Evaluating the Use of Generalized Dynamic Weighted Ordinary Least Squares for Individualized Hiv Treatment Strategies

Larry Dong, Erica E. M. Moodie, Laura Villain, Rodolphe Thiébaut

▶ **To cite this version:**

Larry Dong, Erica E. M. Moodie, Laura Villain, Rodolphe Thiébaut. Evaluating the Use of Generalized Dynamic Weighted Ordinary Least Squares for Individualized Hiv Treatment Strategies. Annals of Applied Statistics, 2023, 17 (3), pp.2432-2451. 10.1214/22-aoas1726 . hal-04344321

HAL Id: hal-04344321
https://hal.science/hal-04344321

Submitted on 14 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

arXiv:2109.01218v1 [stat.AP] 2 Sep 2021

Evaluating the Use of Generalized Dynamic Weighted Ordinary Least Squares for Individualized HIV Treatment Strategies

Larry Dong[1,2], Erica E. M. Moodie[1], Laura Villain[2]
and Rodolphe Thiébaut[2]

[1]Department of Epidemiology, Biostatistics and Occupational Health, McGill University
[2]University of Bordeaux, INSERM U1219 Bordeaux Population Health, Inria SISTM

Abstract

Dynamic treatment regimes (DTR) are a statistical paradigm in precision medicine which aim to optimize patient outcomes by individualizing treatments. At its simplest, a DTR may require only a single decision to be made; this special case is called an individualized treatment rule (ITR) and is often used to maximize short-term rewards. Generalized dynamic weighted ordinary least squares (G-dWOLS), a DTR estimation method that offers theoretical advantages such as double robustness of parameter estimators in the decision rules, has been recently extended to now accommodate categorical treatments. In this work, G-dWOLS is applied to *longitudinal* data to estimate an optimal ITR, which is demonstrated in simulations. This novel method is then applied to a population affected by HIV whereby an ITR for the administration of Interleukin 7 (IL-7) is devised to maximize the duration where the CD4 load is above a healthy threshold (500 cells/μL) while preventing the administration of unnecessary injections.

1 Introduction

In precision or personalized medicine, the central paradigm lies in adopting patient-centric medical practices rather than disease-centric approaches [30]. Because people are intrinsically different, the effects of available treatments can be highly variable from one patient to another. Particularly in the clinical management of chronic illnesses, the evolving nature of ailments such as cancer, depression and substance abuse amongst many other long-term conditions often calls for treatments to be adapted to patient response, characteristics and overall well-being [3, 30].

Dynamic treatment regimes (DTR) are a statistical paradigm in precision medicine that optimizes an outcome of interest through sequential decision making. In general, the procedure of optimizing for multi-stage decision problem requires knowledge of the underlying data-generating mechanism. Although the mechanism is often unknown, it – or components of it – can be estimated using data-driven methods. This is where evidence-based medicine meets statistics, where the development of

a reliable theoretical framework is fundamental in finding optimal treatment regimes and in quantifying their uncertainty. The primary goals of DTR are twofold: comparing expected utilities of different deterministic treatments and obtaining optimal individualized treatment plans for patients [3]. Formally speaking, a DTR is a function that receives patient history as input and outputs an optimal decision vector which itself is comprised of one or multiple decisions, depending on the nature of the problem. At its simplest, a DTR is an estimation procedure that optimizes for a single decision; such single-stage DTR are referred as individualized treatment rules (ITR). One of the biggest strengths of the DTR framework is its ability to optimize long-term outcomes which involve medical interventions that are performed in a cascading fashion.

The primary objective of this current work is to demonstrate how generalized dynamic weighted ordinary least squares (generalized dWOLS or G-dWOLS), a DTR estimation method, can be used to optimize short-term or immediate outcomes from longitudinal data; we refer to such treatment regimes as myopic due to the nearsightedness of the maximization procedure. Although DTR is known for its ability to optimize multi-stage decision problems, we illustrate that developing myopic treatment regimes can be done for problems where it is reasonable to assume no delayed treatment effects. While such global optimal treatment strategies are preferable to locally optimal regimes, the benefits in solving for myopic treatment regimes are its simpler interpretation and a more lightweight statistical estimation procedure. By performing this methodological simplification, we remove the recursive process in the estimation of an optimal multi-stage DTR [3]. The underlying correlation between observations from a same individual can be accounted for by applying G-dWOLS within the generalized estimating equation (GEE) framework. We can accommodate the correlation between observations in the same individuals in a linear modelling framework; this works well with the G-dWOLS approach since it solves for parameters of interest in a weighted least squares (more details in section 3.2).

This methodological extension of G-dWOLS is motivated by the therapeutic benefits of Interleukin 7 (IL-7) injections in low immunological responders - HIV-infected individuals who are characterized by a low CD4 count despite a lack of viral load following antiretroviral treatment [41]. Clinical trials such as INSPIRE 2 and 3, whose protocol is available in Appendix A of the Supplementary Material, have been conducted to evaluate the effect of exogenous IL-7 injections on the immune system in low immunological responders who have been treated with antiretrovirals [16, 41]. Using a simulation study available in Appendix B of the Supplementary Material, we show that the theoretical properties of G-dWOLS still hold when devising an ITR with longitudinal data and multiple treatment categories. An application of our proposed method is then performed on data from INSPIRE clinical trials to see if tailoring treatment with respect to subject-level information would be clinically relevant.

Following the presentation of the INSPIRE data in section 2, a summary of our proposed method are stated in section 3. In section 4, we explain how G-dWOLS will be applied to data from the INSPIRE studies and data analysis results are elaborated in section 5. Lastly, the benefits and limitations of the

work are considered in section 6.

2 INSPIRE Trials and Related Work

HIV is characterized by a depletion of CD4+ T cells (CD4), which are a pillar of the immune system ([5]). Highly active antiretroviral therapy (HAART), the reference treatment for HIV, inhibits the replication of the virus and it is most often followed by an increase of CD4 T cells ([5]). It has been shown that failure to reconstitute of CD4 cells in HIV-infected individuals to a threshold greater or equal to 500 cells/μL is associated with a higher mortality rate and an increased risk in developing opportunistic infections ([17, 19, 23]). However, despite having no detectable viral load, 15% to 30% of individuals receiving HAART are known as poor or low immunological responders due to their inability to increase their CD4 counts and hence reconstitute their immunological system ([19, 11]).

Clinical trials INSPIRE 1, 2 and 3 were conducted to investigate the benefits of administering injections of IL-7 in low immunological responders by attempting to increase participants' CD4 count above 500 cells/μL [15, 42]. IL-7 is a cytokine that plays an essential role in the survival and maintenance of CD4 cells [14, 38, 39, 41]. While this protein is naturally produced in stromal and epithelial cells of the bone marrow and thymus, multiple studies conducted both in humans and in animals suggest that the administration of IL-7 leads to the reconstitution of the immune system [20, 33, 36, 39]. The main finding of INSPIRE 1 was that, after receiving injections of IL-7, participants exhibited improvements of many immunological markers including a sustained increase of CD4 cells [42]. However, a gradual decrease in CD4 counts followed which motivated the assessment of repeated cycles of IL-7 injections in INSPIRE 2 and 3. Three doses of IL-7 (10, 20 and 30 μg/Kg) were used in INSPIRE 1 whereas injections in INSPIRE 2 and 3 were all of 20 μg/Kg dosage due to its compromise in treatment effectiveness and side effects [15, 43]. In this study, only data from INSPIRE 2 and 3 were used due to the lack of baseline patient characteristics in INSPIRE 1.

INSPIRE 2 was a single-arm clinical trial where participants were drawn from an adult population of people living with HIV who have been receiving HAART for over a year and exhibiting suboptimal CD4 counts, i.e. between 100 and 400 cells/μL [12, 16]. The protocol entails providing three injections of IL-7 at one week intervals, monitoring patient T-cell response quarterly and providing another set of injections after 12 months of follow-up only if their CD4 count fell below 550 cells/μL [16, 42]. INSPIRE 3 was a clinical study where participants were randomized into a IL-7 arm and a control arm at a 3:1 ratio. With eligibility criteria defined similarly to those used in INSPIRE 2, the clinical protocol in the IL-7 arm entailed beginning patients with three injections, and readministrating three injections if patients presented a CD4 count $<$ 550 cells/μL at any quarterly evaluations. Patients in the control arm had their CD4 count measured for one year without any IL-7 injections. After one year, similarly to participants in the IL-7 arm, injections were administered if patients in the control arm presented a CD4 count below 550 cells/μL [42]. The discrepancy between the protocol CD4 threshold for treatment administration of 550 cells/μL and "healthy" threshold of 500 cells/μL accounts for possible measurement error and an anticipated immediate decrease in CD4 counts be-

low 500 cells/μL if participants exhibited CD4 levels between 500 and 550 cells/μL [41, 42]. From INSPIRE 2 and 3, a combined total of 113 patients received 198 sets of injections; two participants withdrew from the study. Over the course of the study, 1300 adverse events related to the administration of IL-7 have been recorded, although most of them were not deemed serious [42]. Of all reported occurrences, 77.6% were grade ≤ 1, 20.7% were grade 2 and 1.7% were grade ≥ 3; three patients experienced serious side effects and there was no significant difference in the number of adverse events for each visit [42]. A figure visualizing the INSPIRE protocols is available in Appendix A of the Supplementary Material.

