RESEARCH

Open Access

Group sequential designs for pragmatic clinical trials with early outcomes: methods and guidance for planning and implementation

Nick R. Parsons[1*], Joydeep Basu[1] and Nigel Stallard[1]

Abstract

Background Group sequential designs are one of the most widely used methodologies for adaptive design in randomized clinical trials. In settings where early outcomes are available, they offer large gains in efficiency compared to a fixed design. However, such designs are underused and used predominantly in therapeutic areas where there is expertise and experience in implementation. One barrier to their greater use is the requirement to undertake simulation studies at the planning stage that require considerable knowledge, coding experience and additional costs. Based on some modest assumptions about the likely patterns of recruitment and the covariance structure of the outcomes, some simple analytic expressions are presented that negate the need to undertake simulations.

Methods A model for longitudinal outcomes with an assumed approximate multivariate normal distribution and three contrasting simple recruitment models are described, based on fixed, increasing and decreasing rates. For assumed uniform and exponential correlation models, analytic expressions for the variance of the treatment effect and the effects of the early outcomes on reducing this variance at the primary outcome time-point are presented. Expressions for the minimum and maximum values show how the correlations and timing of the early outcomes affect design efficiency.

Results Simulations showed how patterns of information accrual varied between correlation and recruitment models, and consequentially to some general guidance for planning a trial. Using a previously reported group sequential trial as an exemplar, it is shown how the analytic expressions given here could have been used as a quick and flexible planning tool, avoiding the need for extensive simulation studies based on individual participant data.

Conclusions The analytic expressions described can be routinely used at the planning stage of a putative trial, based on some modest assumptions about the likely number of outcomes and when they might occur and the expected recruitment patterns. Numerical simulations showed that these models behaved sensibly and allowed a range of design options to be explored in a way that would have been difficult and time-consuming if the previously described method of simulating individual trial participant data had been used.

Keywords Group sequential designs, Early outcomes, Randomized controlled trials, Planning

*Correspondence:
Nick R. Parsons
nick.parsons@warwick.ac.uk
[1] Warwick Clinical Trials Unit (WCTU), Warwick Medical School, University of Warwick, CV4 7AL Coventry, UK

Background

Group sequential designs (GSD) are one of the most widely used methodologies for adaptive design in randomized clinical trials [1]. In GSD researchers collect data and undertake sequential analyses with the opportunity to either reject the null hypothesis, stop the study

Parsons *et al. BMC Medical Research Methodology*   (2024) 24:42

Page 2 of 17

for futility or continue recruitment at an interim look, before reaching the planned sample size [2]. Despite the self evident gains in efficiency that GSD and other adaptive designs offer due to the possibility of stopping early, the perception in much of the statistical community is that they are still underused and where they are used they are used only within niche therapeutic areas where there is expertise and experience in implementation (e.g. in pharmaceutical trials testing drugs in oncology) [1]. There has been much discussion of the reasons why this is the case and how the barriers to uptake might be overcome, in particular the lack of knowledge, experience, statistical expertise and opportunity in the clinical trials community, outside of specialist teams [3]. A recent publication showed that GSD are feasible and are likely to be considerably more efficient than fixed sample size designs for pragmatic clinical trials, an application area where adaptive designs are generally never used [4]. Pragmatic trials typically test complex interventions (e.g. surgery, exercise, cognitive behavioural therapy) in routine clinical practice and are characterised by relatively large sample sizes and long follow-up periods [5, 6]. In such settings, GSD that use data from not only the final (primary) study outcome but also from early outcomes at interim analyses to inform stopping decisions have particular attraction due in large part to the use of patient-reported outcome measures (PROMs) that show strong associations between early and final outcomes [4]. This approach is exemplified by the START:REACTS trial that used this methodology to assess a novel intervention for repair of rotator cuff tendon tears [7]. The initial design and planning of this study which was based on simulating individual trial participant data, from a multivariate distribution [8], under an assumed model for study recruitment patterns [9] in order to assess likely information accumulation during a proposed trial, is a very general and highly effective method. However, such simulations are complex and time-consuming to set-up and implement and therefore provide an additional barrier, amongst many others previously identified [3], to the widespread use of GSD, particularly for trialists and those statisticians who are not specialists in this area. If we are willing to make some modest assumptions about the distribution of the outcomes, the likely correlation structure and recruitment patterns we might expect, then we can derive relatively simple analytic expressions for information accrual during a trial. This would allow us to explore a range of options for the timing and number of interim analyses, in a routine way without the need for simulating individual participant data, and as such make the methods much more accessible to potential non-expert users. In order to do this, we propose a number of recruitment models and two contrasting correlation models for the

temporal sequence of outcomes observed for individual study participants. The recruitment and correlation models together provide expressions for the variance of the treatment effect estimate and a natural means to distinguish and make explicit the contribution to the information fraction of the early and primary outcome data at an interim analysis. Previous work has discussed the timing of follow-up measurements for a single early outcome, using a simple linear model for the decay in the correlation between the final and early outcomes over time [10]. The models we develop here allow us to explore this issue in the general case of more than one early endpoint using an information adaptive group sequential approach to improve decision making at interim analyses. More generally, others have also suggested using information from prognostic baseline covariates (e.g. from baseline scores, comorbidities and patient demographics) in addition to early outcomes to inform interim decision making [11, 12]. Our focus is on stopping for treatment efficacy or futility, therefore we do not consider other adaptions that might be made to the trial design (e.g. sample size re-estimation) or more generally issues around inference and how to obtain unbiased estimates of treatment effects for group-sequential trials that stop early [13–15]. Also, given the overwhelming predominance of continuous outcomes in pragmatic trials of complex interventions, we will not discuss binary or time-to-event outcomes. Although the motivation for the work is from our own experiences with pragmatic trials, the methodological approaches described here are applicable much more widely to GSD in any application area where the issues and design characteristics we highlight are important.

We structure the paper as follows. In "Longitudinal outcomes" section, we describe a model for longitudinal outcomes with an assumed approximate multivariate normal distribution. "Recruitment and follow-up models" section develops three contrasting simple models for recruitment of participants into a clinical trial, and "Correlation models" section develops the models from "Longitudinal outcomes" and "Recruitment and follow-up models" sections for a uniform and an exponential correlation model. "Numerical examples" section provides some numerical examples to illustrate the models. The paper concludes in "Discussion" section with a discussion, including details of the availability of software for implementation of the methods described.

Longitudinal outcomes
A group sequential trial
Consider a two-arm randomized controlled trial where participants are randomized to either a treatment or a control arm, followed-up, assessed and the primary outcome observed at a sequence of s occasions at time-points

$d_1, \ldots, d_s$, ordered such that $d_s > \cdots > d_1$. In such a setting, the primary interest of the trial is often to estimate the effect of the treatment on the study outcome at time-point d_s, the primary or final study outcome time-point. At some time t during the study, the total number of participants with data at follow-up occasion r ($r = 1, \ldots, s$) is $N0_r + N1_r$, where $N0_r$ is the number in the control arm and $N1_r$ is the number in the treatment arm. Due to the ordering of the follow-up occasions, prior to the completion of trial follow-up, assuming data are complete, the number of participants with outcome data are structured such that $N0_1 \geq N0_2 \geq \cdots \geq N0_{s-1} \geq N0_s$ and $N1_1 \geq N1_2 \geq \cdots \geq N1_{s-1} \geq N1_s$. For instance, if the primary study outcome time-point is at 12 months after recruitment, with early outcomes at 3 and 6 months, then at all times prior to completion of follow-up we would expect to have more 3 month data than 6 month data, and more 3 and 6 month data than 12 month data.

If the full study sample of N participants is recruited in a period of time of length T_R (the recruitment period) and the primary outcome is observed at time d_s (after recruitment) then study follow-up is complete, and the trial ends, at time $d_s + T_R$. Importantly in this setting, there is a period of time between primary outcome data being available for analysis and the end of recruitment ($d_s < t < T_R$). During this so-called *window of opportunity*, there is the possibility of undertaking interim analyses, potentially stopping the study early for either treatment futility or efficacy. In such settings, if the interim analyses use final outcome data only, then the opportunities for stopping are likely to be extremely limited as trial recruitment will often have been completed before there is sufficient final outcome data available for informed stopping decisions to be made [4, 8]. However, if the early outcomes for trial participants (at occasions $d_r; r = 1, \ldots, s-1$) are correlated with their final outcomes (at d_s), then a group sequential analysis [2] which uses information from both the early and final outcomes to estimate the treatment effect at d_s is likely to lead to considerable increases in statistical power and also to make early stopping feasible [10, 16].

A number of authors have investigated this problem [8, 10, 17] and more generally the use of group-sequential analysis for longitudinal data [18, 19]. In the most simple possible setting, for instance the double-regression method described by Engel and Walstra [17], there is a final (long-term) and a single early (short-term or concomitant) endpoint that are correlated for individuals at the two time points. The main motivation for using information from the early outcomes in addition to the final outcomes in a clinical setting is that it allows us to conduct the trial in a more efficient manner by potentially reaching a conclusive result more quickly and

limiting patient exposure to ineffective or unsafe treatments, if the study ultimately provides little support for the efficacy of the intervention under test. Stallard [16], for instance, showed that using early outcome data, in the setting of a seamless phase II/III clinical trial with treatment selection, results in an increase in statistical power when data are correlated with the final outcome. A general approach in the setting of a sequential clinical trial, with a number of interim analyses, with a single long-term and potentially many short-term endpoints for a two-arm trial was first suggested by Galbraith and Marschner [10] and discussed further by Parsons et al. for a clinical trial in shoulder surgery [8], including extensive simulations for a prospective sample size calculation, and for surgical trials in general [4]. These authors rely in all cases on the independent increments argument, based on an asymptotic joint multivariate normal distribution for the sequential test statistics, for construction of valid group sequential designs for the longitudinal models used, e.g. linear mixed-effects and generalized least squares models [2, 20]. Due to the nature of the applications described, the focus here is purely on using early outcomes only to inform decision making. More generally, others have described approaches in settings where baseline (prognostic) covariates are available in addition to or in preference to early outcomes [12, 21].

Data model

Let y_{ijr} be the outcome for the ith of N participants ($i = 1, \ldots, N$), at follow-up occasion r ($r = 1, \ldots, s$) recruited into intervention arm j (0 = control and 1 = treatment) of the group sequential trial. We assume hereafter independence between the trial participants and that the distribution of outcomes ($y_{ij1}, \ldots, y_{ijs}$) is multivariate normal, with mean ($\mu_{j1}, \ldots, \mu_{js}$) and $s \times s$ covariance matrix

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{12} & \ldots & \sigma_1\sigma_s\rho_{1s} \\ \sigma_2\sigma_1\rho_{21} & \sigma_2^2 & \ldots & \sigma_2\sigma_s\rho_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_s\sigma_1\rho_{s1} & \sigma_s\sigma_2\rho_{s2} & \ldots & \sigma_s^2 \end{pmatrix}, \quad (1)$$

where σ_r is the standard deviation of the outcome at occasion r and $\rho_{rr'}$ is the correlation between endpoints at occasions $r = 1, \ldots, s$ and $r' = 1, \ldots, s$. Noting also that Σ can be expressed as $\Sigma = S^{1/2}RS^{1/2}$, for correlation matrix R and (diagonal) variance matrix S.

