex eligibility criteria such as comorbidities or patient safety [43] using semantic AI to improve model performance in real-world clinical settings, automate the generation of NIH-required Inclusion Enrollment Report for investigators to design more achievable plans which embeds NIH principles of DEI, and evaluate the model in future trial recruitment for further improvement.

APPENDIX

TARGET POPULATION ESTIMATION

The target population of an RCT is often estimated from surveillance datasets such as NHANES or EHRs. In this study, we calculate the target population from NHANES 2015-2016 using the R survey () package [44] with additional filters based on researched disease conditions and eligibility criteria. To estimate the target population of the SPRINT study, we first collected information of NHANES participants in years 2015-2016 who are over 20 years old with available systolic blood pressure readings and anti-hypertensive medication information and who got the fasting test. Then we applied a set of inclusion criteria

1) Age $\geq$ 50 years old
2) Systolic blood pressure (SBP) Criteria
   a) 130–180 mm Hg on 0 or 1 antihypertensive medication class
   b) 130–170 mm Hg on up to 2 antihypertensive medication classes
   c) 130–160 mm Hg on up to 3 antihypertensive medication classes
   d) 130–150 mm Hg on up to 4 antihypertensive medication classes
3) High CVD risk Criteria (one or more of the following)
   a) Clinical coronary heart disease (CHD)
   b) Estimated glomerular filtration rate (eGFR) 20–59 mL/min/1.73 m^2
   c) Framingham risk score for 10-year CVD risk $\geq$ 15
   d) Age $\geq$ 75 years old

We classified the prescribed drugs taken by the NHANES participants into 10 categories according to the SPRINT protocol document: diuretics (thiazide-type, loop, and potassium-sparing), angiotensin converting-enzyme (ACE) inhibitors, angiotensin receptor blockers (ARBs), calcium channel blockers (CCBs), beta-blockers, vasodilators, alpha 2 agonists, adrenergic inhibitor, renin inhibitors, and alpha-blockers. This information is used to estimate eligible NHANES participants for the SPRINT study.

We did not apply the exclusion criteria (i.e., diabetes, history of stroke, proteinuria > 1 g in 24 h, heart failure, eGFR$<$ 20 ml/min/1.73 m^2 or dialysis) in order to keep enough sample size of subgroups of interest for further analysis.

ACKNOWLEDGMENT

REFERENCES

[1] National Academies of Sciences, Engineering, and Medicine, *Strategies for Ensuring Diversity, Inclusion, and Meaningful Participation in Clin. Trials: Proc. of a Workshop*. Washington, D.C., USA: Nat. Acad. Press2016.

[2] National Institute of Health, "Guidelines for the review of inclusion on the basis of sex/gender, race, ethnicity, and age in clinical research," National Institute of Health, Bethesda, Maryland, USA. Mar. 18, 2019. [Online]. Available: https://grants.nih.gov/grants/peer/guidelines_general/Review_Human_subjects_Inclusion.pdf

[3] U.S. Food and Drug Administration, "Enhance EQUITY initiative," Food and Drug Administration, Dec. 9, 2021. [Online]. Available: https://www.fda.gov/consumers/minority-health-and-health-equity/enhance-equity-initiative

[4] National Institute of Health, "Inclusion enrollment report," National Institute of Health, Maryland Bethesda, USA. Dec. 1, 2021. [Online]. Available: https://era.nih.gov/erahelp/assist/Content/ASSIST_Help_Topics/3_Form_Screens/PHS_HS_CT/Incl_Enroll_Rprt.htm

[5] National Institute of Health, "Inclusion across the lifespan," National Institute of Health, Maryland Bethesda, USA. Dec. 29, 2021. [Online]. Available: https://grants.nih.gov/policy/inclusion/lifespan.htm

[6] National Institute of Health, "Inclusion of women and minorities as participants in research involving human subjects," National Institute of Health, Bethesda, Maryland, USA. Dec. 29, 2021. [Online]. Available: https://grants.nih.gov/policy/inclusion/women-and-minorities.htm

[7] U. S. "Food and Drug Administration, Guidance for industry. E9 statistical principles for clinical trials," Food and Drug Administration Sep. 1998. [Online]. Available: https://www.fda.gov/media/71336/download

[8] U. S. "Department of Health and Human Services, Part 46 - protection of human subjects," Health and Human Services, Jan. 2017. [Online]. Available: https://www.ecfr.gov/on/2018-07-19/title-45/subtitle-A/subchapter-A/part-46

[9] B. Nazha et al., "Enrollment of racial minorities in clinical trials: Old problem assumes new urgency in the age of immunotherapy," *Amer. Soc. Clin. Oncol. Educ. Book*, vol. 39, pp. 3–10, Jan. 2019, doi: 10.1200/EDBK_100021.