The goal of the data analysis presented in section 4 is to devise a myopic treatment rule that optimizes the number of injections while preventing the administration of unnecessary ones using G-dWOLS [27]. There are costs – both financial and clinical (such as treatment fatigue and risk of adverse effects) – to injections, and so it is of interest to find the smallest number of injections needed to ensure that CD4 concentration lies above 500 cells/μL.

3 ITR Estimation for Longitudinal Data

ITR can be used in the optimization of long-term outcomes through the maximization of short-term or immediate rewards. The overarching idea behind this approach bears strong similarities with the greedy algorithm whereby locally optimal decisions are made in the attempt to maximize an overall or terminal outcome [13]. The essence of this method stems from the trade-off between the simpler formulation of the problem-solving framework and a potentially non-optimal solution that is relatively close to the best one [13].

The process of estimating optimal ITR, also known as myopic optimal rules, can involve less model-based extrapolation while yielding more precise (i.e. narrower) confidence intervals [27]. Rather than evaluating regimes through a sequence of decision rules, one of many ways to approach the study design to treat every stage as a single observation. In doing so, the sequential nature of interventions is eliminated and the optimization procedure ensues with the maximization of short-term rewards.

Of course, there are many examples where a greedy algorithm or a myopic treatment strategy may fail to provide a reasonable policy to optimize long-term outcomes. In Appendix C in the Supplementary Material), we outline a sufficient set of conditions – requiring assumptions on the immediacy of the effect of treatments and potential interactions between treatments over time – that are needed to suggest the myopic strategy is adequate, and relate these to the INSPIRE setting, which is the focus of this work.

3.1 Overview of Regression-Based Approaches for ITR Estimation

When working with ITR, positing simpler models such as linear regressions to estimate contrasts of utilities offers simpler interpretability of estimated parameters. A **blip** function denoted by $\gamma(\boldsymbol{x}, a; \boldsymbol{\psi})$ is defined to be the expected gain in outcome if treatment $A = a$ were to be chosen instead of a reference or baseline treatment a^{ref}. For instance, in the binary case where $A \in \{0, 1\}$, $a^{\text{ref}} = 0$ often represents an absence of treatment. Non-parametrically, this is given by

$$\gamma(\boldsymbol{x}, a; \boldsymbol{\psi}) = \mathbb{E}\left[Y\left(\boldsymbol{x}, a\right) - Y\left(\boldsymbol{x}, a^{\text{ref}}\right) \mid \boldsymbol{X} = \boldsymbol{x}, A = a\right]$$

where $Y(\boldsymbol{x}, a)$ denotes the counterfactual outcome under treatment a with covariates $\boldsymbol{x}$. In other words, the blip function characterizes the expected gain in the outcome upon providing treatment $A = a$ compared to the reference treatment a^{ref}. When modelling for the outcome variable Y, the model is often written as the sum of two expressions: the treatment-free model and the blip function, which are parametrized by the respective coefficients $\boldsymbol{\beta}$ and $\boldsymbol{\psi}$. The **treatment-free** model $G(\boldsymbol{x}; \boldsymbol{\beta})$ is the portion of the outcome model that is independent of the treatment:

$$\underbrace{\mathbb{E}\left[Y \mid \boldsymbol{X} = \boldsymbol{x}, A = a; \boldsymbol{\beta}, \boldsymbol{\psi}\right]}_{\text{outcome model}} = \underbrace{G(\boldsymbol{x}; \boldsymbol{\beta})}_{\text{treatment-free model}} + \underbrace{\gamma(\boldsymbol{x}, a; \boldsymbol{\psi})}_{\text{blip function}}.$$

The nomenclature for the models in the outcome model is due to the clear separation of linear expressions due to the terms which interact with the treatment variable of interest. In the binary case where $A \in \{0, 1\}$, the blip function can be understood as an expected gain in utility or reward when receiving treatment $A = 1$ compared to $A = 0$; it is assumed that $A = 0$ is the baseline or reference treatment. However, when more than 2 treatments are possible, adaptations to blip function and hence the estimating function must be made. Keeping the treatment-free model untouched, the sum of $m-1$ contrast terms will comprise the "new", multi-treatment blip function. In other words, a linear term $\boldsymbol{x}^{\psi}\psi_{\ell}$ must be posited for each non-reference treatment a_{ℓ}. When A is a categorical variable where $\mathcal{A} \equiv \{a_0, a_1, \ldots, a_{m-1}\}$, assume without loss of generality that a_0 is chosen to be the reference treatment. The blip function can be written as a sum of treatment contrasts, each of which represents an expected gain in utility compared to the counterfactual situation in which a_0 was the assigned treatment:

$$\gamma(\boldsymbol{x}, a; \boldsymbol{\psi}) = \sum_{\ell=1}^{m-1} \mathbb{1}_{a=a_{\ell}} \boldsymbol{x}^{\psi} \boldsymbol{\psi}_{\ell}$$

where $\mathbb{1}_{a=a_{\ell}} = 1$ if $a = a_{\ell}$. Linear contrast functions $\gamma_{\ell}(\cdot)$ are often posited for convenience, but, in theory, any functional form which equals 0 when evaluated at the reference treatment is permissible. ITR estimation methods such as G-dWOLS also calls for a treatment model, also referred as (generalized) propensity score or conditional treatment density function, denoted by $\pi(\boldsymbol{x}, a)$, which can be estimated using (multinomial) logistic regression [1, 32]:

$$\pi(\boldsymbol{x}, a; \boldsymbol{\alpha}) = \mathbb{E}\left[A = a \,|\, \boldsymbol{X} = \boldsymbol{x}; \boldsymbol{\alpha}\right] .$$

The treatment model parameters are estimated separately, then weights are constructed and plugged into the regression used to estimate outcome model parameters in a two-step estimation procedure. Whether in a dynamic treatment regime setting or, as is our focus here, a repeated ITR setting where a myopic regime is estimated via longitudinal data, it is typical to model the propensity score in each treatment interval independently, without taking into account within-person correlation. Under the assumption of no unmeasured confounding, this approach is justified. In this work, the three models detailed above – treatment, treatment-free and blip model – are respectively parameterized by coefficients $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\psi}$. Regression-based approaches for ITR estimation posit linear models for the latter two, i.e. $G(\boldsymbol{x}; \boldsymbol{\beta}) = \boldsymbol{x}^{\beta}\boldsymbol{\beta}$ and $\gamma(\boldsymbol{x}, a; \boldsymbol{\psi}) = \boldsymbol{x}^{\psi}\boldsymbol{\psi}$. Superscripts β and ψ are used to label explanatory variables with respect to their respective "submodel" within the outcome model, whereby $\boldsymbol{X}^{\beta}$ and $\boldsymbol{X}^{\psi}$ are typically subsets of subject-level covariates denoted by $\boldsymbol{X}$.

3.2 Generalized Dynamic Weighted Ordinary Least Squares

When working with ITR, positing simple models to model contrasts of utilities offers straightforward interpretability of estimated parameters [3]. Recent work has examined the use of statistical learning models such as decision trees [40] and deep neural networks ([18]). However, we will not pursue such approaches but rather focus on G-dWOLS, a regression-based approach in the estimation of treatment regimes. Other related DTR estimation methods include Q-Learning [22, 21], G-estimation [31] and dynamic weighted ordinary least squares (dWOLS) [44], the estimation method on which G-dWOLS builds upon Schultz and Moodie (2020) [34].

The G-dWOLS method is a weighted regression-based approach for the estimation of optimal treatment regimes first introduced by [44] and further extended by Schultz and Moodie (2020)[34]. This subsection highlights the theoretical underpinnings and practical considerations detailed in the latter article. Recent work in the dWOLS literature is now also able to handle survival outcomes [35] and continuous treatments [34], but this work focuses on categorical treatments. Other DTR estimation procedures for categorical treatments include Qi et al. (2020) [29] and Xue et al. (2020) [46]. The main advantages of using G-dWOLS compared to other methods are threefold: its relatively intuitive implementation, its ability to accommodate categorical treatments and its statistically robust estimation procedure (more details in the next subsection). G-dWOLS provides doubly-robust blip parameters by relying on having weights $w(\boldsymbol{x}, a)$ satisfying the balancing property below. This estimation procedure also has the potential of being less daunting in implementation due to the familiar function form of estimating function, as it is equivalent to minimizing a weighted least squares. This equivalence in methodological procedure allows a relatively straightforward implementation of the G-dWOLS algorithm using built-in regression functions with weights $w(\cdot)$ such that $\pi\left(\boldsymbol{x}, a\right) w\left(a, \boldsymbol{x}\right) = \pi(\boldsymbol{x}, a')w\left(a', \boldsymbol{x}\right)$ is satisfied for all $a, a' \in \mathcal{A}$ where $\pi\left(\boldsymbol{x}, a\right) = P\left(A = a \,|\, \boldsymbol{X} = \boldsymbol{x}\right)$ [34]. This equality stems from the theorem of the balancing property in dWOLS and G-dWOLS liter-

ature and a detailed proof of this is provided in Wallace and Moodie (2015) [44] for binary treatments and Schultz and Moodie (2020)[34] for categorical treatments.