Expressing as a linear longitudinal model with correlated errors, under the assumption of multivariate normality (MVN), the vector of outcomes $\mathbf{y}_i$, for participant i, has distribution $\mathbf{y}_i \sim \text{MVN}(X_i\beta, \Sigma_i)$, where Σ_i is the $r \times r$ covariance matrix of $\mathbf{y}_i$, for the r observed outcomes for participant i, characterised by covariance

parameters σ_r and $\rho_{rr'}$. X_i is a $r \times 2s$ design matrix and $\boldsymbol{\beta}$ is a $2s \times 1$ vector of unknown model parameters, where for inferential purposes the most important is β_s the effect of the treatment on the study outcome at time-point d_s, the primary study endpoint.

The maximum likelihood estimator for $\boldsymbol{\beta}$, under the multivariate normal assumption, for known Σ, is the generalized least squares estimator [22]

$$\boldsymbol{\beta} = \left(\sum_{i=1}^{N} X_i' \Sigma_i^{-1} X_i \right)^{-1} \left(\sum_{i=1}^{N} X_i' \Sigma_i^{-1} y_i \right), \qquad (2)$$

with variance given by

$$\mathrm{var}(\boldsymbol{\beta}) = \left(\sum_{i=1}^{N} X_i' \Sigma_i^{-1} X_i \right)^{-1}. \qquad (3)$$

Estimates of model parameters $\boldsymbol{\beta}$ and their variances $\mathrm{var}(\boldsymbol{\beta})$, and consequently information, follow naturally given Σ, which is obtained from estimates of $\boldsymbol{\rho}$ and $\boldsymbol{\sigma}$. The covariance parameters could, in principle, be fixed to known or expected values but are generally estimated from accumulating data as a trial progresses. For instance, Galbraith and Marschner [10] use mixed-effects models for analysis of correlated data to estimate $\boldsymbol{\rho}$ and $\boldsymbol{\sigma}$. In practice this can be implemented, for example, by fitting separate fixed-effects for each study outcome time d_r with an unstructured error covariance using the function `lme` in R [23] package `nlme`. However, for practical reasons during trial planning and monitoring we prefer to use the generalized least squares model function `gls` in R package `nlme`, which unlike the mixed-effects model, provides explicit estimates of the covariance parameters [24]. Either the mixed-effects or generalized least squares formulation provides consistent and unbiased estimates of model parameters [4], under an assumed multivariate normal distribution with a general covariance structure, common follow-up times for each individual and missing outcomes that are assumed to be a consequence of the shortened follow-up duration.

Trial planning and monitoring
The primary interest of the clinical trial is to estimate β_s and its variance $\mathrm{var}(\beta_s)$. Easily interpretable explicit expression for $\mathrm{var}(\beta_s)$ do not exist for general s, and general covariance matrix Σ. However, expressions for $\mathrm{var}(\beta_s)$ can be obtained directly for the most simple cases, under the structured data assumptions of "Longitudinal outcomes" section, where there are one ($s = 2$) and two ($s = 3$) early outcomes [4]. For instance, for the simplest possible case $s = 2$,

$$\mathrm{var}(\beta_2) = \sigma_2^2 \left[\frac{(N0_2 + N1_2)(1 - \rho_{12}^2)}{N0_2 N1_2} + \frac{(N0_1 + N1_1)\rho_{12}^2}{N0_1 N1_1} \right]. \qquad (4)$$

Of particular practical importance when planning an information adaptive group sequential study is to understand how information on the treatment effect at a time t, $\mathrm{I}(t) = 1/\mathrm{var}(\beta_s(t))$, is likely to accumulate during recruitment and follow-up. Typically pre-set expected information thresholds are used to trigger interim analyses, and to construct lower and upper stopping boundaries at the interim analyses, with stopping decisions being made based on estimates of β_s and $\mathrm{var}(\beta_s)$ [8]. Clearly, the information at some time t during recruitment depends on the covariance parameters $\rho = \rho_{12}, \ldots, \rho_{rr'}$ and σ_s, and the number of participants ($N0_r$ and $N1_r$) with data at each follow-up occasion r.

In order to plan how a trial might be implemented and if, and when, interim analyses should take place, we need to understand how information is likely to accumulate as the study proceeds. To do this we need to make some *a priori* assumptions about both the expected patterns of recruitment and the correlation structure between the early and final outcomes. In the most general settings we might imagine, with complex patterns of recruitment and accrual of study data and unstructured correlations between outcomes, simulation methods may be the only way to proceed at the planning stage [8]. Such an approach is hard to implement, time-consuming and often provides little or no insight into the general principles at play and how these might guide us when we make future modifications to the design or when planning future studies. However, if we are willing to make some reasonable assumptions about the likely patterns of recruitment and the structure of the correlations then we can obtain explicit analytic expressions for $\mathrm{var}(\beta_s)$ much more quickly and simply, and use these as a means to plan the study.

When planning a trial we assume that a fixed number of early outcomes are available throughout the study (for instance, a primary outcome at 12 months, with two early outcomes at 3 and 6 months) and that, in principle, the timings of the early outcomes could be changed (for instance, to 4 and 8 months). Typically, the exact correlations between early and primary outcomes are unknown. However, we can speculate on the likely correlation structure as a means to understand how information might be accumulated as follow-up proceeds. Two widely used correlation models for longitudinal data are described in "Correlation models" section. At the design stage, for some arbitrarily selected time-point t during recruitment we will generally not know the exact number of participants recruited or the number of participants ($N0_r$ and $N1_r$) with data at each follow-up occasion r. We discuss simple models for predicting recruitment in "Recruitment and

follow-up models" section. Given correlation and recruitment models, together with an estimate of σ_s (e.g. from previously reported studies or pilot data), we can predict how var(β_s), and therefore information, will vary during study follow-up and use this to motivate our choice of the number and timings of early outcome assessments and interim analyses.

When monitoring a study, often more important than the information itself is the information fraction or information time $\tau(t)$ at an interim analysis at time t, defined by $\tau(t) = I(t)/I$, where $I(t)$ and I are the information levels at time t and the study end, respectively [25]. Knowing the information fraction $\tau(t)$ allows us to determine lower and upper boundaries (for planned futility and efficacy stopping) and boundary crossing probabilities at an interim analysis, for some given boundary crossing probabilities under the null hypothesis, based on canonical joint distribution properties for group sequential trials [2]. Boundaries and probabilities can, for instance, be calculated using appropriate functions from the gsDesign package in R [26]. An example of how this might be implemented in practice is provided in "Numerical examples" section, using as an exemplar the START:REACTS study of sub-acromial spacer for tears affecting rotator cuff tendons [7, 27].

Recruitment and follow-up models

Assuming the data are structured as in "A group sequential trial" section and are complete, consistent with what we would do during planning and sample size calculations in a conventional trial design based on a single primary endpoint. We can write a general expression for the number of participants providing outcome data from follow-up occasion r at time t as $N_r(t, d_r) = kg_r(t, d_r)$, where k is a constant depending on the planned sample size N and recruitment period T_R only and $g_r(t, d_r)$ is some function of t and the follow-up time point d_r, measured in the same units as t. For notational convenience, we define $r = 0$ to be the recruitment occasion and thus $g_0(t, d_0)$ is the result of the function $g_r(t, d_r)$ when $r = 0$, that is at the time-point when recruitment occurs at $d_r = 0$, such that $N_0(t, d_0) = kg_0(t, d_0)$ is the number of participants recruited at time t. For $d_r < t \leq d_r + T_R$, the number of participants is $N_r(t, d_r) = kg_r(t, d_r)$ and at $t < d_r$ prior to outcome data becoming available is $N_r(t, d_r) = 0$ and at $t > d_r + T_R$ when data collection has been completed for outcome r is $N_r(t, d_r) = N$. We also note that $n_{rr'}(t) = N_r(t, d_r)/N_{r'}(t, d_{r'})$, the ratio of the number of study participants with outcome data from follow-up occasion d_r to study participants with outcome data from follow-up occasion $d_{r'}$ at time t, is equal to $g_r(t, d_r)/g_{r'}(t, d_{r'})$. Introducing a weight $0 < \phi < 1$ that allows for unequal group sizes, gives intervention group sizes of $N0_r(t, d_r) = \phi N_r(t, d_r)$ and $N1_r(t) = (1 - \phi)N_r(t, d_r)$.

Fixed rate

In the simplest possible situation, setting $k = N/T_R$ and $g_r(t, d_r) = (t - d_r)$ in the expression $N_r(t, d_r) = kg_r(t, d_r)$ leads to a model with a fixed rate of recruitment (λ_f) and follow-up data accrual, where $\lambda_f = N/T_R$ participants are recruited into the study for each of t study days, if T_R is measured in days. The total number of participants recruited into the study at time t for the fixed model is given by $N_0(t) = Nt/T_R$; Fig. 1a shows total recruitment and follow-up data accrual curves for this model.

Linearly increasing rate

Setting $k = N/\{T_R(T_R + 1)\}$ and $g_r(t, d_r) = (t - d_r)((t - d_r) + 1)$ in the expression $N_r(t, d_r) = kg_r(t, d_r)$ leads to a model with an increasing rate of recruitment given by $\lambda_i(t) = 2Nt/\{T_R(T_R + 1)\}$. In this model the mean rate of recruitment across the whole recruitment period for this model is N/T_R, the same as the fixed rate parameter model λ_f, with the starting rate (at $t = 1$) given by $\lambda_i(1) = 2N/T_R(T_R + 1)$ and the end rate (at $t = T_R$) by $\lambda_i(T_R) = 2N/(T_R + 1)$, noting that $\lambda_i(1) < \lambda_f$ and $\lambda_i(T_R) > \lambda_f$. The total number of participants recruited into the study at time t for the increasing rate model is given by $N_0(t) = Nt(t + 1)/\{T_R(T_R + 1)\}$; Fig. 1b shows total recruitment and follow-up data accrual curves for this model.

Linearly decreasing rate

Formulating deliberately as a contrast to the model of "Linearly increasing rate" section, setting $k = N/\{T_R(T_R + 1)\}$ and $g_r(t, d_r) = (t - d_r)(2T_R - (t - d_r) + 1)$ leads to a model with a decreasing rate of recruitment given by $\lambda_d(t) = 2N(T_R - t + 1)/\{T_R(T_R + 1)\}$. In this model the mean rate of recruitment across the whole recruitment period for this model is also N/T_R, the same as the fixed rate parameter model λ_f, with the starting rate (at $t = 1$) given by $\lambda_d(1) = 2N/\{(T_R + 1)\}$ and the end rate (at $t = T_R$) by $\lambda_d(T_R) = 2N/\{T_R(T_R + 1)\}$, noting that, in a reversing of the relationship for the increasing rate model, $\lambda_d(1) > \lambda_f$ and $\lambda_d(T_R) < \lambda_f$. The total number of participants recruited at time t for the decreasing rate model is given by $N_0(t) = Nt(2T_R - t + 1)/\{T_R(T_R + 1)\}$; Fig. 1c shows total recruitment and follow-up data accrual curves for this model.