[10] F. Varse et al., "Challenges in the design, conduct, analysis, and reporting in randomized clinical trial studies: A systematic review," *Med. J. Islamic Republic Iran.*, vol. 33, Apr. 2019, Art. no. 37, doi: 10.34171/mjiri.33.37.

[11] E. Gkioni et al., "A systematic review describes models for recruitment prediction at the design stage of a clinical trial," *J Clin. Epidemiol.*, vol. 115, pp. 141–149, Nov. 2019, doi: 10.1016/j.jclinepi.2019.07.002.

[12] J. Kutner et al., "Methodological challenges in conducting a multi-site randomized clinical trial of massage therapy in hospice," *J. Palliat. Med.*, vol. 13, no. 6, pp. 739–744, Jun. 2010, doi: 10.1089/jpm.2009.0408.

[13] B. Kasenda et al., "Prediction of RECRUITment in randomized clinical trials (RECRUIT-IT)–rationale and design for an international collaborative study," *Trials*, vol. 21, no. 1, pp. 1–7, Aug. 2020, doi: 10.1186/s13063-020-04666-8.

[14] M. Qi et al., "Quantifying representativeness in randomized clinical trials using machine learning fairness metrics," *JAMIA Open*, vol. 4, no. 3, Sep. 2021, doi: 10.1093/jamiaopen/ooab077.

[15] "National center for health statistics, national health and nutrition examination survey data," 1999. [Online]. Available: https://wwwn.cdc.gov/nchs/nhanes/

[16] X. Zhang and Q. Long, "Modeling and prediction of subject accrual and event times in clinical trials: A systematic review," *Trials*, vol. 9, no. 6, pp. 681–688, Dec. 2012, doi: 10.1177/1740774512447996.

[17] K. D. Barnard, L. Dent, and A. Cook, "A systematic review of models to predict recruitment to multicentre clinical trials," *BMC Med. Res. Methodol.*, vol. 10, no. 1, pp. 1–8, Jul. 2009, doi: 10.1186/1471-2288-10-63.

[18] G. Mijoule, S. Savy, and N. Savy, "Models for patients' recruitment in clinical trials and sensitivity analysis," *Statist. Med.*, vol. 31, no. 16, pp. 1655–1674, Jul. 2012, doi: 10.1002/sim.4495.

[19] V. V. Anisimov, "Using mixed poisson models in patient recruitment in multicentre clinical trials," in *Proc. Int. Conf*, London, U.K., 2008, pp. 1046–1049.

[20] V. Anisimov, D. Downing, and V. Fedorov, "Recruitment in multicentre trials: Prediction and adjustment," in *Proc. Adv. Model-Oriented Des. Anal.*, 2007, pp. 1–8.

[21] V. Anisimov, "Modern analytic techniques for predictive modeling of clinical trial operations," in *Quantitative Methods in Pharmaceutical Research and Development*. Berlin, Germany: Springer, 2020, pp. 361–408.

[22] Y. Lan, G. Tong, and D. F. Heitjan, "Statistical modeling and prediction of clinical trial recruitment," *Statist. Med.*, vol. 38, no. 6, pp. 945–955, Nov. 2018, doi: 10.1002/sim.8036.

[23] M. Schepps et al., "Optimizing patient enrollment in clinical trials with metaheuristics," *Biostatistics*, vol. 11, no. 2, 2021.

[24] I Abbas, J. Rovira, and J. Casanovas, "Clinical trial optimization: Monte carlo simulation Markov model for planning clinical trials recruitment," *Contemporary Clin. Trials*, vol. 28, no. 3, pp. 220–231, May 2007, doi: 10.1016/j.cct.2006.08.002.

[25] National Institute of Health, "NINDS guidelines for monitoring in clinical trials," NIH. Maryland Bethesda, USA. Dec. 31, 2019. [Online]. Available: https://www.ninds.nih.gov/funding/apply-funding/application-support-library/clinical-research-guidelines/ninds-guidelines-monitoring-clinical-trials

[26] S. F. Molloy and P. Henley, "Monitoring clinical trials: A practical guide," *Trop. Med. Int. Health*, vol. 21, no. 12, pp. 1602–1611, Dec. 2016, doi: 10.1111/tmi.12781.

