

(*) finally, in sect 2.3, we describe how the prev. method can be used to.. *some hints at intro?*

2. Methods

In Section 2.1, we review current methods for incorporating non-concurrent controls into the analysis of platform trials and discuss the differences and similarities, as well as their limitations. These methods include model-based frequentist approaches, the "Time Machine" Bayesian approach and the network meta-analysis approach.

In Section 2.2, we extend the frequentist model approaches for treatment-control comparisons to more complex platform trials than those considered so far. Furthermore, we explore different variants of the models by considering different definitions of the **time covariate** and propose novel methods that permit further flexibility in the adjustment.

(*) In Section 2.3 we expand the application of the methods to treatment-treatment comparisons between treatments that do not overlap in the trial, but have common concurrent arms.

this might not always be true, right? *previous-mentioned platform*

2.1. Current Methods for Incorporating Non-concurrent Controls

Consider a platform trial design with K experimental treatment arms and a shared control group, where the experimental arms enter and leave the trial at different time points. The duration of the trial can then be split into periods, defined as time intervals bounded by any treatment arm either entering or leaving the platform. Moreover, the time can be divided into equidistant discrete units of calendar time.

We denote the experimental treatment arms by k ($k = 1, \dots, K$, ordered by entry time) and the control group by $k = 0$. Furthermore, we denote the periods the trial consists of by s ($s = 1, \dots, S$) and the calendar time units by c ($c = 1, \dots, C$). The total number of patients in the trial is given by N . The response for patient j ($j = 1, \dots, N$) is indicated by y_j , the arm they were allocated to by k_j and the period and calendar time corresponding to their entry time by s_j and c_j , respectively. The observed time of recruitment of patient j is given by t_j .

The focus of this thesis is on continuous endpoints, hence the frequentist model-based approaches will only be discussed in this context. The Bayesian Time Machine, however, has only been proposed for binary endpoints so far and therefore it will be described for this type of endpoints.

2.1.1. Frequentist Model-based Approaches

Focusing on a simple platform trial with $K = 2$ experimental treatment arms, Lee and Wason [1] investigated linear regression models that allow to include non-concurrent

2. Methods

control data to the analysis. In particular, they assumed a trial, which starts with only one treatment arm and the control group. The second treatment arm is added to the trial later on and both treatment arms finish at the same time. The resulting trial consists of 2 periods, as illustrated in Fig. 2.1. Moreover, time trends of various strengths and shapes may be present in the trial.

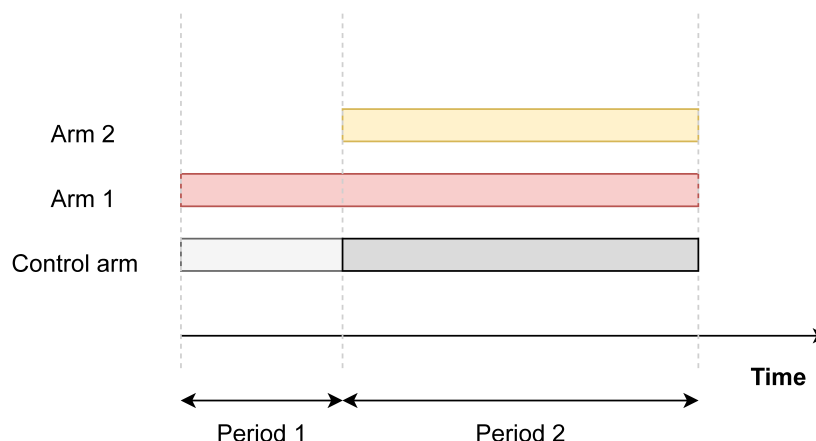


Figure 2.1.: Scheme of the platform trial considered in [1] and [2].