Correlation models

We consider two common single parameter correlation models; the uniform and exponential models [22], the latter is also known as the first-order autoregressive (AR1) model. These models offer contrasting views on the likely correlation structure between early and final outcomes. We choose to use α and γ for the parameters for the uniform and exponential models, in the following
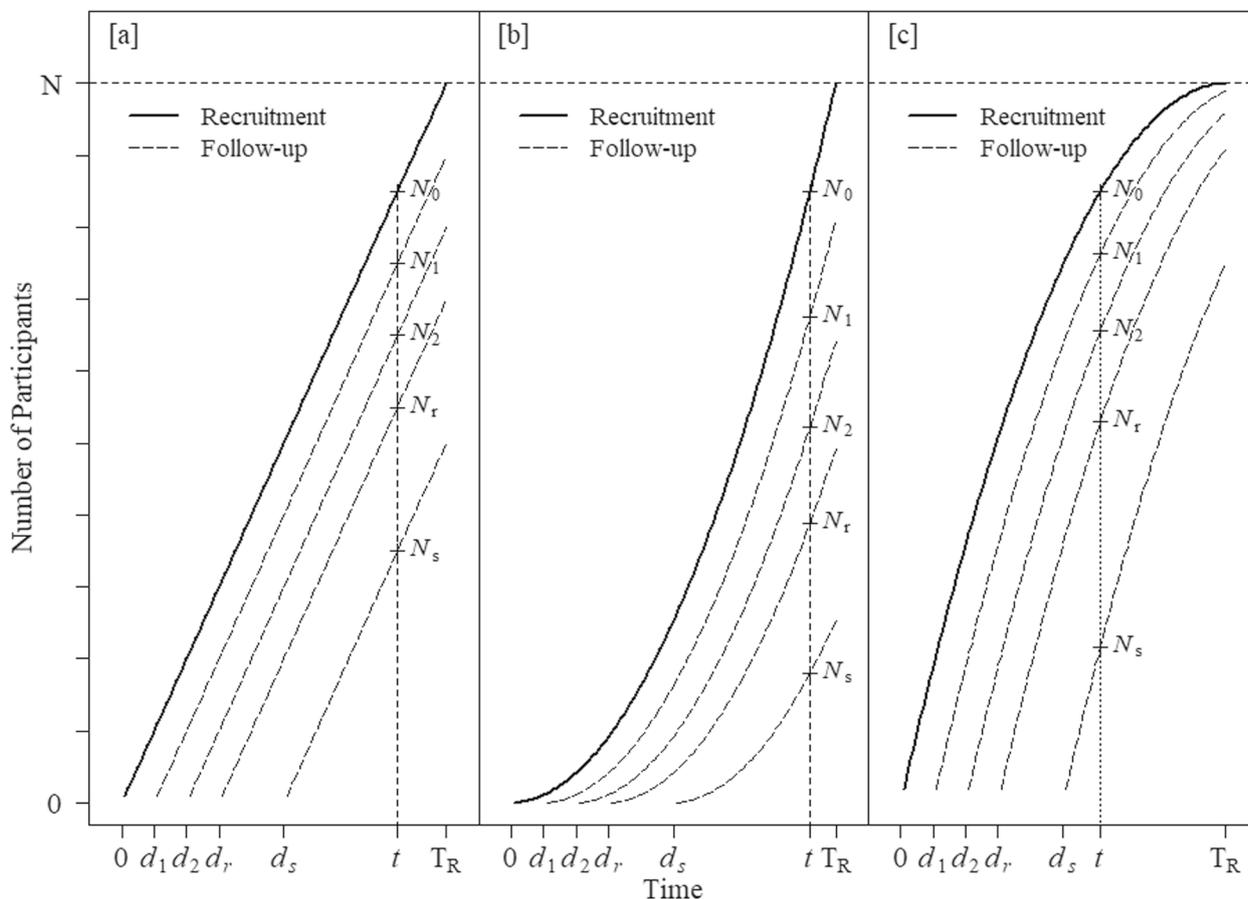
Fig. 1 Total recruitment and follow-up accrual curves for the primary and all early follow-up endpoints at times $d_1, d_2, \ldots, d_r, \ldots, d_s$ for (**a**) the fixed rate recruitment model, **b** the linearly increasing rate recruitment model and (**c**) the linearly decreasing rate recruitment model. With annotation showing numbers of participants recruited N_0 and with follow-up data at $N_1, N_2, \ldots, N_r, \ldots, N_s$ at times t, set such that N_0 is the same for each setting

descriptions, to reflect the fact that they have quite different interpretations.

Uniform

The uniform correlation model is the natural basis for a random-effects model, which we can motivate in our setting by thinking of the covariance structure of the data as a consequence of random variation amongst (unobserved) subject-specific characteristics of participants in a clinical trial. The uniform correlation model is widely seen in trials using PROMs, where participants are asked to assess

on a unit (participant in a trial) at time-points $r = 1, \ldots, s$ and $r' = 1, \ldots, s$ given by $\rho_{rr'} = \alpha$ when $r \neq r'$ and $\rho_{rr'} = 1$ when $r = r'$.

Expression for **var**(β_s)

For the uniform correlation model, assuming that the number of participants with outcome data are structured in the manner described in "Longitudinal outcomes" and "Recruitment and follow-up models" sections, the variance of the treatment effect on the study outcome at time-point s (the primary study endpoint) is given by

$$\text{var}(\beta_s^{\text{unif}}) = \sigma_s^2 \left[\frac{N0_1 + N1_1}{N0_1 N1_1} + \sum_{m=1}^{s-1} \frac{\det(R_{m+1})}{\det(R_m)} \left(\frac{N0_{m+1} + N1_{m+1}}{N0_{m+1} N1_{m+1}} - \frac{N0_m + N1_m}{N0_m N1_m} \right) \right],$$

their own status or functional abilities [4]. It assumes the correlations between measurements are constant regardless of how far apart in time they are, with measurements

where $\det(R_m) = (1 - \alpha)^{m-1}(1 + (m - 1)\alpha)$ is the determinant of the $m \times m$ correlation matrix R_m (see Appendix A1 for details). Therefore we can also write as follows;

$$\mathrm{var}(\beta_s^{\mathrm{unif}}) = \sigma_s^2 \left[\frac{N0_1 + N1_1}{N0_1 N1_1} + \right.$$
$$\left. \sum_{m=1}^{s-1} \frac{(1-\alpha)(1+m\alpha)}{(1+(m-1)\alpha)} \left(\frac{N0_{m+1} + N1_{m+1}}{N0_{m+1} N1_{m+1}} - \frac{N0_m + N1_m}{N0_m N1_m} \right) \right]. \tag{5}$$

For the simplest possible case where $s = 2$, $\det(R_1) = 1$ and $\det(R_2) = (1-\alpha)(1+\alpha) = 1 - \alpha^2$ then, as we would expect, expressions (4) and (5) are equal (i.e. $\mathrm{var}(\beta_2) = \mathrm{var}(\beta_2^{\mathrm{unif}})$), when $\alpha = \rho_{12}$. At the extremes of α, we note that when $\alpha = 1$, then $\mathrm{var}(\beta_s^{\mathrm{unif}}) = \sigma_s^2 (N0_1 + N1_1)/N0_1 N1_1$ and we conclude that data (information) from the early outcome (at $s = 1$) only are important and when $\alpha = 0$, then $\mathrm{var}(\beta_s^{\mathrm{unif}}) = \sigma_s^2 (N0_s + N1_s)/N0_s N1_s$ and data at earlier times provide no information.

Recruitment models and var(β_s)
Substituting the expressions for $N0_r$ and $N1_r$, from "Recruitment and follow-up models" section, into Eq. (5) gives the following expression for the variance of the treatment effect on the primary outcome at time t, where we require $d_s < t \leq T_R$;

$$\mathrm{var}(\beta_s^{\mathrm{unif}}(t)) = \frac{\sigma_s^2}{k\phi(1-\phi)} \left[\frac{1}{g_1(t,d_1)} + \right.$$
$$\left. \sum_{m=1}^{s-1} \frac{(1-\alpha)(1+m\alpha)}{(1+(m-1)\alpha)} \left(\frac{1}{g_{m+1}(t,d_{m+1})} - \frac{1}{g_m(t,d_m)} \right) \right]. \tag{6}$$

When the correlation is zero ($\alpha = 0$) between early and primary outcomes, then $\det(R_{m+1}) = \det(R_m) = 1$ for all m and noting that

$$\sum_{m=1}^{s-1} \left(\frac{1}{g_{m+1}(t,d_{m+1})} - \frac{1}{g_m(t,d_m)} \right) = \frac{1}{g_s(t,d_s)} - \frac{1}{g_1(t,d_1)},$$

then the variance when there is no correlation, $\mathrm{var}(\beta_s^0)$, is given by

$$\mathrm{var}(\beta_s^0(t)) = \frac{\sigma_s^2}{k\phi(1-\phi)g_s(t,d_s)}, \tag{7}$$

where we note that $kg_s(t,d_s)$ is equal to N_s the number of study participants with primary outcome data at time t. We can construct a measure $V_s(t)$ to be the relative effect of the early outcomes, due to their correlation with the primary outcome, on reducing the variance of the primary outcome by dividing $\mathrm{var}(\beta_s(t))$ by $\mathrm{var}(\beta_s^0(t))$. For the uniform model this is

$$V_s^{\mathrm{unif}}(t) = n_{s1}(t) + \sum_{m=1}^{s-1} \frac{(1-\alpha)(1+m\alpha)}{(1+(m-1)\alpha)} (n_{s(m+1)}(t) - n_{sm}(t)). \tag{8}$$

Where, $n_{s1}(t) \leq V_s^{\mathrm{unif}}(t) \leq 1$, with the lower constraint (giving the maximum possible benefit from the early outcomes) occurring when $\alpha = 1$ and information on the primary outcome (at d_s) comes entirely from the first early outcome (at d_1) and the upper constraint occurring when $\alpha = 0$ and there is no information from any of the early outcomes. More generally, for values of the correlation parameter between these limits, then $V_s^{\mathrm{unif}}(t)$ varies as a function of both t and d_r ($r = 1, \ldots, s$), and their relative spacings.

Minimum and maximum of $V_s^{\mathrm{unif}}(t)$
For a given value of α, the minimum of $V_s^{\mathrm{unif}}(t)$ occurs trivially for the uniform correlation model when we maximize the data available for each intermediate early outcome by moving them all towards the earliest outcome at d_1. In this setting $d_m \to d_1$ and functions $g_m(t,d_m) \to g_1(t,d_1)$ and consequently $n_{sm}(t) \to n_{s1}(t)$ for all $m = 2, \ldots, s-1$ and then from expression (8), noting that $n_{ss} = 1$, the minimum of $V_s^{\mathrm{unif}}(t)$ is given by

$$\min_{d_2,\ldots,d_{s-1}} \left(V_s^{\mathrm{unif}}(t) \right) = n_{s1}(t) + \frac{(1-\alpha)(1+(s-1)\alpha)}{(1+(s-2)\alpha)} (1 - n_{s1}(t)). \tag{9}$$

The maximum occurs when all the intermediate early outcomes are moved towards the final outcome at d_s. In this setting $d_m \to d_s$ and functions $g_m(t,d_m) \to g_s(t,d_1)$ and consequently $n_{sm}(t) \to 1$ for all $m = 2, \ldots, s-1$ and then from expression (8) the maximum of $V_s^{\mathrm{unif}}(t)$ is given by

$$\max_{d_2,\ldots,d_{s-1}} \left(V_s^{\mathrm{unif}}(t) \right) = n_{s1}(t) + (1-\alpha^2)(1 - n_{s1}(t)). \tag{10}$$

The terms $n_{rr'}$ in expression (8) represent the effects of the changing sample size, with different recruitment models, and as such are independent of the correlation parameter α. The effect of the correlation α on $V_s^{\text{unif}}(t)$ is fixed and independent of the spacing or differences between the early endpoints.