According to the balancing property, doubly-robust blip parameters are ensured by using correctly specified propensity scores in the weights associated to each observation in the weighted least squares minimizing. The formulation of the theorem is general, in that the product $\pi(\boldsymbol{x}, a)\, w(\boldsymbol{x}, a)$ needs to bear the same value for all treatments $a \in \mathcal{A}$. Examples of multinomial weights adhering to the balancing property that will be discussed in this work include inverse probability of treatment (IPT), $w(\boldsymbol{x}, a) = \{\pi(\boldsymbol{x}, a)\}^{-1}$, and overlap weights, $w(\boldsymbol{x}, a) = \{\pi(\boldsymbol{x}, a) \cdot \sum_{\ell=1}^{m} \pi(\boldsymbol{x}, a_\ell)^{-1}\}^{-1}$. Notice that the overlap weights are in fact IPT weights divided by a stabilizing term $\sum_{\ell=1}^{m} \pi(\boldsymbol{x}, a_\ell)^{-1}$ that solely depends on $\boldsymbol{x}$. This dissimilarity in weight forms allows overlap weights to be bounded between 0 and 1, whereas IPT weights which are unbounded from above. In Schultz and Moodie (2020) [34], other weight forms for categorical treatments adhering to the balancing property are also provided, but will not be addressed in this paper as they offer no discernible advantage over the IPT and overlap weights.

Without the balancing weights, the G-dWOLS estimation process simplifies to Q-learning in a single-stage setting [44]. In practice, when using G-dWOLS for ITR estimation, the selection of the weights comes down to potential gain in efficiency and researcher's preference. Standard error estimates can vary depending on the form of weights since the asymptotic variance of blip parameters depend on $w(\cdot)$ [34, 44]. The G-dWOLS estimation procedure for ITRs is described in the five following steps:

1. Select (possibly identical) subsets of covariates $\boldsymbol{X}^{\alpha}, \boldsymbol{X}^{\beta}$ and $\boldsymbol{X}^{\psi}$ from $\boldsymbol{X}$ for the treatment model, the treatment-free model and the blip function respectively. We let covariate matrices $\boldsymbol{X}_i^{\beta}$ and $\boldsymbol{X}_i^{\psi}$ be of $(p \times 1)$ and $(r \times 1)$ dimension.
2. Propose a treatment model $\mathbb{E}[A \mid \boldsymbol{X}]$ and define a weight w such that the balancing condition is satisfied.
3. Posit treatment-free model $f(\cdot)$ in $\mathbb{E}[Y \mid \boldsymbol{X}, A] = f\left(\boldsymbol{X}^{\beta}\boldsymbol{\beta}\right) + \sum_{\ell=1}^{m-1} \mathbb{1}_{A=a_\ell} \boldsymbol{X}^{\psi}\boldsymbol{\psi}_\ell$. Typically, linear regressions are used due to their appealing statistical properties and simplicity, i.e. $f(\boldsymbol{X}^{\beta}, \boldsymbol{\beta}) = \boldsymbol{X}^{\beta}\boldsymbol{\beta}$.
4. Solve the following system of estimating equations to obtain the blip function parameter estimates $\widehat{\boldsymbol{\psi}}$:

$$\mathbf{0}_{(p+(m-1)r)\times 1} = \sum_{i=1}^{n} \begin{pmatrix} \boldsymbol{X}_i^{\beta} \\ \mathbb{1}_{A_i=a_1} \boldsymbol{X}_i^{\psi} \\ \vdots \\ \mathbb{1}_{A_i=a_{m-1}} \boldsymbol{X}_i^{\psi} \end{pmatrix} w_i \left(Y_i - \boldsymbol{X}^{\beta}\boldsymbol{\beta} - \sum_{\ell=1}^{m-1} \mathbb{1}_{A_i=a_\ell}\ \boldsymbol{X}^{\psi}\boldsymbol{\psi}_\ell \right).$$

   The solution to the score function is a $(p+(m-1)r)\times 1$-dimensional parameter vector: $\left(\boldsymbol{\beta}, \boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_{m-1}\right)$.

5. Evaluate the optimal treatment plan for each subject given estimates $\widehat{\boldsymbol{\psi}}$ as follow:

$$\widehat{A}^{\text{opt}}(\boldsymbol{x}^{\psi}) = \begin{cases} \underset{\ell \in \{1,\ldots,m-1\}}{\text{argmax}} \; \boldsymbol{x}^{\psi}\widehat{\boldsymbol{\psi}}_{a_\ell} & \text{if } \underset{\ell \in \{1,\ldots,m-1\}}{\max} \boldsymbol{x}^{\psi}\widehat{\boldsymbol{\psi}}_{a_\ell} > 0 \\ a_0 & \text{otherwise.} \end{cases}$$

3.3 Generalized Estimating Equation Framework

When solving multi-stage decision problems using a myopic regime study design, each participant i contributes n_i observations into an agglomerated dataset; in a clinical context, each observation can be referred to as patient-stage. Because there are many observations for a single individual, an intuitive assumption is to suppose that there is some degree of correlation between measurements contributed by the same person. Many statistical concepts surrounding generalized estimating equations (GEE) have been developed to accurately estimate parameters while acknowledging underlying correlation structures. This general framework provided in [4] allows statistical inference to be performed on linear coefficients $\boldsymbol{\beta}$ and $\boldsymbol{\psi} = \left(\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_{m-1}\right)$ under the assumption of different correlation structures for repeated measurements [4]. Notation and terminology are largely inspired from Chapter 4 in [4].

To account for the correlation of repeated measurements, different structures for the variance matrix can be assumed $V_i = \text{Var}(\boldsymbol{Y}_i)$, which is also referred as the working covariance matrix and may differ from the true variance structure [4]. The primary benefit in using a GEE framework is that the estimators of linear coefficients are asymptotically normal even when misspecifying the working covariance structure V_i [4]. This is particularly useful in situations where the underlying correlation structure is unknown. Based on previous research [45, 26], we advocate for the use of an independence working correlation, however an autoregressive structure could also be a reasonable choice.

Although longitudinal data call for a GEE or similar approach, in G-dWOLS, the procedure for estimating the standard errors of blip coefficients for an ITR also needs to incorporate the variability due to estimating generalized propensity scores (or parameters of a treatment model). The G-dWOLS method incorporates estimated rather than known weights into the estimation procedure of DTR. In practice, these weights are estimated values, which means that an adjustment to the robust standard error formula is required for them to provide valid inference. One way to circumvent this issue is to estimate standard errors via bootstrapping [6]. In G-dWOLS, bootstrap estimation of standard errors can be beneficial in that it accounts for the uncertainty attributable to the estimation of the generalized propensity score. However, it is also prone to suffer from small sample size issues since resampling must be performed on individuals rather than observations to retain the underlying correlation structure due to repeated measurements across bootstrap iterations.

3.4 Summary of Proposed Method

Combining elements from G-dWOLS and methods to account for repeated measurements, an overview of the proposed method for ITR estimation thesis is provided below.

For each individual i with $n_i \geq 1$ observations, given a sequence of outcome variables $\boldsymbol{Y}_i = \{Y_{i1}, \ldots, Y_{in_i}\}^\top$, categorical treatment $\boldsymbol{A}_i = \{A_{i1}, \ldots, A_{in_i}\}^\top$ with $A_{i\cdot} \in \mathcal{A} \equiv \{a_0, \ldots, a_{m-1}\}$ and time-varying subject-level characteristics $\boldsymbol{X}_i = (\boldsymbol{X}_{i1}, \ldots, \boldsymbol{X}_{in_i})^\top$, we wish to understand the treatment effects of each treatment option $a_j \in \mathcal{A}$ on the outcome Y. We first concatenate all observations into a single set denoted by $\mathcal{D} = \{(Y_{ij}, A_{ij}, \boldsymbol{X}_{ij})\}$. As detailed in section 3.1, we propose the following linear model for the outcome variable:

$$\mathbb{E}\left[Y_i \,|\, \boldsymbol{X}_i, A_i; \boldsymbol{\beta}, \boldsymbol{\psi}\right] = \boldsymbol{X}_i^\beta \boldsymbol{\beta} + \sum_{\ell=1}^{m-1} \mathbb{1}_{A_i = a_\ell} \boldsymbol{X}_i^\psi \boldsymbol{\psi}_\ell.$$

Compared to other similar methods in DTR or ITR estimation, dWOLS and G-dWOLS offer both double-robustness of blip estimators and relative ease in implementation given the estimating function is a weighted least squares [44]. Likewise, another key advantage of these estimation methods is that these approaches are regression-based. Instead of using the multi-stage framework in DTR to address longitudinal data, a simpler study design can be employed: a myopic ITR can be estimated in which repeated measurements can be accounted for by specifying a covariance structure. Combining both G-dWOLS balancing weights w_i adhering to the balancing property and proposed covariance structure $V_i \in \mathbb{R}_{n_i \times n_i}$, the estimating function in which $\boldsymbol{\psi}$ needs to be solved for is similar to the one provided in step 4 of section 3.2 in which $\text{diag}(w_i)$ is multiplied by V_i.

Estimation of standard errors can be done robustly using robust or sandwich estimator (although this ignores variability due to estimation of the balancing weights) or bootstrap resampling on individuals. By performing the resampling of individuals rather than observations, the correlation structure induced by the repeated measures is preserved. However, because the number of observations varies from one individual to another, sample sizes can also differ between bootstrap iterations.

4 Analysis of INSPIRE Data

For this work, interest lies in determining the smallest number of injections to ensure reconstitution of CD4 load above a healthy threshold of 500 cells/μL. In other words, we are interested in obtaining the optimal number of IL-7 injections by preventing the administration of unnecessary ones. Because the INSPIRE 2 and 3 protocol calls for injections if the CD4 load falls below 550 cells/μL at quarterly evaluations, the optimal number of injections will be determined at 90 day time intervals, which was the usual duration between two visits as per trial protocol [41].