Lee and Wason [1] proposed two linear regression models for evaluating the efficacy of the second treatment arm, which use all data available in the trial (i.e. data from both experimental treatment arms and the control arm). The potential time trends are adjusted for by including the time as a covariate to the model, either in terms of the patient recruitment time (given by (2.1)) or the period indicator (given by (2.2)), as follows:

$$y_j = \eta_0 + \sum_{k=1,2} \theta_k \cdot I(k_j = k) + \gamma \cdot j + \varepsilon_j \quad (2.1)$$

$$y_j = \eta_0 + \sum_{k=1,2} \theta_k \cdot I(k_j = k) + \tau \cdot I(t_j \in s_2) + \varepsilon_j \quad (2.2)$$

where the model intercept η_0 is given in terms of the response in the control group in the first period and θ_k is the treatment effect for treatment k . The time effect is estimated by the parameters γ and τ . In (2.1), γ represents the effect of continuous recruitment, while in (2.2) τ denotes the time effect between period 1 and 2. The residuals ε_j are assumed to be normally distributed with mean 0 and variance σ^2 ($\varepsilon_j \sim \mathcal{N}(0, \sigma^2)$). Note that both models implicitly assume that the time trend is equal across all arms. However, the first model (2.1) assumes that the time trend increases linearly with time, while the second one (2.2) models the time trend as constant in all periods.

Using these models to test the hypothesis $H_{0,2} : \theta_2 = 0$ against the one-sided alternative $H_{1,2} : \theta_2 > 0$, it is shown that both regression models improve the statistical power as

Lee and Wason investigated the operating characteristics of testing $H_{0,2}$ using (2.1) and (2.2) in trials with linear and unknown time trends equal across arms. Prof. M. Hoag et al. further investigated these models in trials with time trends that might vary between arms, and stated the conditions under which these lead to unbiased tests and TTE control.

2.1. Current Methods for Incorporating Non-concurrent Controls

compared to the separate analysis and control the type I error, if the functional form of the time trend is correctly specified, the assumption of equal time trends across arms holds and the time effect is additive on the model scale. Moreover, the model with stepwise adjustment (2.2) maintains the type I error even under misspecification of the functional form of the time trend (e.g., if this is linear instead of stepwise). ~~These models were further investigated by Bofill Roig et al. [2],~~ ^{show also} ~~again~~ in the context of trials with 2 treatment arms, ~~demonstrating~~ ^{when} that the type I error control is lost if the time trend in treatment arm 1 differs, and thus the assumption of equal time trends is violated.

simulation in

2.1.2. Bayesian Time Machine

so-called

A Bayesian method that adjusts for potential time trend in the platform trial, the Bayesian Time Machine, has been introduced by Saville et al. [3]. The Time Machine uses the division of the trial into calendar time intervals of equal length ("buckets"), which are indexed backwards in time, so that the most recent time interval is denoted by $c = 1$ and the time interval corresponding to the beginning of the trial by $c = C$.

This method has been discussed only in the context of binary endpoints, ~~so far~~ and can be described in terms of a generalized linear model as follows:

$$g(E(y_j)) = \eta_0 + \theta_{k_j} + \alpha_{c_j} \quad (2.3)$$

where y_j is the binary response for patient j and $g(\cdot)$ is the logit link function, which maps the expected value of the patient response to the linear predictors in the model. The model intercept η_0 denotes the response of the control group at time of the analysis, θ_{k_j} is the effect of the treatment arm k that patient j was enrolled in, relative to control. For the parameters η_0 and θ_{k_j} , normal prior distributions are assumed, with mean 0 and variances $\sigma_{\eta_0}^2$ and σ_{θ}^2 , respectively:

$$\eta_0 \sim \mathcal{N}(0, \sigma_{\eta_0}^2)$$

$$\theta_{k_j} \sim \mathcal{N}(0, \sigma_{\theta}^2)$$

In the time machine, the time is represented by c , which

it ~~is~~ α_{c_j} represents the change in the response for the time bucket c that corresponds to the enrollment of patient j , relative to $c = 1$ and is modelled using a Bayesian second-order normal dynamic linear model that creates a smoothing over the control response, such that closer time buckets are modelled with closer response rates:

this

$$\alpha_1 = 0$$

$$\alpha_2 \sim \mathcal{N}(0, 1/\tau)$$

$$\alpha_c \sim \mathcal{N}(2\alpha_{c-1} - \alpha_{c-2}, 1/\tau), \quad 3 \leq c \leq C$$

where τ denotes the drift parameter that controls the degree of smoothing over the time buckets and is assumed to have a Gamma hyperprior distribution:

2. Methods

$$\tau \sim \text{Gamma}(a, b)$$

Saville et al. [3] examined the operating characteristics of the Time Machine in a simulation study considering a scenario with $K = 5$ treatment arms, focusing only on evaluating the efficacy of the last treatment arm ($k = 5$) against the shared control arm (testing $H_{0,5} : \theta_5 = 1$ vs. $H_{1,5} : \theta_5 > 1$). They show that the Time Machine can lead to superior performance as compared to the frequentist model with categorical adjustment for time in scenarios with linear time trend. However, in situations with sudden changes in the time trend, the frequentist model-based approach is preferable, as the time effect is modelled independently for each time interval in this case.

Similarly to the frequentist model that adjusts for time by including it as a categorical variable, the Bayesian Time Machine relies on the assumption that the time trend affects all arms in the trial equally and is additive on the model scale. Hence there should be no interaction between treatment and time [4]. The Time Machine may also lead to inferior results as compared to the frequentist model in settings with little to no overlap between the treatment arms, as there is not enough data to provide a link between concurrent and non-concurrent controls. Moreover, the performance of the method also depends on the choice of the time buckets, the prior distributions and the values of their corresponding parameters, which need to be chosen individually for a given endpoint, disease and population.

2.1.3. Network Meta-analysis

Marschner and Schou [5] proposed to analyse platform trials using meta-analysis techniques, which allows to conduct both, treatment-control as well as treatment-treatment comparisons for non-concurrent arms. They argued that even though the design of a platform trial is usually adapted over time (e.g. by changing randomization allocations, patient recruitment, by adding and dropping treatment arms, etc.), the periods between these adaptations may still be viewed as separate trials with fixed design. In particular, the randomization and allocation ratio is preserved in each period, hence unbiased direct comparisons can be made between concurrent arms. Moreover, indirect comparisons between two arms across multiple stages can be performed by combining two direct comparisons to a common reference arm, that is concurrent to both arms of interest at some point (i.e. there is an overlap of the reference arm and the arms being compared). From this perspective, the platform trial can be viewed as a network of direct (concurrent) and indirect (non-concurrent) comparisons.

Let $\hat{\theta}$ be a vector containing all direct contrast estimates from all periods, i.e., period-wise estimates of treatment-control comparisons, and also of treatment-treatment comparisons of concurrent treatments. The proposed method linearly combines the estimators from $\hat{\theta}$ to obtain the network estimates $\hat{\theta}^{network}$ of the treatment-control comparisons. This effect estimator can then be written as:

$$\hat{\theta}^{network} = (\mathbf{X}^T \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W} \hat{\theta} \quad (2.4)$$

2.2. Extensions to the Frequentist Models for Treatment-Control Comparisons

where \mathbf{X} is a design matrix specifying the possible treatment contrasts in all periods the trial and \mathbf{W} is a weight matrix, chosen based on the treatment effect standard errors in each period. This weight matrix also quantifies the contribution of the direct evidence to the overall evidence.

As all previously described approaches for incorporating non-concurrent controls, the validity of the network meta-analysis relies on the assumption that the underlying difference between the compared arms remains the same for all direct and indirect comparisons over time. This assumption can be assessed using formal tests of inconsistency, which were developed in the context of meta-analysis. *originally*

2.1.4. Open Questions

Commentary -> [4]

I think a short summary here would be enough. i.e., there:

*→ conditions in common
→ time adjustment*

and then just a few lines to motivate your research :)

2.2. Extensions to the Frequentist Models for Treatment-Control Comparisons

We extend the frequentist models from [1] and [2] to platform trials with K experimental treatment arms ($K > 1$), indexed by entry order, and a shared control group (for illustration see Fig. 2.2). Moreover, we propose methods that adjust for time trend by using periods, as in [1] and [2], as well as using calendar time intervals as in [3]