Exponential

The exponential model, in contrast to the uniform model, assumes that the correlation between pairs of measurements on the same subject decays to zero as the time separation between them increases. This model is widely used for longitudinal outcomes [22] and evidence from our own work suggests that it is a useful working assumption for modelling the association between serial measurements of PROMs for many large pragmatic clinical trials [4]. In the exponential model the correlation between a pair of measurements on a unit (participant in a trial) at time-points r and r' tends towards zero as the time between measurements increases $\rho_{rr'} = \gamma^{|d_r - d_{r'}|}$ [22]. Where the d_r are increasing ordered times that indi-

In this setting, without loss of generality, we can set the timings of the follow-up assessments d_r such that in all settings $d_1 = 1$ and $d_s = s$ such that, for instance, if $d_1 = 1, d_2 = 3, d_3 = 7/2$ and $d_4 = 4$ then this might represent assessments at 1, 3, 3.5 and 4 years or 4, 12, 14 and 16 months, depending on whether the base unit of time is 1 year or 4 months. Although, clearly the correlation parameter γ will generally differ depending on whether we are considering the former or latter settings. In the most general case, similar arguments can be applied if we wish to make follow-up assessments such that d_s is not a multiple of d_1. For instance, if assessments are planned at 4, 12, 18 and 22 months, then setting $d_1 = 1, d_2 = 7/3, d_3 = 10/3$ and $d_4 = 4$ ensures that, as we would expect given the relative distances, correlations between 18 and 22 month assessments $\gamma^{(d_4 - d_3)} = \gamma^{2/3}$ are the square of those between 4 and 12 month assessments $\gamma^{(d_2 - d_1)} = \gamma^{4/3}$, for a given value of γ. If the assessments are equally spaced in our model (i.e. when $d_r = r$ for $r = 1, \ldots, s$) and $(d_s - d_{s-1}) = \cdots = (d_2 - d_1) = 1$, $(d_s - d_{s-2}) = \cdots = (d_3 - d_1) = 2, \cdots, (d_s - d_1) = s - 1$, then

$$
\text{var}(\beta_s^{\text{exp}}) = \sigma_s^2 \left[\frac{(N0_s + N1_s)(1 - \gamma^2)}{N0_s N1_s} + \right.
$$
$$
\left. (1 - \gamma^2) \sum_{m=1}^{s-2} \frac{(N0_{s-m} + N1_{s-m})\gamma^{2m}}{N0_{s-m}N1_{s-m}} + \frac{(N0_1 + N1_1)\gamma^{2(s-1)}}{N0_1 N1_1} \right].
$$

(12)

cates the relative times of assessment. The parameter γ expresses the strength of association, for unit separation (i.e. where $|d_r - d_{r'}| = 1$), and for the applications discussed here for ease of interpretation is such that $0 \leq \gamma < 1$.

Expression for var(β_s)

For an assumed exponential model and a known, or an assumed, value of γ, then for data structured in the manner described in "Longitudinal outcomes" and "Recruitment and follow-up models" sections, var(β_s^{exp}) is given after some algebraic manipulation (see Appendix A2 for details), for $s \geq 3$, by

$$
\text{var}(\beta_s^{\text{exp}}) = \sigma_s^2 \left[\frac{(N0_s + N1_s)\{1 - \gamma^{2(d_s - d_{s-1})}\}}{N0_s N1_s} + \right.
$$
$$
\sum_{m=1}^{s-2} \frac{(N0_{s-m} + N1_{s-m})\{1 - \gamma^{2(d_{s-m} - d_{s-m-1})}\}\gamma^{2(d_s - d_{s-m})}}{N0_{s-m}N1_{s-m}} +
$$
$$
\left. \frac{(N0_1 + N1_1)\gamma^{2(d_s - d_1)}}{N0_1 N1_1} \right].
$$

(11)

For the case of a single early and a final outcome, then var(β_2^{exp}) is given, by dropping the middle term in the square brackets and setting $s = 2$ in which case, as we might expect, var$(\beta_2) = \text{var}(\beta_2^{\text{unif}}) = \text{var}(\beta_2^{\text{exp}})$, if $\alpha = \gamma$.

Recruitment models and var(β_s)

Substituting the expressions for $N0_r$ and $N1_r$, from "Recruitment and follow-up models" section, into Eq. (11) gives the following expression for the variance of the treatment effect on the primary outcome at time t;

$$
\text{var}(\beta_s^{\text{exp}}(t)) = \frac{\sigma_s^2}{k\phi(1 - \phi)} \left[\frac{\{1 - \gamma^{2(d_s - d_{s-1})}\}}{g_s(t, d_s)} + \right.
$$
$$
\left. \sum_{m=1}^{s-2} \frac{\{1 - \gamma^{2(d_{s-m} - d_{s-m-1})}\}\gamma^{2(d_s - d_{s-m})}}{g_{s-m}(t, d_{s-m})} + \frac{\gamma^{2(d_s - d_1)}}{g_1(t, d_1)} \right].
$$

(13)

Noting that the variance when there is no correlation ($\gamma = 0$) between early and primary outcomes var$(\beta_s^0(t))$, is given by expression (7), then the effect of the correlation, due to the early outcomes, on reducing the variance of the primary outcome for the exponential model is given by

$$V_s^{\exp}(t) = 1 - \gamma^{2(d_s - d_{s-1})} +$$
$$\sum_{m=1}^{s-2} n_{s(s-m)}(t)(1 - \gamma^{2(d_{s-m} - d_{s-m-1})})\gamma^{2(d_s - d_{s-m})} + n_{s1}(t)\gamma^{2(d_s - d_1)}. \tag{14}$$

Where $n_{s1}(t) \leq V_s^{\exp}(t) \leq 1$, with the lower constraint occurring when $\gamma = 1$ and information on the primary outcome (at d_s) comes entirely from the first early outcome (at d_1) and the upper constraint occurring when $\gamma = 0$ and there is no information from any of the early outcomes. More generally, for values of the correlation parameter between these limits, then $V_s^{\exp}(t)$ varies as a function of both t and d_r $(r = 1, \ldots, s)$, and their relative spacings.

Minimum and maximum of $V_s^{\exp}(t)$

The maximum of $V_s^{\exp}(t)$, for known γ, occurs when $d_{s-1} \to d_1$ and as a consequence all the intermediate terms move also towards d_1; i.e. $d_{s-1} \to d_{s-2} \to \cdots \to d_2 \to d_1$, is given by

$$\max_{d_2, \ldots, d_{s-1}} \left(V_s^{\exp}(t) \right) = n_{s1}(t) + (1 - \gamma^{2(d_s - d_1)})(1 - n_{s1}(t)). \tag{15}$$

In contrast to the uniform correlation model, there is no simple expression for the minimum of $V_s^{\exp}(t)$. As we move intermediate outcomes towards the earliest outcome $d_m \to d_1$ then functions $g_m(t, d_m) \to g_1(t, d_1)$ for all $m = 2, \ldots, s-1$ and we have more data available which, all other things being equal, will minimise $V_s^{\exp}(t)$. However, when we increase the amount of data available by moving intermediate outcomes towards the earliest outcome we also increase the distances $d_s - d_{s-m}$ which, from expression (14), clearly acts to increase $V_s^{\exp}(t)$ by making terms $\gamma^{2(d_s - d_{s-m})} \to 0$.

The settings of $d_2, \ldots, d_{s-1}$ that minimise $V_s^{\exp}(t)$, will vary with the correlation parameter γ and s. In general, minimums of $V_s^{\exp}(t)$ can be obtained numerically using linearly constrained optimization methods; e.g. using function constrOptim in R, with gradients set to be the derivatives $\partial V_s^{\exp}/\partial d_m$, which are relatively simple to calculate (see Appendix A3 for details) [23].

Information

The information fraction, when the correlation between early outcomes and the final outcome is zero ($\tau 0$) at time t, is given by the information at time t divided by the information at the study end which is $\tau 0(t) = \text{var}(\beta_s^0(t = d_s + T_R))/\text{var}(\beta_s^0(t))$ [25]; noting that at the study end $t = d_s + T_R$. From expression (7), this can be written more simply as $\tau 0(t) = N_s(t, d_s)/N$,

the proportion of participants with final outcome data at time t, or alternatively as $\tau 0(t) = g_s(t, d_s)/g_s(d_s + T_R, d_s)$.

From previously, the relative effect of the early outcomes, due to their correlation with the primary outcome, on reducing the variance of the primary outcome is $V_s(t) = \text{var}(\beta_s(t))/\text{var}(\beta_s^0(t))$. Therefore, the information fraction τ at time t, for the full longitudinal model including the contribution of the early outcomes, is given by

$$\tau(t) = \tau 0(t)/V_s(t). \tag{16}$$

This allows us to make explicit the distinction between information that comes directly from observation of the final outcome ($\tau 0$) and information that comes from the early outcomes (V_s) at time t. It also makes clear that $1/V_s(t)$ is the proportionate increase in the information fraction τ at time t due to the early outcomes. For instance, if $V_s(t) = 0.8$, then we have 1.25 times as much information at time t than we would have had if the early outcomes were uncorrelated with the primary outcome.

Numerical examples

Uniform correlation model

To understand the properties of the uniform correlation model in the setting described, we calculate V_s^{unif} for typical values of $s = 2, 3, 4, 5, 6$ for the recruitment models of "Recruitment and follow-up models" section. Without loss of generality, we set $d_s = 2$ and $d_1 = 1$ and arbitrarily set the recruitment period T_R to be a fixed multiple of d_s such that the information fraction $\tau 0(t_w)$ when $\alpha = 0$ (i.e. the proportion of participants with final outcome data) at three equally spaced interim analyses ($w = 1, 2, 3$), are $\tau 0(t_1) = 0.15$, $\tau 0(t_2) = 0.30$ and $\tau 0(t_3) = 0.45$, which we nominally refer to hereafter as *early*, *mid* and *late*. The actually timings of the interim analyses, relative to T_R, will depend on the selected recruitment model; see Appendix A4 for details. We do this to allow us to make simple comparisons between the recruitment models and values of s at each the interim analyses.

Plots showing the difference in minimum and maximum values and the empirical distribution of V_s^{unif} for equal group sizes ($\phi = 0.5$), with varying $d_2, \ldots, d_{s-1}$, for correlations in the range $0 \leq \alpha < 1$ and the decreasing, fixed and increasing rate recruitment models are available in the supplementary files (see Additional file 1, Figs. S1 to S9). The pattern of differences between recruitment models and interim analyses are consistent across values

of α, with values for V_s^{unif} decreasing monotonically with increasing α; the larger the correlation, the greater the information available from the early outcomes. Picking a typical value of $\alpha = 0.5$ for illustration purposes, Table 1 shows the effects of s, recruitment model and timing of interim analysis on V_s^{unif} and minimum and maximum values of V_s^{unif} and also, as a means of comparison the equal spacing model where $d_r = 1 + (r-1)/(s-1)$ for all $r = 1, \ldots, s$, which we denote by $\tilde{V}_s^{unif}$.