4.1 An Outcome Balancing Benefits and Patient Inconvenience

The first challenge was to conceive a scalar-valued outcome consisting of a utility accounting for a patient's immune response and penalization for excessive injections, so that the outcome of interest was a single measurement representing an entire 90-day windows which we call a treatment stage. We also define such treatment stages to be time intervals where patients are eligible to receive injections, i.e. the first CD4 measurement of candidate treatment stages or CD4_1 must be below 550 cells/μL [42]. Given that participants are enrolled for at least one year, CD4 dynamics need to be estimated. For simplicity, we define $\text{CD4}(t)$ to be the estimated trajectory of patients cell count obtained via linear interpolation (see Appendix D in the Supplementary Material for an example). More sophisticated statistical methods such as mechanistic models [12, 28, 43] have been used to explicitly estimate the CD4 T-cell dynamics. In Figure 1, we illustrate an example of treatment stage labelling alongside an estimated CD4 trajectory for a given patient.

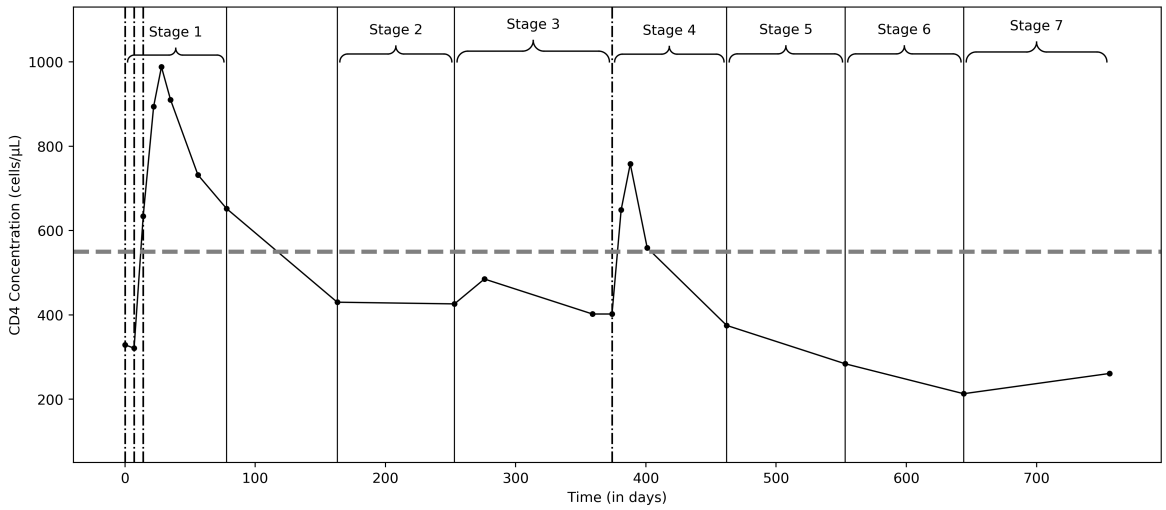

Figure 1: Estimated CD4 dynamics of an INSPIRE 2 participant using their observed CD4 counts with seven treatment stages labelled.

Example 4.1. Consider the CD4 dynamics of a patient from INSPIRE 2 with the labelling of treatment stages in Figure 1. Here, this patient receives two sets of injections: a first set consisting of three injections for treatment stage 1 spanning from day 0 to day 374 and a second set consisting of a single injection for treatment stage 4 spanning from day 374 to day 756. Both treatment intervals are split into multiple treatment stages: the first treatment interval contains 3 treatment stages whereas the second cycle contains 4 treatment stages. In the first treatment interval, the participant receives an injection cycle consisting of 3 injections in the first treatment stage and this is followed by 3 treatment stages where no injections are administered. In the second treatment interval, the participant receives an injection cycle consisting of a single injection in the first treatment stage (4th overall) and this is followed by 3 treatment stages with no injections. Notice that the interval between day 78 and day 163 does not constitute as a treatment stage because the CD4 count measured at day 78 is above 550 cells/μL.

An outcome of interest denoted by U^g that captures immune response is defined to be the proportion of time in a treatment stage where a participant had CD4 level above 500 cells/μL. By definition, $U^g \in [0, 1]$ and $U^g = 0$ if all observed measurements are below 500 cells/μL. The cost $U^{\text{inj}} \in \{-3, -2, -1, -0\}$ is defined to be the negative of the number of injections. In other words, the purpose of penalizing for the number of injections on top of a low CD4 cell count is to prevent the administration of superfluous injections. Inspired by Pasin et al. (2018) [25], we define the outcome variable of interest, denoted by $U(\eta)$, to be a convex sum of two previously defined utilities, U^g and U^{inj}:

$$U(\eta) = \eta U^g + (1 - \eta)U^{\text{inj}} \quad \text{for } \eta \in [0, 1].$$

The parameter η allows the focus of the constructed outcome variable to vary depending on the chosen value of η. Thus, an η of 0 would suggest that the utility is simply the negative of the number of injections, which would be maximized by never injecting participants. Conversely, setting η to 1 would generate a treatment rule designed to maximize CD4 response, without any consideration of the number of injections.

It is worth noting that the possible values that A^{opt} can take depends on η due to the definition of U. Because U^g is a proportion, i.e. $U^g \in [0, 1]$, we have that $U \geq 0$ under $A = 0$, hence the penalization term U^{inj} would restrict the domain of A^{opt} depending on η. For instance, under $A > 0$, we have that:

$$\max_{U^g} U(\eta) = \eta \cdot 1 + (1 - \eta)U^{\text{inj}} \leq \eta + (1 - \eta)(-1) = 2\eta - 1 < 0 \text{ for } \eta < 0.5\,.$$

Likewise, A =1 is only eligible to be the optimal number of injections for $\eta \geq 0.5$, A =2 for $\eta \geq 2/3$ and A =3 for $\eta \geq 3/4$ (see Appendix E for more details).

4.1.1 Q-TWiST Method

The results from the statistical analysis using ITR modelling will provide the ideal number of injections according to a patient's profile and response to previous injections. While we consider the potential onset of side effects important in assessing the necessity of each IL-7 injection, data related to the occurrence of adverse effects was not provided in this analysis. As such, a penalization term for the number of injections will serve as a surrogate for the risk of side effects in the statistical analysis of an optimal treatment rule.

The idea behind the definition outcome variable $U(\cdot)$ shares similar objectives with the Q-TWiST method, short for Quality-adjusted Time Without Symptoms of disease or Toxicity of treatment [9, 8]. The overarching goal behind this method is to define a single outcome which captures different utilities, all of which bears clinical relevance in evaluating the trade-off between quality and quantity of life. An outcome of interest to be investigated is formed from a weighted sum of three utilities: treatment toxicity, time spent devoid of symptoms or side effects and time after relapse [8]. The

difference in importance allocated to factors surrounding quality or quantity of life can be reflected by varying weights values attributed to the three utilities. For instance, patients who are more prone to adverse effects may put a stronger emphasis on symptom-free time whereas others may prefer stronger treatments if it decreases the chances of illness relapse. The Q-TWiST method provides a framework to compare different treatment options for patients with potentially different clinical preferences. Particularly in the clinical management of chronic illnesses and palliative care, a quality of life index is of prime importance in assessing various treatment possibilities for late-stage diseases [8]. In individualizing the number of IL-7 injections, rather than considering three utilities as in the Q-TWiST method, a single outcome capturing the two quantities U^g and U^{inj} can be optimized under fixed values of the utility weight parameter η.

4.1.2 Tailoring Variables

In practice, the idea behind longitudinal studies as a whole is to investigate the effects of studied interventions over time. In this analysis, we allow for previous immune responses to affect future ones in the modelling procedure by conditioning on individual patients' histories. That is, in addition to positing a correlation structure between treatment stages of the same patient, one way to account for historical injection information is to define tailoring variables that can embody potential prior biological response to the administration of IL-7. A patient's historical treatment information, denoted by Hx, is a dichotomous variable indicating if a patient has received injections in a prior treatment stage. A patient's response to previous treatment is denoted by Resp, and is defined as followed:

$$\text{Resp} = \begin{cases} 0 & \text{if Hx } = 0 \\ \frac{1}{\#\text{ prev inj}} \left(\max_k \left(\left\{ \text{CD4}_k^{\text{prev}} \right\}_{k=1}^{n_{\text{prev}}} \right) - \text{CD4}_1^{\text{prev}} \right) & \text{if Hx } = 1, \end{cases}$$

where the "prev" superscript refers to the most recent preceding treatment stage in which injections were administered to the individual. The k subscript indexes the set of observed CD4 counts in that "prev" treatment stage, hence $\text{CD4}_1^{\text{prev}}$ refers to its first CD4 reading and n_{prev} is the number of observations in that treatment stage. For instance, in Figure 1, "prev" refers to stage 1 for stages 2, 3, and 4 whereas, for stages 5, 6 and 7, "prev" refers to treatment stage 4. Resp is defined to be the largest increase in CD4 counts compared to the "prev" baseline value $\text{CD4}_1^{\text{prev}}$ and the multiplicative coefficient $\{\#\text{ prev inj}\}^{-1}$ serves the purpose of adjusting for the number of injections administered. The idea is to highlight the sensitivity of a participants' immune response to the quantity of IL-7 provided. Since Resp exhibited considerable skew, a transformation which we denote logResp = log(Resp + 1) was employed in the G-dWOLS analysis to reduce the impact of outlying values. Data on participants' characteristics collected in the INSPIRE studies 2 and 3 can be used to tailor the number of injections to HIV-infected patients within the G-dWOLS framework. In the outcome model, such quantities can be incorporated into the blip function, and include the following information: age, sex, BMI, ethnicity and logResp. A table summarizing these variables is given in Table 1 in section 4.2.1.