and explore further modeling approaches

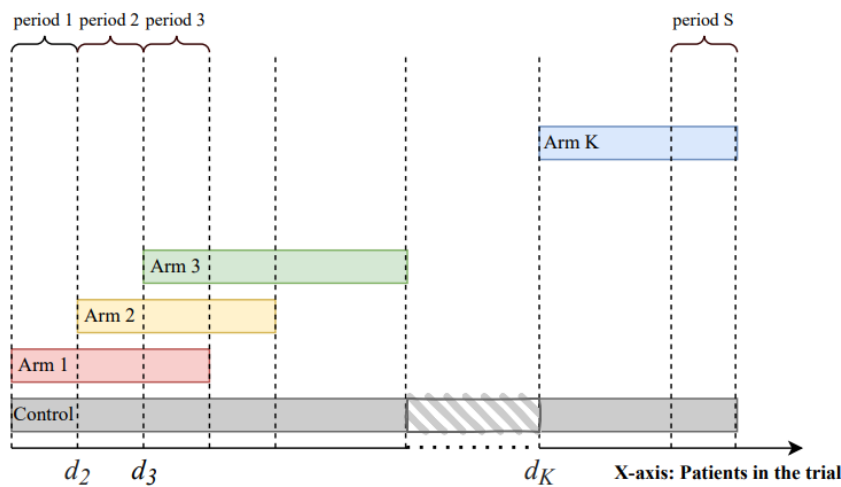


Figure 2.2.: Scheme of the considered general platform trial.

First, we divide the trial into periods, defined as time intervals bounded by adding or dropping experimental treatment arms. Formally, let \mathbf{T}_S be the set of all periods in the trial, $\mathbf{T}_S = \{T_{S1}, T_{S2}, \dots, T_{SS}\}$, where the s -th period is defined by

$$T_{S_s} = (\min\{\max_{\text{IP}(k_j=l')>0}\{t_j\}, \min_{\text{IP}(k_j=l)>0}\{t_j\}\}, \min\{\max_{\text{IP}(k_j=l')>0}\{t_j\}, \min_{\text{IP}(k_j=l+1)>0}\{t_j\}\}) \text{ for } s \in$$

*↑
{j: π_{kj} > 0}, where π_{kj} is the allocation rate of treatment k_j.*

2. Methods

Notation	Definition
K	number of experimental treatment arms
$k = 0, \dots, K$	arm indicator
N	total number of patients in the trial
$j = 1, \dots, N$	patient index
S	total number of periods
$s = 1, \dots, S$	period indicator
C	total number of calendar time units
$c = 1, \dots, C$	calendar time unit indicator
y	patient response
t	patient recruitment time
θ	treatment effect
λ	strength of the time trend
d	timing of adding treatment arms
M	currently evaluated treatment arm
S_M	period, in which arm M left the trial
C_M	calendar time unit, in which arm M left the trial
\mathcal{K}_M	set of treatment arms active in the trial prior or up to S_M or C_M

Table 2.1.: Summary of the used notation.

$\{1, \dots, S\}$ and $l, l' \in \{1, \dots, K\}$, $l \neq l'$. Here $\mathbb{P}(k_j = l)$ denotes the probability that patient j will be allocated to treatment l .

Dividing the trial into equidistant discrete units of calendar time \mathbf{T} (e.g. months, where $\mathbf{T} = 1$ would correspond to the first month of the trial) is given by the set of intervals $\mathbf{T}_{\mathbf{C}} = \{T_{C_1}, T_{C_2}, \dots, T_{C_C}\}$, where $T_{C_i} = (i - 1, i)$, for $i \in \{1, \dots, C\}$.

We summarize the notation in Table 2.1

We focus on evaluating treatment arms that enter when the trial is already ongoing and therefore NCC data is available for these arms. We aim at comparing the efficacy of each treatment against the control as soon as the treatment arm is completed. Consider the one-sided null hypothesis $H_{0,M} : \theta_M \leq 0$ for arm M under study, where θ_M denotes the treatment effect size for treatment M . To test the null hypothesis $H_{0,M}$, we propose model-based approaches adjusting for time as fixed or random factor, where time stratified either period or by calendar time. Our proposed models use all data from the trial until treatment M leaves the platform (i.e., all data in the set $\mathcal{D}_M = \{(y_j, k_j, t_j), j = 1, \dots, N; t_j \leq \mathbf{T}_M\}$, where \mathbf{T}_M denotes the time in which arm M finishes). Note that also data from unfinished arms is included to the models. **The considered models are summarized in Table 2.2.**

τ maybe + information?