Table 1 and expressions (8), (9) and (10) show that variation in V_s^{unif}, for some fixed s, is due solely to differences in n_{s1}, the ratio of the number of study participants with final outcome data to study participants with first early outcome data at time t. The value of n_{s1} depends on both the timing of the interim analysis and the recruitment model. As $n_{s1} \to 1$, when all participants have final outcome data, then also $V_s^{unif} \to 1$ and there is no additional information available from the early outcomes. This is apparent in the trend for larger values in V_s^{unif}, for all settings, as we move from early to mid to late interim analyses and n_{s1} increases in value. As we increase the number of early outcomes for fixed α, with increasing s, then the term $D = (1-\alpha)(1 + (s-1)\alpha)/(1 + (s-2)\alpha)$ in expression (8) decreases in value towards a minimum of $(1 - \alpha)$ as $s \to \infty$. Values of D in Table 1, for $\alpha = 0.5$,

decrease with increasing s rapidly initially (e.g. from $s = 2$ to $s = 3$), but slower later (e.g. from $s = 5$ to $s = 6$). This suggests that there is little to be gained in increasing information fraction by increasing s much beyond the values we use here.

Exponential correlation model

The correlation between the primary outcome at d_s and the first outcome at d_1 is given by $\gamma^{d_s - d_1}$, and the parameter γ can be set such that the known or expected correlation between d_1 and d_s is ρ_{1s}. Therefore, as a means to produce consistency between uniform and exponential correlation models we set $\gamma = \alpha^{1/(d_s - d_1)}$, where α is the reference correlation from the uniform model. As $d_s = 2$ for all s, then $\gamma = \alpha$. Using these parametrisations for d_r and γ, we note that by replacing γ in expression (15) by α, $d_s = 2$ and $d_1 = 1$, results in $\max(V_s^{unif}) = \max(V_s^{exp})$.

Plots showing the difference in minimum and maximum values and the empirical distribution of V_s^{exp} for equal group sizes ($\phi = 0.5$), with varying $d_2, \ldots, d_{s-1}$, for correlations in the range $0 \le \gamma < 1$ and the decreasing, fixed and increasing rate recruitment models are available in the supplementary files (see Additional file 1, Figs. S10 to S18). The pattern of differences between recruitment models and interim analyses

Table 1 Minimum and maximum values of V_s^{unif} and values for the equal spacing model $\tilde{V}_s^{unif}$, for $\alpha = 0.5$ for early (t_1), mid (t_2) and late (t_3) interim analyses, where $\tau0_1 = 0.15$, $\tau0_2 = 0.30$ and $\tau0_3 = 0.45$ respectively, for s from two to six for the fixed, increasing and decreasing rate recruitment models. Also shown are values of $D = (1-\alpha)(1 + (s-1)\alpha)/(1 + (s-2)\alpha)$ and n_{s1} for each setting

s	D	Early (t_1) n_{s1}	$\tilde{V}_s^{unif}$	V_s^{unif} min - max	Mid (t_2) n_{s1}	$\tilde{V}_s^{unif}$	V_s^{unif} min - max	Late (t_3) n_{s1}	$\tilde{V}_s^{unif}$	V_s^{unif} min - max
[a] Fixed										
2	0.75	0.55	-	0.89-0.89	0.71	-	0.93-0.93	0.78	-	0.95-0.95
3	0.67	0.55	0.86	0.85-0.89	0.71	0.91	0.90-0.93	0.78	0.94	0.93-0.95
4	0.62	0.55	0.85	0.83-0.89	0.71	0.90	0.89-0.93	0.78	0.93	0.92-0.95
5	0.60	0.55	0.84	0.82-0.89	0.71	0.90	0.88-0.93	0.78	0.92	0.91-0.95
6	0.58	0.55	0.83	0.81-0.89	0.71	0.89	0.88-0.93	0.78	0.92	0.91-0.95
[b] Increasing										
2	0.75	0.59	-	0.90-0.90	0.68	-	0.92-0.92	0.72	-	0.93-0.93
3	0.67	0.59	0.88	0.86-0.90	0.68	0.90	0.89-0.92	0.72	0.92	0.91-0.93
4	0.62	0.59	0.86	0.84-0.90	0.68	0.89	0.88-0.92	0.72	0.91	0.90-0.93
5	0.60	0.59	0.85	0.83-0.90	0.68	0.89	0.87-0.92	0.72	0.90	0.89-0.93
6	0.58	0.59	0.85	0.83-0.90	0.68	0.88	0.87-0.92	0.72	0.90	0.88-0.93
[c] Decreasing										
2	0.75	0.42	-	0.86-0.86	0.62	-	0.91-0.91	0.74	-	0.93-0.93
3	0.67	0.42	0.82	0.81-0.86	0.62	0.88	0.87-0.91	0.74	0.92	0.91-0.93
4	0.62	0.42	0.80	0.78-0.86	0.62	0.87	0.86-0.91	0.74	0.91	0.90-0.93
5	0.60	0.42	0.79	0.77-0.86	0.62	0.86	0.85-0.91	0.74	0.91	0.89-0.93
6	0.58	0.42	0.78	0.76-0.86	0.62	0.86	0.84-0.91	0.74	0.90	0.89-0.93

are consistent across values of γ, with values for V_s^{exp} decreasing monotonically with increasing γ; the larger the correlation, the greater the information available from the early outcomes. Therefore, as for the uniform model, we pick a typical value of $\gamma = 0.5$ to illustrate in Table 2 the effects of s, recruitment model and timing of interim analysis on V_s^{exp}.

Table 2 shows minimum and maximum values of V_s^{exp} and also the equal spacing model. The values of $\tilde{V}_s^{exp}$ in Table 2 are always smaller than values of $\tilde{V}_s^{unif}$ in Table 1. This is due to the settings we adopt here to force $\max(V_s^{unif}) = \max(V_s^{exp})$, that make the correlations in the exponential model stronger. For instance, for the model where $s = 3$, the correlation between the early outcome at d_2 and the final outcome at d_3 is given by α in the uniform model and $\sqrt{\alpha}$ (i.e. $\alpha^{d_3-d_2} = \alpha^{2-1.5}$) in the exponential model. The observed decreases in V_s^{exp} as we move from $s = 2$ to $s = 6$, across all settings, are of a similar magnitude to those observed for the uniform model, suggesting some gains in information with increasing numbers of early outcomes, but with those gains diminishing as s increases. Of particular note for the exponential model is that, for the selected correlation $\gamma = 0.5$, $\tilde{V}_s^{exp}$ is quite close to $\min(V_s^{exp})$, across recruitment models and interim analyses. This suggests that for

a moderate correlation, having equally spaced outcomes is very close to the best possible design setting. As correlations become larger and γ approaches one, then clearly the spacing of the outcomes becomes unimportant. Conversely, as the correlation becomes smaller then the spacing of the outcomes has a larger relative impact on V_s^{exp}, for the recruitment models we explore here, and suggests that moving the early outcomes to be nearer to the final outcome provides more information.

START:REACTS clinical trial

The START:REACTS study was a double-blind, group-sequential, randomised controlled trial for rotator cuff tendon (shoulder) tears comparing arthroscopic debridement of the subacromial space with biceps tenotomy (control group) with the same procedure but including insertion of a sub-acromial spacer balloon (treatment group) [7, 27]. At the planning stage, individual participant data were simulated, for 10000 trials, and the models described in "Data model" section were fitted in order to estimate treatment effects, test statistics and information for each simulated trial; details of simulations and the how they were implemented are reported in detail by Parsons et al. [8]. The results of the simulations for START:REACTS showed that for 90% power, a minimum

Table 2 Minimum and maximum values of V_s^{exp} and values for the equal spacing model $\tilde{V}_s^{exp}$, for $\gamma = 0.5$ for early (t_1), mid (t_2) and late (t_3) interim analyses, where $\tau_{01} = 0.15$, $\tau_{02} = 0.30$ and $\tau_{03} = 0.45$ respectively, for s from two to six for the fixed, increasing and decreasing rate recruitment models. Also shown are values of n_{s1} for each setting

s	Early (t_1)			Mid (t_2)			Late (t_3)		
	n_{s1}	$\tilde{V}_s^{exp}$	V_s^{exp}	n_{s1}	$\tilde{V}_s^{exp}$	V_s^{exp}	n_{s1}	$\tilde{V}_s^{exp}$	V_s^{exp}
			min - max			min - max			min - max
[a] Fixed									
2	0.55	-	0.89-0.89	0.71	-	0.93-0.93	0.78	-	0.95-0.95
3	0.55	0.81	0.80-0.89	0.71	0.88	0.88-0.93	0.78	0.92	0.91-0.95
4	0.55	0.78	0.78-0.89	0.71	0.87	0.86-0.93	0.78	0.90	0.90-0.95
5	0.55	0.77	0.76-0.89	0.71	0.86	0.85-0.93	0.78	0.90	0.90-0.95
6	0.55	0.76	0.75-0.89	0.71	0.85	0.85-0.93	0.78	0.89	0.89-0.95
[b] Increasing									
2	0.59	-	0.90-0.90	0.68	-	0.92-0.92	0.72	-	0.93-0.93
3	0.59	0.83	0.83-0.90	0.68	0.87	0.87-0.92	0.72	0.89	0.89-0.93
4	0.59	0.81	0.80-0.90	0.68	0.85	0.85-0.92	0.72	0.88	0.87-0.93
5	0.59	0.80	0.79-0.90	0.68	0.84	0.84-0.92	0.72	0.87	0.87-0.93
6	0.59	0.79	0.79-0.90	0.68	0.84	0.84-0.92	0.72	0.86	0.86-0.93
[c] Decreasing									
2	0.42	-	0.86-0.86	0.62	-	0.91-0.91	0.74	-	0.93-0.93
3	0.42	0.75	0.73-0.86	0.62	0.84	0.84-0.91	0.74	0.89	0.89-0.93
4	0.42	0.71	0.69-0.86	0.62	0.82	0.82-0.91	0.74	0.88	0.88-0.93
5	0.42	0.69	0.68-0.86	0.62	0.81	0.80-0.91	0.74	0.87	0.87-0.93
6	0.42	0.67	0.67-0.86	0.62	0.80	0.80-0.91	0.74	0.87	0.86-0.93

of $N = 188$ participants were required, with the expected number of participants providing outcome data and trial information at the interim analyses shown in Table 3[a]. The expected information at the study end was given by $I = N/(4\sigma_s^2) = 188/(4 \times 144) = 0.326$. Table 3[b] shows the observed numbers of participants providing data and the information and the estimated test statistic ($Z = \beta_s/\text{sd}(\beta_s)$) at the first interim analysis when the study was stopped for futility.

The observed correlations between outcomes were larger than expected at the first interim analysis when the trial was stopped; $\alpha \approx 0.75$, not $\alpha = 0.5$ as planned. Although the observed numbers of participants providing outcome data were reasonably similar to the expected numbers, this was more by chance than by design, as the number of recruiting sites used for the trial was actually 24, not the planned 15, and the pattern of site initiations was quite different from the plan.