4.1.3 Analysis Plan

The statistical analysis of INSPIRE data using an ITR framework is summarized in the following protocol. In what follows, the subscript i indexes data points (pre-treatment covariates, number of injections, and utility) from a specific treatment stage rather than an study participant.

1. Compute U^g and U^{inj} for all patients and all stages. A value of 1 for the Origin variable indicates that a participant is of African origin and 0 otherwise (i.e. Caucasian or other); the Sex variable was also defined as dichotomous, where 1 represents male.
2. Fit generalized propensity scores $P(A_i = a_i \,|\, \boldsymbol{X}_i^{\alpha})$ where covariates $\boldsymbol{X}_i^{\alpha}$ consist of the following: Sex, Age, BMI, Origin and CD4_1. Weights w_i were taken to be IPT weights, i.e. $w_i = P(A_i = a_i \,|\, \boldsymbol{X}_i^{\alpha})^{-1}$. We use $A = 0$ as baseline treatment as it is the largest group.
3. Define the treatment-free and blip model covariates to include patient-specific data:
   - $\boldsymbol{X}^{\beta}$: Sex, Age, BMI, Origin, Hx, logResp and CD4_1
   - $\boldsymbol{X}^{\psi}$: Sex, Age, BMI, Origin, Hx and logResp
4. Compute $U(\eta) = \eta U^g + (1 - \eta) U^{\text{inj}}$ for a given weight value η for all individual-stage data. Separate analyses will be performed for each value of η.
5. Apply the G-dWOLS algorithm above using the defined IPT weights W, treatment-free covariates $\boldsymbol{X}^{\beta}$, blip covariates $\boldsymbol{X}^{\psi}$ and outcome variable $U(0)$ using a weighted GEE to perform estimation (here, we opt for an independence working correlation). We use ψ_ℓ for the vector of blip parameters associated with injections for $a \in \{a_1, a_2, a_3\} \equiv \{1, 2, 3\}$ (where, recall, $A = a_0 = 0$ is the baseline treatment category).
6. Estimate empirical standard errors and confidence intervals using the robust variance estimator for the blip parameters $\widehat{\boldsymbol{\psi}}$ obtained from the previous step.
7. Determine the optimal treatment (number of injections) for each patient using the formula provided in step 5 of section 3.2 where $a \in \{0, 1, 2, 3\}$.

Repeat the analytic steps 4-7 for other values $\eta \in [0, 1]$.

4.2 Results

4.2.1 Descriptive Results

Treating patient-stages as observations, a summary of its covariates is available in Table 1. The standardized mean difference (SMD) is a score that measures the imbalance of characteristics across observations in different treatment groups [7, 47]. An SMD value greater than 0.1 is a common criterion used to determine if a particular covariate is imbalanced across treatment groups [37]. For instance, in Table 1, both the Age and Sex variables have an SMD value of 0.12, the smallest SMD value amongst covariates of interest. This descriptive measure shows that there are significant imbalances in all characteristics across observations grouped by treatment category. It is also worth highlighting the low sample size in treatment categories $A = 1$ and $A = 2$. There are 315 observations associated with 0 injections and 150 observations associated with 3 injections whereas there are only 17 and 22 observations associated with 1 and 2 injections respectively. The low number of observations in

the $A = 1$ and $A = 2$ group is a by-product of the INSPIRE protocols where 3 injections should be have administered according to study protocol. The average treatment stage duration is 90.5 days with a standard deviation of 12.3 days. It is assumed that this variability does not substantially affect the analysis, especially in the definition of the immune response utility, where U^g is computed as a proportion.

Table 1: Summary of patient characteristics with respect to the number of injections received

Characteristic	Mean (SD) or Count (Proportion)				SMD
	$A = 0$ ($n = 315$)	$A = 1$ ($n = 17$)	$A = 2$ ($n = 22$)	$A = 3$ ($n = 150$)	
Sex	227 (77%)	12 (71%)	17 (77%)	102 (68%)	0.12
Age	45.4 (8.8)	46.3 (8.8)	44.5 (7.9)	44.9 (8.4)	0.12
BMI	24.4 (3.6)	25.0 (4.5)	25.8 (4.5)	24.3 (3.5)	0.21
Origin	0.40 (0.49)	0.29 (0.47)	0.32 (0.48)	0.48 (0.50)	0.22
$CD4_{init}$	350 (112)	435 (89)	358 (105)	322 (116)	0.56
logResp	3.60 (2.53)	4.89 (1.90)	2.97 (2.79)	2.01 (2.68)	0.64

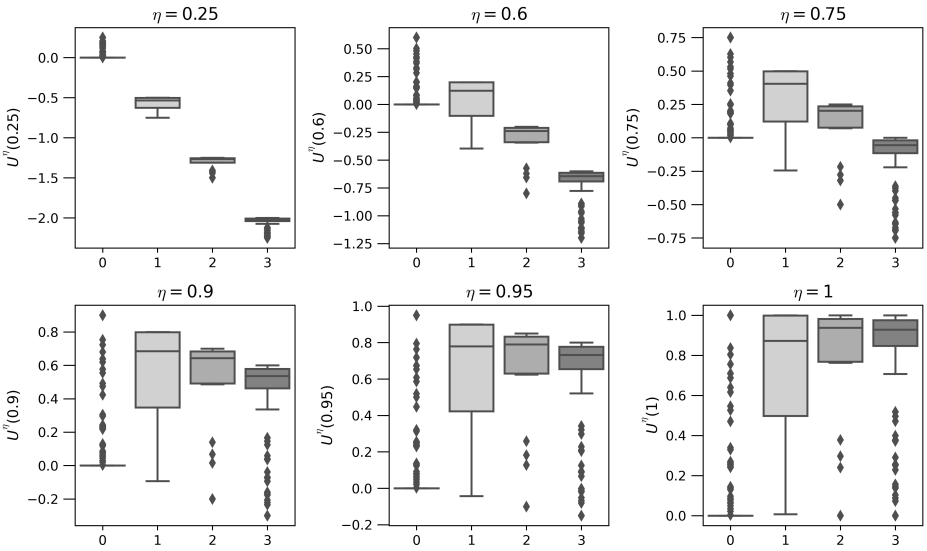

Figure 2: Boxplots for $U(\eta)$ values for $\eta \in \{0.25, 0.6, 0.75, 0.9, 0.95, 1\}$.

Figure 2 provides an overview of the outcome variable $U(\eta)$ distribution with respect to treatment categories and utility weights. For instance, for $\eta < 0.6$, the value of $U(\cdot)$ is largely dominated by the penalizing term U^{inj} for the quantity of administered injections. This result follows from the discrepancy in the widths of ranges of U^g and U^{inj}; $U^g \in [0, 1]$ whereas $U^{\text{inj}} \in \{-3, -2, -1, 0\}$. When more emphasis is put on immune response, i.e. in cases where η conveys a larger value such as 0.9,

0.95 or 1, the outcome values for 1, 2 and 3 injections are closer to each other. In the bottom three plots in Figure 2, utilities capturing immune response information do not seem to be considerably different in patient-stages where at least 1 injection was administered. U^g values for 0 injections are 0 for nearly all patient stages, which explains the lack of variability in the boxplot for $A = 0$.

Generalized propensity scores were fitted using multinomial logistic regression, which were then used to compute the weights w_i needed for the dWOLS analysis. Boxplots for generalized propensity scores $P(A = a \mid \boldsymbol{X})$, IPT and overlap weights are displayed in Appendix F of the Supplementary Material.

The use of either weight functions in G-dWOLS analysis can be argued. While both IPT weights and overlap weights adhere to the balancing property for G-dWOLS, a potential gain in efficiency is possible when using overlap weights due to the restriction in their range. However, the similarity in results produced by our simulation results (see Appendix B in the Supplementary Material) using the different weight forms permits the use of either weight function in the analysis of INSPIRE studies. In fact, IPT weighting for controlling of confounders has garnered much attention from researchers in the causal inference literature over the recent year. The popularity of this well-studied and well-documented method also motivate the use of IPT weights in the data analysis of INSPIRE data [2, 21, 32].

4.2.2 Statistical Summary of Blip Coefficients

Blip parameter estimates of $\widehat{\boldsymbol{\psi}}_1$, $\widehat{\boldsymbol{\psi}}_2$ and $\widehat{\boldsymbol{\psi}}_3$ and their confidence intervals are dependent on the outcome measure $U(\cdot)$ which itself relies on a particular choice of $\eta \in [0, 1]$. Each η value yields different values for the outcome variable, which in turn implies that the estimated ITR would vary accordingly. Because there are many coefficient estimates of interest and the utility weight is a continuous parameter, the overarching purpose of the statistical analysis and its interpretation is to paint a comprehensive portrait of the results while providing sufficient detail to determine key factors that influence treatment recommendation. The statistical analysis results for $\eta = 0.7$ and 0.9 are provided in Tables 2 and 3 respectively on the following page. Coefficient estimates are displayed alongside their 95% confidence intervals; bold estimates are statistically significant at a 5% significance level. From these summary tables, coefficients for Hx and Resp, denoted by $\widehat{\psi}_{\ell,\text{Hx}}$ and $\widehat{\psi}_{\ell,\text{logResp}}$, are statistically significant across contrast functions $\gamma_\ell(\cdot)$, $\ell = 1, 2, 3$ for both $\eta = 0.7$ and $\eta = 0.9$. Other coefficients such as $\widehat{\psi}_{2,\text{Sex}}$, $\widehat{\psi}_{3,\text{Sex}}$ and $\widehat{\psi}_{3,\text{Age}}$ are also statistically significant for $\eta = 0.7$ and $\eta = 0.9$. For residual plots, see Appendices G and H in the Supplementary Material for outcome variables $U(0.7)$ and $U(0.9)$ respectively.