2.2.1. Fixed-effects Models

Firstly, we consider two linear regression models that estimate the effect of treatments, which were active in the trial prior or up to the time unit \mathbf{T}_M and adjust for potential time trends by including time as a categorical covariate. The two approaches differ in the

2.2. Extensions to the Frequentist Models for Treatment-Control Comparisons

Name	Time definition	Time covariate	Section and reference
Fixed-effect model with period adjustment	period	fixed	Section 2.2.1 Equation (2.5)
Fixed-effect model with calendar time adjustment	calendar time	fixed	Section 2.2.1 Equation (2.6)
Mixed-effect model with period adjustment and uncorrelated random effects	period	random	Section 2.2.2 Equation (2.7)
Mixed-effect model with calendar time adjustment and uncorrelated random effects	calendar time	random	Section 2.2.2 Equation (2.8)
Mixed-effect model with period adjustment and autocorrelated random effects	period	random	Section 2.2.2 Equation (2.9)
Mixed-effect model with calendar time adjustment and autocorrelated random effects	calendar time	random	Section 2.2.2 Equation (2.10)

Table 2.2.: Summary of the proposed models.

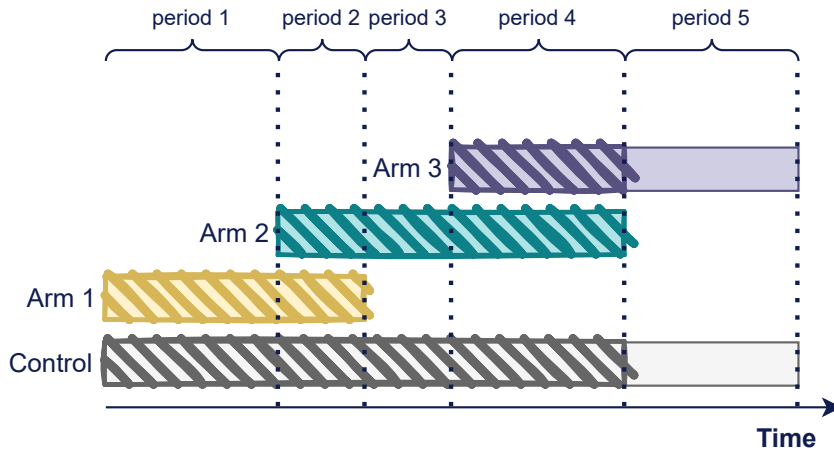


Figure 2.3.: Illustration of the data set \mathcal{D}_2 . Data taken into account for the evaluation of the 2nd treatment arm are highlighted.

2. Methods

way time intervals are defined in the model, one using periods and the other the calendar time.

Period Adjustment

In the first model, we adjust for the time effect by including the factor period to the model, i.e., using a step-function:

$$y_j = \eta_0 + \sum_{k \in \mathcal{K}_M} \theta_k \cdot I(k_j = k) + \sum_{s=2}^{S_M} \tau_s \cdot I(t_j \in T_{S_s}) + \varepsilon_j \quad (2.5)$$

where η_0 is the response in the control arm in the first period; θ_k represents the effect of treatment k compared to control for $k \in \mathcal{K}_M$, where \mathcal{K}_M is the set of treatments that were active in the trial during periods prior or up to \mathbf{T}_M ; τ_s indicates the stepwise period effect between periods 1 and s ($s = 2, \dots, S_M$), where S_M denotes the period, in which arm M left the trial (i.e. the period in which \mathbf{T}_M is included).

This approach is a direct extension of the model (2.2), investigated in [1] and [2], to platform trials with more than two treatment arms and periods. A difference here is that data from treatment arms that entered the trial later than the investigated arm M and are not finished yet is also used in the analysis. Since we assume that the time effect is equal across arms and constant over the period, as in [1] and [2], the model relies on the same assumption, that is, additivity on model scale and equal strength of the time trend across arms.