[27] A. Chandra and J. S. Skinner, "Geography and racial health disparities," in *Critical Perspectives on Racial and Ethnic Differences in Health in Late Life*, Washington, D.C., USA: Nat. Acad. Press, 2004.

[28] V. Paluy and V. Shnaydman, "Portfolio approach to optimize site selection," *Appl. Clin. Trials*, vol. 28, no. 3, pp. 28–34, 2019. [Online]. Available: https://www.appliedclinicaltrialsonline.com/view/portfolio-approach-optimize-site-selection

[29] B. Harper and D. Zuckerman, "Critical success factors for planning for site selection and patient recruitment planning," *BioExecutive Int.*, vol. 2, no. 6, pp. 16–28, Jun. 2006.

[30] Niche, "Site selection - identifying high performing clinical sites: An insider's insight," *Niche Sci. Technol. Ltd.*, 2017. [Online]. Available: http://www.niche.org.uk/asset/insider-insight/Insider-Site-Selection.pdf

[31] D. Kurbegov et al., "Recommendations to streamline and standardize clinical trial site feasibility assessments: An ASCO research statement," *JCO Oncol. Pract.*, vol. 17, no. 1, pp. 41–45, Jan. 2021, doi: 10.1200/OP.20.00821.

[32] M. Gheorghiade et al., "Site selection in global clinical trials in patients hospitalized for heart failure: Perceived problems and potential solutions," *Heart Failure Rev.*, vol. 19, no. 2, pp. 135–152, Mar. 2015, doi: 10.1007/s10741-012-9361-8.

[33] M. Berkelaar et al., "Interface to 'lp_solve' v. 5.5 to solve linear/integer programs." Jan. 24, 2020. [Online]. Available: https://cran.r-project.org/web/packages/lpSolve/lpSolve.pdf

[34] V. H. Murthy, H. M. Krumholz, and C. P. Gross, "Participation in cancer clinical trials race-, sex-, and age-based disparities," *JAMA.*, vol. 291, no. 22, pp. 2720–2726, Jun. 2004. doi: 10.1001/jama.291.22.2720.

[35] Agency for Healthcare Research and Quality, "Toolkit for using the AHRQ quality indicators," Agency for Healthcare Research and Quality, Mar., 2017. [Online]. Available: https://www.ahrq.gov/patient-safety/settings/hospital/resource/qitool/index.htm

[36] A. Sen et al., "GIST 2.0: A scalable multi-trait metric for quantifying population representativeness of individual clinical studies," *J. Biomed Inform.*, vol. 63, pp. 325–336, Oct. 2016. doi: 10.1016/j.jbi.2016.09.003.

[37] Z. He et al., "Simulation-based evaluation of the generalizability index for study traits," in *Proc. AMIA Annu Symp Proc.*, San Francisco, CA, USA, 2015, pp. 594–603.

[38] T. Kennedy-Martin et al., "A literature review on the representativeness of randomized controlled trial samples and implications for the external validity of trial results," *Trials*, vol. 16, Nov., 2015, Art. no. 495, doi: 10.1186/s13063-015-1023-4.

[39] New York State Department of Health, "County health indicators by race/ethnicity (CHIRE)," Nov. 2021. [Online]. Available: https://www.health.ny.gov/statistics/community/minority/county/index.htm

[40] New York State Department of Health, "The burden of cardiovascular disease in New York: Mortality, prevalence, risk factors, costs, and selected populations," [Online]. Available: https://www.health.ny.gov/diseases/cardiovascular/heart_disease/docs/burden_of_cvd_in_nys.pdf

[41] U. S. "Department of Health & Human Services, Examination of clinical trial costs and barriers for drug development," *Health & Hum. Serv.* Jul. 2014. [Online]. Available: https://aspe.hhs.gov/sites/default/files/migrated_legacy_files//44516/rpt_erg.pdf

[42] W. T. Ambrosius et al., "The design and rationale of a multicenter clinical trial comparing two strategies for control of systolic blood pressure: The systolic blood pressure intervention trial (SPRINT)," *Clin. Trials.*, vol. 11, no. 5, pp. 532–546, Oct. 2014, doi: 10.1177/1740774514537404.

[43] J. R. Rogers et al., "Leveraging electronic health record data for clinical trial planning by assessing eligibility criteria's impact on patient count and safety," *J. Biomed. Inform.*, vol. 127, no. 1, Mar., 2022, Art. no. 104032, doi: 10.1016/j.jbi.2022.104032.

[44] T. Lumley, "Analysis of complex survey samples," *J. Statist. Softw.*, vol. 16, Apr. 2004, Art. no. 495, doi: 10.18637/jss.v009.i08.