The expectation in the original study design was that the interim analyses should occur when approximately 25% and 35% of the trial participants had final outcome data; that is when $\tau 0(t_1) = 0.25$ and $\tau 0(t_2) = 0.35$. For the uniform correlation model, setting d_r for $r = 1, 2, 3$ to reflect the spacing of the outcomes at 3, 6 and 12 months (e.g. we can set $d_1 = 1$, $d_2 = 2$ and $d_3 = 4$) and $T_R = 2d_3$ (i.e. the recruitment period is 24 months and the final outcome is at 12 months) allows us to calculate $V_3^{\text{unif}}(t)$ at $t = t_1$ and $t = t_2$. From expression (8), in "Recruitment models and $\text{var}(\beta_s)$" section and setting $\alpha = 0.5$, these are as follows $V_3^{\text{unif}}(t_1) = 0.808$ and $V_3^{\text{unif}}(t_2) = 0.836$ for $t_1 = 6$ (18 months) and $t_2 = 6.8$ (20.4 months), the times when $\tau 0(t_1) = 0.25$ and $\tau 0(t_2) = 0.35$, for the example formulations for d_r. From expression (16), in "Information" section, the information fractions at these interim analyses are $\tau(t_1) = 0.309$ and $\tau(t_2) = 0.419$. The information fractions allow us to calculate bounds and probabilities (power), using for instance functions `gsBound` and `gsProbability` from the R package `gsDesign` [26], for selected values of the overall trial sample size N. Table 4[a] shows the expected numbers of participants providing data and the information and test statistic boundaries at the first and second interim analyses for the fixed recruitment model for $N = 188$. Power for the fixed recruitment rate made is 90.6% for a treatment difference of 6 and $\sigma_3 = 12$, the same as the actual START:REACTS trial.

We can repeat the above calculation quite simply for the decreasing rate recruitment model, recruiting over the same length of time at the same three time-points (3, 6 and 12 months) and, for the same choice of covariance parameters, we get $V_3^{\text{unif}}(t_1) = 0.786$ and $V_3^{\text{unif}}(t_2) = 0.820$ for $t_1 = 5.13$ (15.4 months) and $t_2 = 5.64$ (16.9 months) with information fractions at these interim analyses of $\tau(t_1) = 0.318$ and $\tau(t_2) = 0.427$. Table 4[b] shows the expected numbers of participants providing data and the information and test statistic boundaries at the first and second interim analyses for the decreasing recruitment model for $N = 188$. Power for the decreasing recruitment rate made is 90.7%.

The final trial reported a strong effect in favour of the control group -4·2 (95% CI -8·2 to -0·26), rather than the expected effect in favour of the treatment group [7]. If a result of this magnitude in favour of the control group had been anticipated, then we would have had the probabilities p_1 and p_2 of stopping for futility at the interim analyses shown in Fig. 2a and b, as functions of α and the first analysis time-point t_1. At the extreme of the latter values shown at $t_1 = 4.5$ (13.5 months)

Table 3 START:REACTS study planning and observed trial data. Numbers of participants providing outcome data at 3, 6 and 12 months, information ($I = 1/\text{var}(\beta_s)$) and test statistic ($Z = \beta_s/\text{sd}(\beta_s)$) boundaries in [a] the expected (planned) study design based on extensive simulations and [b] observed in the trial itself. Note that the trial was stopped at the first interim analysis for futility, as the test statistic fell below the lower boundary

[a] Expected					[b] Observed		
Outcome		Interim		End	Interim		End
		1	2		1	2	
12 months	N_3	50	70	188	47	-	-
6 months	N_2	90	110	188	86	-	-
3 months	N_1	120	140	188	112	-	-
Information	I	0.102	0.139	0.326	0.110	-	-
Information fraction	τ	0.311	0.426	1.000	-	-	-
Test statistic	Z	-	-	-	-0.881	-	-
Boundary	lower	-0.706	0.581	1.909	-	-	-
	upper	∞	3.090	1.909	-	-	-

Table 4 Numbers of participants providing outcome data at 3, 6 and 12 months, information ($I = 1/\text{var}(\beta_s)$) and test statistic ($Z = \beta_s/\text{sd}(\beta_s)$) boundaries for putative START:REACTS trial designs where the sample size is $N = 188$ for [a] an expected fixed rate and [b] a decreasing rate of recruitment

[a] Fixed rate					[b] Decreasing rate		
Outcome		Interim		End	Interim		End
		1	2		1	2	
12 months	N_3	47.0	65.8	188	47.0	65.8	188
6 months	N_2	94.0	112.8	188	113.5	127.0	188
3 months	N_1	117.5	136.3	188	138.9	149.8	188
Information	I	0.101	0.137	0.326	0.104	0.139	0.326
Information fraction	τ	0.309	0.419	1.000	0.318	0.427	1.000
Boundary	lower	-0.706	0.581	1.907	-0.706	0.581	1.910
	upper	∞	3.090	1.907	∞	3.090	1.910
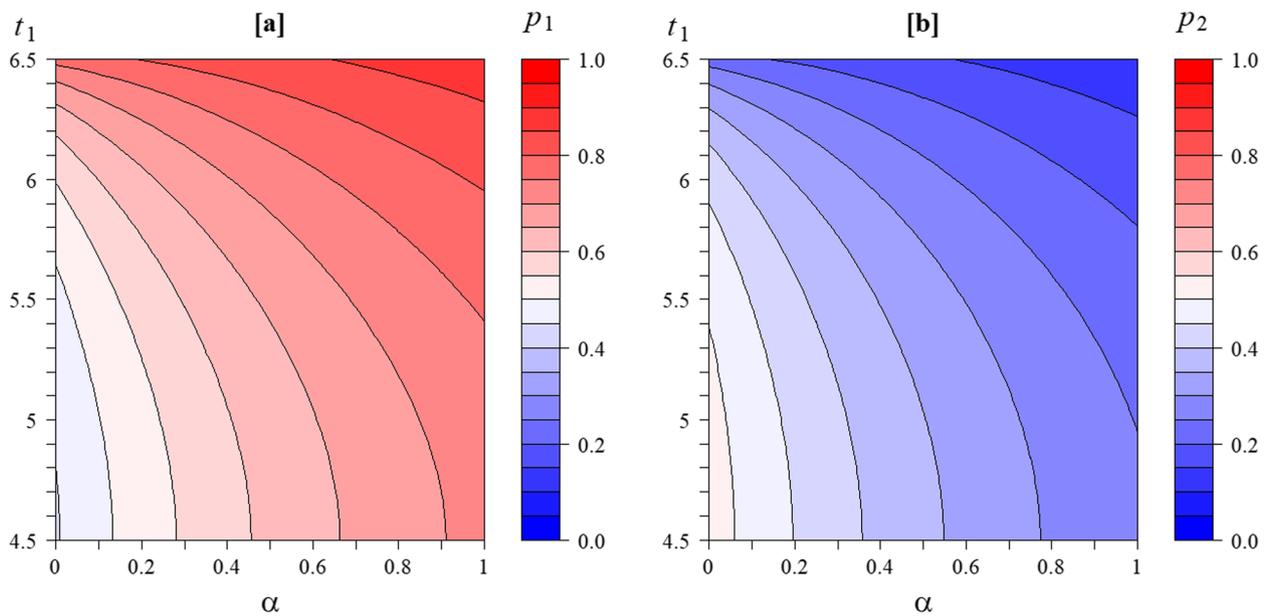
Fig. 2 Contour plots showing futility stopping probabilities for a treatment difference of -4 (in favour of the control group) for the fixed rate recruitment and the uniform correlation model at (**a**) the first interim (p_1) and (**b**) the second interim analysis (p_2) as functions the correlation α and the timing of the first interim analysis (t_1; in 3 monthly base unit, such that for instance $t_1 = 5$ corresponds to 15 months)

there would have been very little outcome data for the fixed rate model ($N_3 = 11.8$) and a consequent modest value of $p_1 = 0.446$, with $p_2 = 0.525$ ($N_3 = 65.8$), assuming no correlation $\alpha = 0$. Increasing the correlation to $\alpha = 0.8$ would provide a considerably larger first interim analysis futility stopping probabilities $p_1 = 0.581$. A later first interim analysis at 19.5 months ($t_1 = 6.5$) would provide much more data ($N_3 = 58.8$) and a larger probability of stopping $p_1 = 0.716$ for $\alpha = 0$ ($p_1 = 0.820$ at $\alpha = 0.8$).

Discussion

The numerical examples of "Numerical examples" section show how patterns of information accrual vary between correlation and recruitment models and provide some general guidance for planning a group sequential trial with early outcome data. Given equal numbers of participants providing early and final outcome data, the stronger the correlation between early and final outcomes the greater the gain in information (reduction in variance) at some interim time-point t. Expressions (8)

and (14) show $V_s(t)$, the relative effect of the early outcomes on reducing the variance of the primary outcome, for the uniform and exponential models to be monotonically increasing in the correlation parameters ($0 \leq \alpha < 1$ and $0 \leq \gamma < 1$) under our model constraints. Therefore, if a number of suitable outcome measures are available then one should choose the one with the strongest correlations between serial time-points, assuming that the measures are equally responsive to change and have similar variances. Making recommendations for the number and, particularly, the timing of early outcomes is more complex. The simulations (Tables 1 and 2) show that there was generally little change in V_s for $s > 5$, suggesting that for the models tested there was little to be gained by having more than four early outcome time-points. For the exponential model, equally spaced early outcomes, for known γ, proved to be a sensible, and often almost optimal, choice to give the greatest information gain (i.e. values of $\tilde{V}_s^{exp}$ are close to $\min(V_s^{exp})$). For smaller values of γ there was some value for the exponential model in moving the early outcomes to be nearer to the final outcome. However, one might argue that such an approach is not sensible for such small values of γ as little would be gained in practice by waiting to collect early outcomes at such a late time-point. For the time invariant uniform correlation model, maximum information gains are made (trivially) when the early outcomes occur as early as possible, simply because more data will be available as more trial participants will been followed-up. Clearly, it is not sensible to assume that the uniform model must apply for any spacing between outcome time-points. Therefore, it is worth emphasising that the results presented in Tables 1 and 2 are strongly dependent on the model assumptions, in the sense that we have assumed that changing the timings of the early outcomes is possible for some fixed value of the correlation parameter. For instance, if there was known to be an equal correlation of magnitude α between four equally-spaced outcome time-points, then would this correlation model still be appropriate if the first two outcomes and the last two outcomes were moved to be almost coincident? It seems highly unlikely. This being the case, we would caution against using numerical experiments (such as those of "Uniform correlation model" and "Exponential correlation model" sections) solely as a means to make decisions on the spacing of outcome time-points. However, we do believe that within reasonable limits it would be profitable to explore the likely information gains that alternate spacing models may offer. Although, such decisions may be dependent on the specific application area and other trial constraints such as when participant follow-up would routinely be available. One aspect of the numerical simulation studies that is clear is that the decreasing

recruitment rate model is preferable, in terms of information gain, to either the increasing or fixed rate models. This is due to a greater proportion of the study participants being available to provide early outcome data for the decreasing rate model. In many instances, there is little one could do about the likely pattern of recruitment into a trial, but in settings where this was possible then clearly it would be advantageous to plan to recruit a large proportion of the target sample size into the trial as early as possible.