Calendar Time Adjustment

In the second approach, we consider a regression model adjusting for the time effect by using calendar time intervals, and thus estimating the calendar time effect rather than period effect:

$$y_j = \eta_0 + \sum_{k \in \mathcal{K}_M} \theta_k \cdot I(k_j = k) + \sum_{c=2}^{C_M} \tau_c \cdot I(t_j \in T_{C_c}) + \varepsilon_j \quad (2.6)$$

Here η_0 represents the control response in the first calendar time unit; θ_k denotes the treatment effect of treatment k compared to control for $k \in \mathcal{K}_M$, and \mathcal{K}_M is the set of treatments that were active up until the arm M left the trial. τ_c is the effect between calendar time units 1 and c ($c = 2, \dots, C_M$), where C_M indicates the calendar time unit, in which arm M left the trial.

The division of the trial into calendar times is depicted in Fig. 2.3. The length of these units is given in terms of the number of enrolled patients and can be specified arbitrarily.

PENDING - Add calendar times to Fig. 2.3!

Note that this model, similarly to the Bayesian Time Machine, divides time into bins of equal length. These bins are, unlike the period adjustment considered in [1] and [2], and

2.2. Extensions to the Frequentist Models for Treatment-Control Comparisons

in the extended model (2.5), independent of alternations to the trial design, i.e. adding and dropping arms is not taken into account.

As the adjustment in this case is done for each calendar time unit, it is assumed that the time effect is constant in each unit and equal across all arms. Moreover, the length of the time intervals poses an additional design parameter that allows to adjust also for shorter time intervals than are given by the periods.

2.2.2. Mixed Models

In models (2.5) and (2.6), time is treated as a fixed factor. Alternatively, patients within different periods or calendar time units could be considered as different clusters, having a period- or calendar time-specific random intercepts. In the next section, we include the time variable to the models as a random factor. In this way, possible correlation between the random effects associated with different periods or calendar times can also be taken into account.

what follows

Under such models, the potential

Mixed Models with Uncorrelated Random Effects

First, we consider simple mixed-effect models, where the effects of the given time intervals (period or calendar time units) are assumed to be uncorrelated with the effects of neighbouring intervals. The mixed-effect model with period adjustment has the following form:

$$y_j = \eta_0 + \sum_{k \in \mathcal{K}_M} \theta_k \cdot I(k_j = k) + \sum_{s=2}^{S_M} u_s \cdot I(t_j \in T_{S_s}) + \varepsilon_j \quad (2.7)$$

whereas the model adjusting for calendar time units is given by:

$$y_j = \eta_0 + \sum_{k \in \mathcal{K}_M} \theta_k \cdot I(k_j = k) + \sum_{c=2}^{C_M} u_c \cdot I(t_j \in T_{C_c}) + \varepsilon_j \quad (2.8)$$

where, y_j and θ_k has the same interpretation as in the fixed effect models. The model intercept η_0 is in this case given in terms of the control response across the whole trial up until S_M or C_M , respectively. This interpretation of the intercept corresponds to the Time Machine approach. u_s and u_c denote the random effect associated with the intercept for period s or calendar time unit c .

The period-specific random effects in (2.7) are assumed to be normally distributed with mean 0 and constant variance σ_{period}^2 :

$$\mathbf{u} \sim \mathcal{N}(0, \sigma_{period}^2 \cdot I_{S_M \times S_M})$$

The random effects in (2.8) associated with calendar time units are distributed analogously:

add some comments regarding the interpretation of the random effects as compared to prev. models.

2. Methods

I think you could also mention here some comments regarding the estimation

$$\mathbf{u} \sim \mathcal{N}(0, \sigma_{calendar}^2 \cdot I_{C_M \times C_M})$$

Note that in this case the correlation between any two period or calendar time effects is 0.

In both models, the distribution of the residuals ε_j , associated with the response of individual patient j is assumed to be the same for all treatments:

$$\varepsilon \sim \mathcal{N}(0, \sigma^2 \cdot I_{N \times N})$$

The random effects, u_s or u_c , and the residuals, ε_j , are assumed to be independent.