The optimal number of injections can also be calculated for each patient-stage in the dataset used for the analysis. Estimation of contrast values $\gamma_\ell(\boldsymbol{x}^\psi)$ can be done by plugging in covariate-specific information $\boldsymbol{x}^\psi$ for each patient-stage in the processed dataset. $\widehat{A}^{\text{opt}}$ can be obtained by comparing values of the blip functions as discussed in section 3.2. The number of observations having $\widehat{A}^{\text{opt}}$ being equal to 0, 1, 2 or 3 injections are displayed with respect to η values ranging from 0 to 1 in

Table 2: Summary of estimated blip coefficients for a G-dWOLS analysis of outcome $U(0.7)$

Characteristic	Estimates (95% C.I.)					
	$A=1$		$A=2$		$A=3$	
Intercept	-0.11	$(-0.64, 0.41)$	-0.10	$(-0.83, 0.63)$	0.17	$(-0.11, 0.45)$
Sex	-0.02	$(-0.11, 0.06)$	**0.22**[a]	$(0.00, 0.43)$	**−0.06**	$(-0.12, 0.00)$
Age[b]	0.08	$(-0.01, 0.17)$	-0.07	$(-0.18, 0.04)$	**−0.04**	$(-0.07, 0.00)$
BMI[c]	-0.03	$(-0.16, 0.10)$	0.03	$(-0.13, 0.19)$	**−0.12**	$(-0.21, -0.04)$
Origin	0.14	$(-0.02, 0.29)$	0.11	$(-0.08, 0.30)$	-0.02	$(-0.08, 0.05)$
Hx	**−1.79**	$(-2.69, -0.89)$	**−1.35**	$(-2.15, -0.55)$	**−0.65**	$(-1.08, -0.21)$
logResp	**0.29**	$(0.13, 0.45)$	**0.24**	$(0.10, 0.39)$	**0.10**	$(0.03, 0.18)$

Table 3: Summary of estimated blip coefficients for a G-dWOLS analysis of outcome $U(0.9)$

Characteristic	Estimates (95% C.I.)					
	$A=1$		$A=2$		$A=3$	
Intercept	0.14	$(-0.54, -0.82)$	0.44	$(-0.50, 1.4)$	**1.08**	$(0.72, 1.44)$
Sex	-0.03	$(-0.14, 0.08)$	**0.28**	$(0.04, 0.56)$	**−0.08**	$(-0.20, 0.00)$
Age[b]	0.10	$(-0.01, 0.21)$	-0.09	$(-0.23, 0.06)$	**−0.05**	$(-0.10, 0.00)$
BMI[c]	-0.04	$(-0.20, 0.13)$	0.04	$(-0.17, 0.25)$	**−0.16**	$(-0.28, -0.05)$
Origin	0.18	$(-0.03, 0.38)$	0.14	$(-0.11, 0.38)$	-0.02	$(-0.11, 0.06)$
Hx	**−2.30**	$(-3.45, -1.14)$	**−1.73**	$(-2.77, -0.70)$	**−0.83**	$(-1.39, -0.27)$
logResp	**0.37**	$(0.17, 0.57)$	**0.31**	$(0.12, 0.50)$	**0.13**	$(0.033, 0.23)$

[a] **Bold** indicates statistical significance at a 0.05 significance level.
[b] For every 10 years
[c] For every 10kg/m^2

Figure 3. One important thing to observe is that, when $\eta = 1$, i.e. when there is no penalization for the number of injections, a considerable number of patient-stages are still being recommended 1 or 2 injections rather than 3. Although other factors such as a patient's age, sex, ethnic origin may affect the estimated ITR, recall that the observed immune response utility U^g seems to be comparable across groups $A = 1$, $A = 2$ and $A = 3$ (see Figure 2).

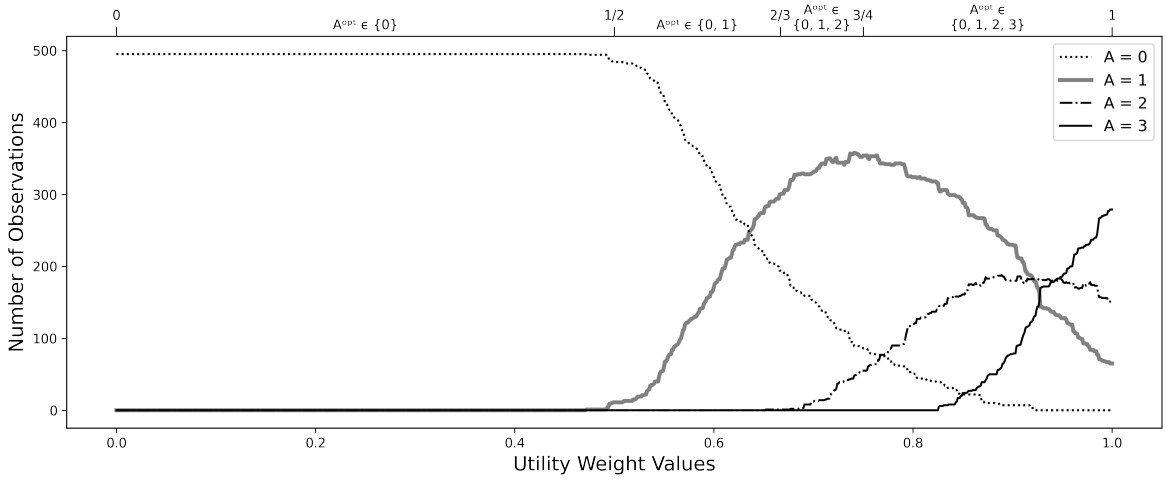

Figure 3: Number of observations having A^{opt} as each treatment type with respect to η values. Possible values of A^{opt} as a function of η are indicated on the top horizontal axis.

4.2.3 Treatment Recommendation for Specific Patient Profiles

Treatments depend on the chosen weight values associated with utilities U^g and U^{inj}; however the individualization of the treatment recommendation also depends on characteristics of patients, as the blip model is a function of covariates $\boldsymbol{X}^{\psi}$. To give an idea of the optimal number of injections, plots showing the estimated value of $\gamma_\ell(\cdot, a_\ell)$ for $a_\ell = 1, 2, 3$ for four chosen patient profiles are shown in Figure 4. The chosen patient profiles are as follow:

Profile 1: 25-year-old man of non-African ethnic background, with a BMI of 25 and without any IL-7 injection history;

Profile 2: 40-year-old woman of African ethnic background, with a BMI of 35 and with a recent injection cycle with a Resp value of 200;

Profile 3: 60-year-old man of African ethnic background, with a BMI of 24 and without any IL-7 injection history;

Profile 4: 80-year-old woman of non-African ethnic background, with a BMI of 30 and with a recent injection cycle with a Resp value of 400.

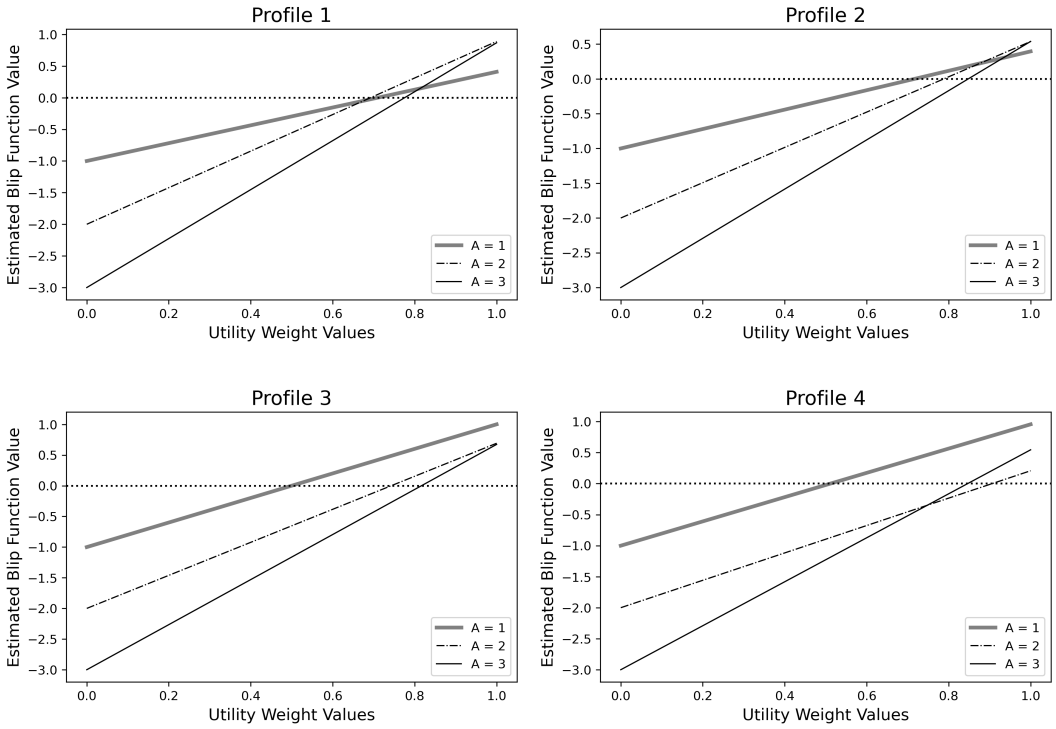

Figure 4: Contrast function utility for different number of injections with respect to utility weights $\eta \in [0, 1]$ for four patient profiles.