The STARTS:REACTS trial provided a real example of a study using an information adaptive approach for a group sequential trial with early outcome data [7]. The study was originally planned based on a large study that simulated individual participant data [8]. What the results of "START:REACTS clinical trial" section show is that the trial could have been planned, with little effort and without the need for time-consuming coding and simulation, based on the models described here. If that had been done, then, as Tables 3 and 4 show, the final design would have been almost exactly equivalent to that used in the original study, at considerably less effort and cost. Additionally, a range of other options (e.g. changing the number and timing of interim analysis) could have been explored with none of the considerable extra work and effort that would have been required if we were simulating outcomes for individual participants. An advantage of using the approaches to design outlined here, rather than simulation models, is that the procedures are absolutely specified and explicit, and therefore easily checked and replicated by others. Whereas, in a simulation study much relies on the availability of the code and the readability and competence of the coding and also the assumptions made by those developing the code, that are often not explicitly stated. For these, and many other reasons, we would strongly recommend that those wishing to use the group sequential designs described here use the one of the selection of models in "Correlation models" section. Clearly, if one believes that the study setting is of a completely different type, or does not approximate to one of the settings described here then simulation may be the only option to determining appropriate trial sample sizes.

This study presents certain limitations. The choices of fixed, decreasing and increasing recruitment rate models ("Recruitment and follow-up models" section) was essentially arbitrary and used mainly as a means of showing a range of contrasting options. The fixed rate model might represent for instance a situation where participants were identified and recruited into a study at a fixed rate across one or more recruitment centres. Whereas the increasing rate model might arise naturally if the number of centres was likely to increase during recruitment and each centre recruited at the same

fixed rate, resulting in an overall recruitment rate that followed the profile seen in Fig. 1b. The decreasing rate model Fig. (1c) provides the type of recruitment profile that might be observed where there is an existing pool of participants who are available to enter a trial quickly, resulting in a rapid rise that is followed by a slowing rate of accrual after the pool is exhausted and we have to rely solely on a new (incident) cases of a condition being identified. All of these scenarios are amongst the many we have observed within our own trials experience. Although we fully accept that the settings we present may not cover every possible option that trialists using these methods may wish to consider. However, we believe it would be relatively easy to suggest and implement a range of other sensible models, provided they followed the general structures and properties we outline in "Recruitment and follow-up models" section. Similarly, although clearly it would be possible to consider more complex correlation models than those described in "Correlation models" section, we chose the uniform and exponential correlation models mainly because they are very widely used for longitudinal outcomes, are often good approximations to observed data and also because they lead to simple analytic expressions for $\text{var}(\beta_s)$, and as such allow us to illustrate some key ideas about the methodological approach described here [4, 22]. In practice, if we wished to assume that outcomes followed an exponential correlation model based on limited data, then we could reason as follows. For instance, consider a study that is being planned with four outcomes at 3, 6, 12 an 18 months, with the latter as the final (primary) outcome, and the others as early outcomes. Data from another study suggests that the correlation between outcomes at 3 and 12 months is approximately $\rho_{3m,12m} = 0.5$, and therefore by noting that $\gamma = \rho_{rr'}^{1/|d_r - d_{r'}|}$, we can write $\gamma = 0.5^{1/3}$, by setting $d_1 = 1, d_2 = 2, d_3 = 4$ and $d_4 = 6$ to model the outcome spacings. In this model, $\rho_{3m,6m} = 0.79$, $\rho_{3m,12m} = 0.50$, $\rho_{3m,18m} = 0.31$, $\rho_{6m,12m} = 0.63$, $\rho_{6m,18m} = 0.40$ and $\rho_{12m,18m} = 0.63$. The simple expressions for $\text{var}(\beta_s)$ for the exponential and uniform correlation models are due to the fact that general expressions are available for R_s^{-1}; see Appendix A1 and A2. Therefore, if similar general expressions were available for alternate correlation models, then in principle we believe it would be possible to provide analytic expressions for such models.

Currently, those who wish to exploit the methodology reported here will need to implement the results themselves. However, work is ongoing to develop a package of R functions [23] to implement the models in a form that will make it easy for the user to explore all the design options described in a simple and interactive manner.

Conclusions

We have developed models for information accrual during recruitment into a group sequential clinical trial using early outcomes to augment the information available from the trial primary outcome measures as a means to make decisions about whether to stop prior to the completion of recruitment [4, 8]. The analytic solutions provided in "Correlation models" section are based on some simple, but we believe realistic and useful, models of recruitment into the study and the serial correlation between the early and final outcome measures reported during participant follow-up. Although in general the correlations may be unknown at the planning stage, we can speculate on the likely correlation structure. In an analogous way to what we might do for variances in a conventional trial. At some arbitrarily selected point during recruitment we will not (in general) know the exact number of participants recruited or the number of participants ($N0_r$ and $N1_r$) with data at each follow-up occasion *r*. However, we can speculate on the likely recruitment rates and therefore the likely number of participants providing follow-up data at any point during the trial. Given the above we can predict how $\text{var}(\beta_s)$, and therefore information, will vary during the study and use this to motivate the choice and timings of the interim analyses. The models provide analytic expressions for information accrual that can be routinely used at the planning stage of a putative trial, based on some modest assumptions about the likely number of outcomes and when they might occur and the expected recruitment patterns. Numerical simulations show that these models behave sensibly (i.e. in a manner that we would expect) and allow us to explore a range of design options in a way that would have been considerably more difficult and time-consuming if we had to use the previously described method of simulating individual trial participant data.

Supplementary Information

Additional file 1: Figure S1. The feasible region of V_s^{unif} for s = 2 (shaded areas), bounded above by the maximum and below by the minimum, for correlations in the range 0 ≤ α < 1 and equal group sizes (φ = 0.5) for the decreasing, fixed and increasing rate recruitment models with lines for the setting where the time-points are given by dr = 1+(r−1)/(s−1) (r = 1, 2) for [a] early (τ 01 = 0.15), [b] mid (τ 02 = 0.30) and [c] late (τ 03 = 0.45) interim analyses. **Figure S2.** The feasible region of V_s^{unif} for s = 3 (shaded areas), bounded above by the maximum and below by the minimum, for correlations in the range 0 ≤ α < 1 and equal group sizes (φ = 0.5) for the decreasing, fixed and increasing rate recruitment models with lines for the setting where the time-points are given by dr = 1+(r−1)/(s−1) (r = 1, 2, 3; i.e. equal spacing) for [a] early (τ 01 = 0.15), [b] mid (τ 02 = 0.30) and [c] late (τ 03 = 0.45) interim analyses. **Figure S3.** The empirical distribution (nsim = 10000) of ΔV_s^{unif}, the difference from the median value of V_s^{unif}, with varying 1 < dr < 2 (r = 2) for s = 3, with shading showing quantiles 0-5%, 5-25%, 25-50%, 50-75%, 75-95% and 95-100%, for

correlations in the range $0 \leq \alpha < 1$ and equal group sizes ($\varphi = 0.5$) for [a] early ($\tau 01 = 0.15$), [b] mid ($\tau 02 = 0.30$) and [c] late ($\tau 03 = 0.45$) interim analyses, for the (i) increasing, (ii) fixed and (iii) decreasing rate recruitment models. **Figure S4.** The feasible region of V_s^{unif} for s = 4 (shaded areas), bounded above by the maximum and below by the minimum, for correlations in the range $0 \leq \alpha < 1$ and equal group sizes ($\varphi = 0.5$) for the decreasing, fixed and increasing rate recruitment models with lines for the setting where the time-points are given by dr = 1+(r−1)/(s−1) (r = 1, 2, 3, 4; i.e. equal spacing) for [a] early ($\tau 01 = 0.15$), [b] mid ($\tau 02 = 0.30$) and [c] late ($\tau 03 = 0.45$) interim analyses. **Figure S5.** The empirical distribution (nsim = 10000) of Δ , the difference from the median value of V_s^{unif} , with varying 1 < dr < 2 (r = 2, 3) for s = 4, with shading showing quantiles 0-5%, 5-25%, 25-50%, 50-75%, 75-95% and 95-100%, for correlations in the range $0 \leq \alpha < 1$ and equal group sizes ($\varphi = 0.5$) for [a] early ($\tau 02 = 0.30$) and [c] late ($\tau 03 = 0.45$) interim analyses, for the (i) increasing, (ii) fixed and (iii) decreasing rate recruitment models. **Figure S6.** The feasible region of V_s^{unif} for s = 5 (shaded areas), bounded above by the maximum and below by the minimum, for correlations in the range $0 \leq \alpha < 1$ and equal group sizes ($\varphi = 0.5$) for the decreasing, fixed and increasing rate recruitment models with lines for the setting where the time-points are given by dr = 1+(r−1)/(s−1) (r = 1, 2, 3, 4, 5; i.e. equal spacing) for [a] early ($\tau 01 = 0.15$), [b] mid ($\tau 02 = 0.30$) and [c] late ($\tau 03 = 0.45$) interim analyses. **Figure S7.** The empirical distribution (nsim = 10000) of ΔV_s^{unif}, the difference from the median value of V_s^{unif} , with varying 1 < dr < 2 (r = 2, 3, 4) for s = 5, with shading showing quantiles 0-5%, 5-25%, 25-50%, 50-75%, 75-95% and 95-100%, for correlations in the range $0 \leq \alpha < 1$ and equal group sizes ($\varphi = 0.5$) for [a] early ($\tau 01 = 0.15$), [b] mid ($\tau 02 = 0.30$) and [c] late ($\tau 03 = 0.45$) interim analyses, for the (i) increasing, (ii) fixed and (iii) decreasing rate recruitment models. **Figure S8.** The feasible region of V_s^{unif} for s = 6 (shaded areas), bounded above by the maximum and below by the minimum, for correlations in the range $0 \leq \alpha < 1$ and equal group sizes ($\varphi = 0.5$) for the decreasing, fixed and increasing rate recruitment models with lines for the setting where the time-points are given by dr = 1+(r−1)/(s−1) (r = 1, 2, 3, 4, 5, 6; i.e. equal spacing) for [a] early ($\tau 01 = 0.15$), [b] mid ($\tau 02 = 0.30$) and [c] late ($\tau 03 = 0.45$) interim analyses. **Figure S9.** The empirical distribution (nsim = 10000) of ΔV_s^{unif} , with varying 1 < dr < 2 (r = 2, 3, 4, 5) for s = 6, with shading showing quantiles 0-5%, 5-25%, 25-50%, 50-75%, 75-95% and 95-100%, for correlations in the range $0 \leq \alpha < 1$ and equal group sizes ($\varphi = 0.5$) for [a] early ($\tau 01 = 0.15$), [b] mid ($\tau 02 = 0.30$) and [c] late ($\tau 03 = 0.45$) interim analyses, for the (i) increasing, (ii) fixed and (iii) decreasing rate recruitment models. **Figure S10.** The feasible region of Vexp s for s = 2 (shaded areas), bounded above by the maximum and below by the minimum, for correlations in the range 0 $\leq \gamma < 1$ and equal group sizes ($\varphi = 0.5$) for the decreasing, fixed and increasing rate recruitment models with lines for the setting where the time-points are given by dr = 1+(r−1)/(s−1) (r = 1, 2) for [a] early ($\tau 01 = 0.15$), [b] mid ($\tau 02 = 0.30$) and [c] late ($\tau 03 = 0.45$) interim analyses. **Figure S11.** The feasible region of Vexp s for s = 3 (shaded areas), bounded above by the maximum and below by the minimum, for correlations in the range $0 \leq \gamma < 1$ and equal group sizes ($\varphi = 0.5$) for the decreasing, fixed and increasing rate recruitment models with lines for the setting where the time-points are given by dr = 1+(r−1)/(s−1) (r = 1, 2, 3; i.e. equal spacing) for [a] early ($\tau 01 = 0.15$), [b] mid ($\tau 02 = 0.30$) and [c] late ($\tau 03 = 0.45$) interim analyses. **Figure S12.** The empirical distribution (nsim = 10000) of ΔVexp s , the difference from the median value of Vexp s , with varying 1 < dr < 2 (r = 2) for s = 3, with shading showing quantiles 0-5%, 5-25%, 25-50%, 50-75%, 75-95% and 95-100%, for correlations in the range $0 \leq \gamma < 1$ and equal group sizes ($\varphi = 0.5$) for [a] early ($\tau 01 = 0.15$), [b] mid ($\tau 02 = 0.30$) and [c] late ($\tau 03 = 0.45$) interim analyses, for the (i) increasing, (ii) fixed and (iii) decreasing rate recruitment models. **Figure S13.** The feasible region of Vexp s for s = 4 (shaded areas), bounded above by the maximum and below by the minimum, for correlations in the range 0 $\leq \gamma < 1$ and equal group sizes ($\varphi = 0.5$) for the decreasing, fixed and increasing rate recruitment models with lines for the setting where