Mixed Models with Autocorrelated Random Effects

To account for possible correlation of the random effects, we also consider random effects with first-order autoregressive structure (AR(1)), again with period or calendar time adjustments. The model equations are identical to (2.7) and (2.8):

$$y_j = \eta_0 + \sum_{k \in \mathcal{K}_M} \theta_k \cdot I(k_j = k) + \sum_{s=2}^{S_M} u_s \cdot I(t_j \in T_{S_s}) + \varepsilon_j \quad (2.9)$$

$$y_j = \eta_0 + \sum_{k \in \mathcal{K}_M} \theta_k \cdot I(k_j = k) + \sum_{c=2}^{C_M} u_c \cdot I(t_j \in T_{C_c}) + \varepsilon_j \quad (2.10)$$

The model (2.9) considers period adjustments, while (2.10) adjusts for calendar times and the model parameters have the same interpretation as above.

There is a difference, however, with respect to the distribution of the random effects, which are now modelled as autocorrelated. The random effects for individual periods are assumed to be normally distributed with mean 0, constant variance σ_{period}^2 and an AR(1) correlation structure:

$$\mathbf{u} \sim \mathcal{N}(0, \sigma_{period}^2 \cdot \Sigma_{S_M \times S_M})$$

$$\Sigma_{S_M \times S_M} = \begin{bmatrix} 1 & \phi & \dots & \phi^{S_M-1} \\ \phi & 1 & \dots & \phi^{S_M-2} \\ \vdots & \vdots & \ddots & \vdots \\ \phi^{S_M-1} & \phi^{S_M-2} & \dots & 1 \end{bmatrix}$$

Analogous distribution is assumed for the random effects associated with different calendar times:

2.2. Extensions to the Frequentist Models for Treatment-Control Comparisons

$$\mathbf{u} \sim \mathcal{N}(0, \sigma_{\text{calendar}}^2 \cdot \Sigma_{C_M \times C_M})$$

$$\Sigma_{C_M \times C_M} = \begin{bmatrix} 1 & \phi & \dots & \phi^{C_M-1} \\ \phi & 1 & \dots & \phi^{C_M-2} \\ \vdots & \vdots & \ddots & \vdots \\ \phi^{C_M-1} & \phi^{C_M-2} & \dots & 1 \end{bmatrix}$$

The parameter ϕ denotes the correlation between two adjacent periods or calendar time units. Note that ϕ can range from -1 to 1 and the correlation of periods that are w units apart is equal to ϕ^w , such that the correlation is weaker for periods that are further apart.

The residuals are again assumed to be normally distributed (i.i.d.) with mean 0 and constant variance:

$$\varepsilon \sim \mathcal{N}(0, \sigma^2 \cdot I_{N \times N}) \quad (2.11)$$

2.2.3. Generalized Additive Models

The models in Sections 2.2.1 and 2.2.2 assumed an underlying linear relationship between the patient response and the time. In order to also capture potential non-linearity of the time trend, we consider estimating the patient response using a generalized additive model (GAM) with smoothing splines. The model is given by:

$$y_j = \eta_0 + \sum_{k \in \mathcal{K}_M} \theta_k \cdot I(k_j = k) + f(j) + \varepsilon_j \quad (2.12)$$

where y_j and θ_k are defined as in previous sections. Note that the treatment effect enters the model as a linear predictor.

the treatment effect for treatment k , which enters the model as a linear predictor. The time trend is modelled via a smooth function $f(\cdot)$, which gives the model additional flexibility as compared to the previously considered models. This smooth function is defined in terms of basis functions $b_i(j)$ as follows:

$$f(j) = \sum_{i=1}^k b_i(j) \beta_i$$

In this case, the patient index j , which also specifies their entry time, is used directly in the model and the time does not need to be divided in any way.

The residuals are normally distributed with mean 0 and variance σ^2 :

$$\varepsilon_j \sim \mathcal{N}(0, \sigma^2)$$

2. Methods

PENDING: discuss how the splines are penalized (i.e. choice of k , trade-off between model fit and degree of smoothness) and the selection of the smoothing parameter $\lambda \rightarrow$ GCV.

2.2.4. Toy Example

Add a simple example trial and compare all methods + code.

2.3. Treatment-Treatment Comparisons

 well done! getting there!