The dotted line representing the 0 utility threshold is the value of the blip function for the $A = 0$ category since it was chosen to be reference treatment. Because the blip function compares utilities across treatment options, the number of injections exhibiting the largest utility value for a fixed η value is the optimal treatment determined by the dWOLS analysis. Profile 1 seems to suggest that $A = 2$ is generally better than other treatment option counterparts while treatments seem to offer similar estimated utility values in Profile 2 (with $A = 2$ marginally better than $A = 1, 3$ for $\eta > 0.9$. Estimated contrast function plots for Profiles 3 and 4 show that a single injection seems to perform better than two and three injections for $\eta > 0.5$. For values of $\eta < 0.5$, no injections seem to be the recommended treatment for all four patient profiles. These results align well with the theoretical range of A^{opt} as discussed in Section 4.1 and in Appendix E. From these plots, three conclusions can be drawn.

Firstly, similarly to what [12] have found, there does not seem to be a considerable benefit in administering 3 injections. Instead, either 1 or 2 injections seem to result in comparable or even better immune response since, for $\eta = 1$, $A = 2$ seems to provide better utility than $A = 3$ for patient profiles 1 and 2 whereas $A = 1$ curves are above $A = 3$ for patient profiles 3 and 4. Secondly, although $A^{\text{opt}} = 0$ for $\eta < 0.5$, the decision to administer a non-zero number of injections seems to vary across patient profiles. The $A = 1$ line crosses the dotted line in profile 3 and 4 for a smaller value of η relative to the $A = 2$ line in profiles 1 and 2. The main difference between these pairs of patient profiles is their age: profiles 1 and 2 are respectively 20 and 40 years old whereas profiles 3 and 4 are

respectively 60 and 80 years old. This in turn can imply that, according to this analysis, a first IL-7 injection is more necessary in older participants than in younger counterparts but a second dosage is unnecessary. However, for younger patients, assigning 1, 2 or 3 injections yield comparable results in immune response; a moderate decision of 2 injections may be the best, most conservative approach.

5 Discussion and Conclusion

The goal of this statistical analysis was to show that G-dWOLS can be applied on longitudinal data within a GEE framework in constructing a myopic ITR. Overall, the G-dWOLS analysis provides valuable insight on the ideal number of injections through the design of ITR, one for each value of the utility weight η. Treatment recommendation largely depends on two things: the value of η and patient-specific information. When η is closer to 0, more weight is put on minimizing the number of injections and, as a result, the recommended number of injections is conservative. When η is closer to 1, more importance is put on having a better immune response, hence participants are more likely to be recommended to receive injections. However, despite this, many observations in the dataset are suggested to receive 1 or 2 injections. This result follows immediately from the good immune response outcome for observations associated with these treatment categories, despite their low sample size (see Figure 2). For a relatively large number of treatment stages, the maximal number of injections set by the clinical protocol are not being recommended; this warrants further investigation on the relevance of a third injection first; this question was previously raised by Jarne et al. (2017) [12]. While a third injection may increase the CD4 load to a higher peak, their results show that two injections are sufficient to ensure that the CD4 concentration exceeds the threshold of 500 cells/μL [12, 25].

Previous approaches to optimizing IL-7 treatment protocols have not incorporated covariate-specific information on the ideal number of injections to provide [25, 43]. In this work, intrinsic patient characteristics such as BMI and ethnic origin have not shown to be either statistically or clinically significant in individualizing the number of injections. Their the estimated coefficients did not significantly differ from 0, and do not seem to alter treatment recommendation (see Tables 2 and 3). The coefficient estimates for sex across treatment groups were also low in magnitude, except for $A = 2$; however, this finding was not found to have a strong impact on our results. Although our analyses were limited in power, our results suggest that age may be clinically useful tailoring variable to determine the optimal number of injections. For instance, the findings in the analysis of specific patient profiles recommend a first injection in older patients at lower utility weights than in younger patients. However, for the latter subpopulation, a second injection seems more necessary whereas, in older patients, treatment recommendation seems to follow a "one or nothing" approach to injections. This could potentially point to the risk of treatment toxicity in older people living with HIV, since the penalization term U^{inj} in this analysis served as a proxy for any potential undesirable consequences attributed to IL-7 injections, or simply to a plateauing of response in older patients. The effect of age on treatment recommendation warrants further investigation, particularly as adverse effects to IL-7 were not available to us for this analysis.

Two major limitations that must be raised are the sparsity of observations associated with 1 and 2 injections and the design of the myopic rule. When receiving a cycle of IL-7 injections, participants from the INSPIRE studies were more much more likely to receive 3 injections than they are to receive 1 or 2; cycles consisting of 1 or 2 injections are considered incomplete by the clinical protocols of INSPIRE 2 and 3 [42, 43]. An inherent limitation of this is that we cannot exclude confounding by indication due to participants not receiving all 3 injections, although we consider the possibility unlikely due to the nature of the treatment and outcome. Second, the design of the studied ITR assumes that the effect of IL-7 injections are short-term, in that their lingering influence on the CD4 count across subsequent stages are insignificant. The use of an ITR to investigate injection effects is motivated by the desire to maximize immediate utilities and its simpler interpretation [27]. However, an observable trend in the INSPIRE participants' CD4 dynamics is its gradual decrease over time following injections. An important model-based consideration in future work would to incorporate the duration since prior injections because, with increasing time, patients CD4 load are be more likely to fall below 500 cells/μL. Although we attempted to capture information regarding immune response from previous injections, the temporal aspect was not considered in this analysis.

The initial research question that inspired this study addresses the issue of an ideal number of IL-7 injections to provide to a patient exhibiting poor restoration of CD4 cells despite receiving HAART. The application of G-dWOLS requires the definition of an outcome variable that takes into account information regarding both a desirable immune response and the number of injections for an entire treatment stage. In our formulation of the statistical methodology, the outcome is defined as a sum of utilities which reflect both pieces of information through a weighting parameter η that varies between 0 and 1. As such, the optimal number of injections can only be determined for fixed values of this utility weight η. The choice of this parameter can depend on many aspects surrounding clinical outcomes of IL-7 administration, such as improvements in quality of life, treatment fatigue, side effects and treatment toxicity amongst many others. However, an ideal value of η cannot be determined by simply examining patient characteristics and medical history: it is a value that is subjective to a patient's preference or a medical decision maker's opinion. As argued in literature surrounding the Q-TWiST method, a single clinical measure accounting for multiple health factors provides a useful framework for medical decision-making [8, 10]. The ability to evaluate treatments by shifting focus or changing utility weight values with respect to clinical preference enables decision makers and especially patients to subjectively tailor medical care. The risk-benefit assessment of the Q-TWiST approach suggests that the optimal utility weight value η depends on the clinical implications of the statistical analysis and preferences of a patient [8]. In fact, it may even be the case that these preferences change with time. We posit that it is not possible to provide an optimal η value via data-driven methods. Rather, our analysis allows patients to determine optimal number of injections with respect to their personal (rather than statistical) valuation of the potential risks and benefits of IL-7 therapy, which may change over time as, for example, side effects are experienced and their severity and tolerability assessed.

In this work, we have provided the first ITR analysis of the 'dosing' of IL-7 using a myopic imple-

mentation of G-dWOLS applied to longitudinal data. Our analysis relied on several assumptions, including that of no unmeasured confounding which is plausible in the INSPIRE setting. The decision of an individual patient not to follow trial protocol and receive all three planned injections is highly unlikely to be associated with immune response. There have been some investigations into the use of random effects models to address unmeasured confounding [24], an avenue for future investigation in a repeated measures setting such as this. Another important direction for further research is to more fully determine the necessary and sufficient conditions under which the myopic approach may be used in place of a DTR analysis.

References

[1] AUSTIN, P. C. Assessing the performance of the generalized propensity score for estimating the effect of quantitative or continuous exposures on binary outcomes. *Statistics in Medicine 37*, 11 (2018), 1874–1894.

[2] AUSTIN, P. C., AND STUART, E. A. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Statistics in Medicine 34*, 28 (2015), 3661–3679.

[3] CHAKRABORTY, B., AND MOODIE, E. *Statistical methods for dynamic treatment regimes*. Springer: New York, USA, 2013.

[4] DIGGLE, P. J., HEAGERTY, P., LIANG, K.-Y., HEAGERTY, P. J., ZEGER, S., ET AL. *Analysis of longitudinal data*. Oxford University Press: New York, USA, 2002.

[5] DOUEK, D. C., ROEDERER, M., AND KOUP, R. A. Emerging concepts in the immunopathogenesis of AIDS. *Annual Review of Medicine 60* (2009), 471–484.

[6] EFRON, B. Bootstrap methods: another look at the jackknife. In *Breakthroughs in Statistics*. Springer, 1992, pp. 569–593.