the time-points are given by dr = 1+(r−1)/(s−1) (r = 1, 2, 3, 4; i.e. equal spacing) for [a] early ($\tau 01 = 0.15$), [b] mid ($\tau 02 = 0.30$) and [c] late ($\tau 03 = 0.45$) interim analyses. **Figure S14.** The empirical distribution (nsim = 10000) of ΔVexp s , the difference from the median value of Vexp s , with varying 1 < dr < 2 (r = 2, 3) for s = 4, with shading showing quantiles 0-5%, 5-25%, 25-50%, 50-75%, 75-95% and 95-100%, for correlations in the range $0 \leq \gamma < 1$ and equal group sizes ($\varphi = 0.5$) for [a] early ($\tau 01 = 0.15$), [b] mid ($\tau 02 = 0.30$) and [c] late ($\tau 03 = 0.45$) interim analyses, for the (i) increasing, (ii) fixed and (iii) decreasing rate recruitment models. **Figure S15**. The feasible region of Vexp s for s = 5 (shaded areas), bounded above by the maximum and below by the minimum, for correlations in the range $0 \leq \gamma < 1$ and equal group sizes ($\varphi = 0.5$) for the decreasing, fixed and increasing rate recruitment models with lines for the setting where the time-points are given by dr = 1+(r−1)/(s−1) (r = 1, 2, 3, 4, 5; i.e. equal spacing) for [a] early ($\tau 01 = 0.15$), [b] mid ($\tau 02 = 0.30$) and [c] late ($\tau 03 = 0.45$) interim analyses. **Figure S16**. The empirical distribution (nsim = 10000) of ΔVexp s , the difference from the median value of Vexp s , with varying 1 < dr < 2 (r = 2, 3, 4) for s = 5, with shading showing quantiles 0-5%, 5-25%, 25-50%, 50-75%, 75-95% and 95-100%, for correlations in the range $0 \leq \gamma < 1$ and equal group sizes ($\varphi = 0.5$) for [a] early ($\tau 01 = 0.15$), [b] mid ($\tau 02 = 0.30$) and [c] late ($\tau 03 = 0.45$) interim analyses, for the (i) increasing, (ii) fixed and (iii) decreasing rate recruitment models. **Figure S17.** The feasible region of Vexps for s = 6 (shaded areas), bounded above by the maximum and below by the minimum, for correlations in the range $0 \leq \gamma < 1$ and equal group sizes ($\varphi = 0.5$) for the decreasing, fixed and increasing rate recruitment models with lines for the setting where the time-points are given by dr = 1+(r−1)/(s−1) (r = 1, 2, 3, 4, 5, 6; i.e. equal spacing) for [a] early ($\tau 01 = 0.15$), [b] mid ($\tau 02 = 0.30$) and [c] late ($\tau 03 = 0.45$) interim analyses. **Figure S18.** The empirical distribution (nsim = 10000) of ΔVexp s , the difference from the median value of Vexp s , with varying 1 < dr < 2 (r = 2, 3, 4, 5) for s = 6, with shading showing quantiles 0-5%, 5-25%, 25-50%, 50-75%, 75-95% and 95-100%, for correlations in the range $0 \leq \gamma < 1$ and equal group sizes ($\varphi = 0.5$) for [a] early ($\tau 01 = 0.15$), [b] mid ($\tau 02 = 0.30$) and [c] late ($\tau 03 = 0.45$) interim analyses, for the (i) increasing, (ii) fixed and (iii) decreasing rate recruitment models.

Additional file 1: Appendix A. A.1 Uniform correlation model. A.2 Exponential correlation model. A.3 Partial derivatives of Vexps. A.4 Recruitment and follow-up models. Table A1 Times (t1, t2 and t3) for early τ0(t1) = 0.15, mid τ0(t2) = 0.30 and late τ0(t3) = 0.45 interim analyses, for increasing, fixed and decreasing rate recruitment models.

Authors' contributions

NP conceived and undertook the research with discussion and feedback from NS and JB. NP, JB and NS discussed and commented on the manuscript in draft, and NP finalized the manuscript and prepared the figures and tables for the submission.

Availability of data and materials

Not applicable. as no data or materials were used in this research.

Code availability

The code used to generate the numerical examples is available on request from the corresponding author (NP). Work is ongoing to develop a package of R functions to allow simple implementation of the models described.

Declarations

Ethics approval and consent to participate

Not applicable, as no patient data were used in this research.

Parsons *et al. BMC Medical Research Methodology*   (2024) 24:42

Page 17 of 17

References

1. Hatfield I, Allison A, Flight L, Julious SA, Dimairo M. Adaptive designs undertaken in clinical research: a review of registered clinical trials. Trials. 2016;17(1):150.
2. Jennison C, Turnbull BW. Group sequential methods with applications to clinical trials. Boca Raton: Chapman and Hall; 2000.
3. Dimairo M, Boote J, Julious SA, Nicholl JP, Todd S. Missing steps in a staircase: a qualitative study of the perspectives of key stakeholders on the use of adaptive designs in confirmatory trials. Trials. 2015;16:430.
4. Parsons NR, Stallard N, Parsons H, Haque A, Underwood M, Mason J, et al. Group sequential designs in pragmatic trials: feasibility and assessment of utility using data from a number of recent surgical RCTs. BMC Med Res Methodol. 2022;22(1). https://doi.org/10.1186/s12874-022-01734-2. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85139109334 & doi=10.1186%2fs12874-022-01734-2 &partnerID=40 &md5=39107fa728 feb689a7f31f5569cd006e.
5. Roland M, Torgerson DJ. What are pragmatic trials? BMJ. 1998;316:285.
6. Ford I, Norrie J. Pragmatic trials. N Engl J Med. 2016;375(5):454–63. https://doi.org/10.1056/NEJMra1510059.
7. Metcalfe A, Parsons H, Parsons N, Brown J, Fox J, Gemperle Mannion E, et al. Subacromial balloon spacer for irreparable rotator cuff tears of the shoulder (START:REACTS): a group-sequential, double-blind, multicentre randomised controlled trial. Lancet. 2022;399:1954–63.
8. Parsons N, Stallard N, Parsons H, Wells P, Underwood M, Mason J, et al. An adaptive two-arm clinical trial using early endpoints to inform decision making: design for a study of sub-acromial spacers for repair of rotator cuff tendon tears. Trials. 2019;20(1):694. https://doi.org/10.1186/s13063-019-3708-6. https://www.ncbi.nlm.nih.gov/pubmed/31815651
9. Barnard KD, Dent L, Cook A. A systematic review of models to predict recruitment to multicentre clinical trials. BMC Med Res Methodol. 2010;10:63. https://doi.org/10.1186/1471-2288-10-63.
10. Galbraith S, Marschner IC. Interim analysis of continuous long-term endpoints in clinical trials with longitudinal outcomes. Stat Med. 2003;22(11):1787–805.
11. Jennison C, Turnbull BW. Group-sequential analysis incorporating covariate information. J Am Stat Assoc. 1997;92:1330–41.
12. Van Lancker K, Vandebosch A, Vansteelandt S. Improving interim decisions in randomized trials by exploiting information on short-term endpoints and prognostic baseline covariates. Pharm Stat. 2020;19:583–601.
13. Viele K, McGlothlin A, Broglio K. Interpretation of clinical trials that stopped early. JAMA. 2016;315(15):1646–7.
14. Liu A, Hall W. Unbiased estimation following a group sequential test. Biometrika. 1999;86(1):71–8.
15. Todd S, Whitehead J, Facey KM. Point and interval estimation following a sequential clinical trial. Biometrika. 1996 06;83(2):453–461.
16. Stallard N. A confirmatory seamless Phase II/III clinical trial design incorporating short-term endpoint information. Stat Med. 2010;29:959–71. https://doi.org/10.1002/sim.3863.
17. Engel B, Walstra P. Increasing precision or reducing expense in regression experiments by using information from a concomitant variable. Biometrics. 1991;47(1):13–20. https://doi.org/10.2307/2532491. https://www.jstor.org/stable/2532491
18. Spiessens B, Lesaffre E, Verbeke G, Kim K, DeMets DL. An overview of group sequential methods in longitudinal clinical trials. Stat Methods Med Res. 2000;9(5):497–515.
19. Spiessens B, Lesaffre E, Verbeke G. A comparison of group sequential methods for binary longitudinal data. Stat Med. 2003;22(4):501–15.
20. Kim K, Anastasios AT. Independent increments in group sequential tests: a review. Stat Oper Res Trans. 2020;44(2):223–64.
21. Qian T, Rosenblum M, Qiu H. Improving power in group sequential, randomized trials by adjusting for prognostic baseline variables and shortterm outcomes. Johns Hopkins University, Dept of Biostatistics Working Papers, Working paper 285. 2016;5(5):5.
22. Diggle P, Diggle P. Analysis of longitudinal data. 2nd ed. Oxford statistical science series. Oxford: Oxford University Press; 2002.
23. R Core Team. R: A Language and Environment for Statistical Computing. Vienna; 2022. https://www.R-project.org/. Accessed 3 Nov 2023.
24. Pinheiro JC, Bates D. Mixed-Effects Models in S and S-PLUS. Statistics and Computing. New York: Springer; 2009.
25. Lan KKG, Reboussin DM, DeMets DL. Information and information fractions for design and sequential monitoring of clinical trials. Commun Stat Theory Methods. 1994;23(2):403–20.
26. Anderson K. gsDesign: Group Sequential Design. 2021. https://CRAN.R-project.org/package=gsDesign. Accessed 3 Nov 2023.
27. Metcalfe A, Gemperle Mannion E, Parsons H, Brown J, Parsons N, Fox J, et al. Protocol for a randomised controlled trial of Subacromial spacer for Tears Affecting Rotator cuff Tendons: a Randomised, Efficient, Adaptive Clinical Trial in Surgery (START:REACTS). BMJ Open. 2020;10(5): e036829. https://doi.org/10.1136/bmjopen-2020-036829. https://www.ncbi.nlm.nih.gov/pubmed/32444433

Publisher's Note