[7] FLURY, B. K., AND RIEDWYL, H. Standard distance in univariate and multivariate analysis. *The American Statistician 40*, 3 (1986), 249–251.

[8] GELBER, R. D., COLE, B. F., GELBER, S., AND GOLDHIRSCH, A. Comparing treatments using quality-adjusted survival: the Q-TWiST method. *The American Statistician 49*, 2 (1995), 161–169.

[9] GELBER, R. D., GOLDHIRSCH, A., AND CAVALLI, F. Quality-of-life-adjusted evaluation of adjuvant therapies for operable breast cancer. *Annals of Internal Medicine 114*, 8 (1991), 621–628.

[10] GLASZIOU, P. P., COLE, B. F., GELBER, R. D., HILDEN, J., AND SIMES, R. J. Quality adjusted survival analysis with repeated quality of life measures. *Statistics in Medicine 17*, 11 (1998), 1215–1229.

[11] GRABAR, S., LE MOING, V., GOUJARD, C., LEPORT, C., KAZATCHKINE, M. D., COSTAGLIOLA, D., AND WEISS, L. Clinical outcome of patients with HIV-1 infection according to immunologic and virologic response after 6 months of highly active antiretroviral therapy. *Annals of Internal Medicine 133*, 6 (2000), 401–410.

[12] JARNE, A., COMMENGES, D., VILLAIN, L., PRAGUE, M., LÉVY, Y., THIÉBAUT, R., ET AL. Modeling CD4+ T cells dynamics in HIV-infected patients receiving repeated cycles of exogenous Interleukin 7. *The Annals of Applied Statistics 11*, 3 (2017), 1593–1616.

[13] KLEINBERG, J., AND TARDOS, E. *Algorithm design*. Pearson Education India, 2006.

[14] LAWSON, B. R., GONZALEZ-QUINTIAL, R., ELEFTHERIADIS, T., FARRAR, M. A., MILLER, S. D., SAUER, K., MCGAVERN, D. B., KONO, D. H., BACCALA, R., AND THEOFILOPOULOS,

A. N. Interleukin-7 is required for CD4+ T cell activation and autoimmune neuroinflammation. *Clinical Immunology 161*, 2 (2015), 260–269.

[15] LEVY, Y., LACABARATZ, C., WEISS, L., VIARD, J.-P., GOUJARD, C., LELIÈVRE, J.-D., BOUÉ, F., MOLINA, J.-M., ROUZIOUX, C., AVETTAND-FÉNOËL, V., ET AL. Enhanced t cell recovery in hiv-1–infected adults through il-7 treatment. *The Journal of Clinical Investigation 119*, 4 (2009), 997–1007.

[16] LEVY, Y., SERETI, I., TAMBUSSI, G., ROUTY, J., LELIEVRE, J., DELFRAISSY, J., MOLINA, J., FISCHL, M., GOUJARD, C., RODRIGUEZ, B., ET AL. Effects of recombinant human Interleukin 7 on T-cell recovery and thymic output in HIV-infected patients receiving antiretroviral therapy: results of a phase I/IIa randomized, placebo-controlled, multicenter study. *Clinical Infectious Diseases 55*, 2 (2012), 291–300.

[17] LEWDEN, C., CHÊNE, G., MORLAT, P., RAFFI, F., DUPON, M., DELLAMONICA, P., PELLEGRIN, J.-L., KATLAMA, C., DABIS, F., LEPORT, C., ET AL. HIV-infected adults with a CD4 cell count greater than 500 cells/mm3 on long-term combination antiretroviral therapy reach same mortality rates as the general population. *JAIDS Journal of Acquired Immune Deficiency Syndromes 46*, 1 (2007), 72–77.

[18] LIU, N., LIU, Y., LOGAN, B., XU, Z., TANG, J., AND WANG, Y. Learning the dynamic treatment regimes from medical registry data through deep Q-network. *Scientific reports 9*, 1 (2019), 1–10.

[19] LOGEROT, S., RANCEZ, M., CHARMETEAU-DE MUYLDER, B., FIGUEIREDO-MORGADO, S., ROZLAN, S., TAMBUSSI, G., BEQ, S., COUËDEL-COURTEILLE, A., AND CHEYNIER, R. HIV Reservoir dynamics in HAART-treated poor immunological responder patients under IL-7 therapy. *AIDS 32*, 6 (2018), 715–720.

[20] MACKALL, C. L., FRY, T. J., AND GRESS, R. E. Harnessing the biology of IL-7 for therapeutic application. *Nature Reviews Immunology 11*, 5 (2011), 330–342.

[21] MOODIE, E. E., CHAKRABORTY, B., AND KRAMER, M. S. Q-learning for estimating optimal dynamic treatment rules from observational data. *Canadian Journal of Statistics 40*, 4 (2012), 629–645.

[22] MURPHY, S. A. A generalization error for Q-learning. *Journal of Machine Learning Research 6*, Jul (2005), 1073–1097.

[23] OPPORTUNISTIC INFECTIONS PROJECT TEAM OF THE COLLABORATION OF OBSERVATIONAL HIV EPIDEMIOLOGICAL RESEARCH IN EUROPE (COHERE) IN EUROCOORD. CD4 cell count and the risk of AIDS or death in HIV-infected adults on combination antiretroviral therapy with a suppressed viral load: a longitudinal cohort study from COHERE. *PLoS Medicine 9*, 3 (2012).

[24] PAPADOGEORGOU, G., CHOIRAT, C., AND ZIGLER, C. M. Adjusting for unmeasured spatial confounding with distance adjusted propensity score matching. *Biostatistics 20*, 2 (2019), 256–272.

[25] PASIN, C., DUFOUR, F., VILLAIN, L., ZHANG, H., AND THIÉBAUT, R. Controlling IL-7 injections in HIV-infected patients. *Bulletin of Mathematical Biology 80*, 9 (2018), 2349–2377.

[26] PELS, W. A., ALAM, S., CARPP, L. N., AND MOODIE, E. E. A call for caution in using information criteria to select the working correlation structure in generalized estimating equations. *Epidemiology 29*, 6 (2018), e51–e52.

[27] PETERSEN, M. L., DEEKS, S. G., AND VAN DER LAAN, M. J. Individualized treatment rules: generating candidate clinical trials. *Statistics in Medicine 26*, 25 (2007), 4578–4601.

[28] PRAGUE, M., COMMENGES, D., DRYLEWICZ, J., AND THIÉBAUT, R. Treatment monitoring of hiv-infected patients based on mechanistic models. *Biometrics 68*, 3 (2012), 902–911.

[29] QI, Z., LIU, D., FU, H., AND LIU, Y. Multi-armed angle-based direct learning for estimating optimal individualized treatment rules with various outcomes. *Journal of the American Statistical Association 115*, 530 (2020), 678–691.

[30] RAMASWAMI, R., BAYER, R., AND GALEA, S. Precision medicine from a public health perspective. *Annual Review of Public Health 39* (2018), 153–168.

[31] ROBINS, J. M. Optimal structural nested models for optimal sequential decisions. In *Proceedings of the Second Seattle Symposium in Biostatistics* (2004), Springer, pp. 189–326.

[32] ROSENBAUM, P. R., AND RUBIN, D. B. The central role of the propensity score in observational studies for causal effects. *Biometrika 70*, 1 (1983), 41–55.

[33] ROSENBERG, S. A., SPORTÈS, C., AHMADZADEH, M., FRY, T. J., NGO, L. T., SCHWARZ, S. L., STETLER-STEVENSON, M., MORTON, K. E., MAVROUKAKIS, S. A., MORRE, M., ET AL. Il-7 administration to humans leads to expansion of CD8+ and CD4+ cells but a relative decrease of CD4+ T-regulatory cells. *Journal of Immunotherapy (Hagerstown, Md.: 1997) 29*, 3 (2006), 313.

[34] SCHULZ, J., AND MOODIE, E. E. Doubly robust estimation of optimal dosing strategies. *Journal of the American Statistical Association* (2020), 1–13.

[35] SIMONEAU, G., MOODIE, E. E., NIJJAR, J. S., PLATT, R. W., INVESTIGATORS, S. E. R. A. I. C., ET AL. Estimating optimal dynamic treatment regimes with survival outcomes. *Journal of the American Statistical Association* (2019), 1–9.

[36] SPORTÈS, C., HAKIM, F. T., MEMON, S. A., ZHANG, H., CHUA, K. S., BROWN, M. R., FLEISHER, T. A., KRUMLAUF, M. C., BABB, R. R., CHOW, C. K., ET AL. Administration of rhIL-7 in humans increases in vivo TCR repertoire diversity by preferential expansion of naive T cell subsets. *The Journal of Experimental Medicine 205*, 7 (2008), 1701–1714.

[37] STUART, E. A., LEE, B. K., AND LEACY, F. P. Prognostic score–based balance measures can be a useful diagnostic for propensity score methods in comparative effectiveness research. *Journal of Clinical Epidemiology 66*, 8 (2013), S84–S90.

[38] SURH, C. D., AND SPRENT, J. Regulation of mature T cell homeostasis. *Seminars in Immunology 17*, 3 (2005), 183–191.

[39] SURH, C. D., AND SPRENT, J. Homeostasis of naive and memory T cells. *Immunity 29*, 6 (2008), 848–862.

[40] TAO, Y., WANG, L., AND ALMIRALL, D. Tree-based reinforcement learning for estimating optimal dynamic treatment regimes. *The Annals of Applied Statistics 12*, 3 (2018), 1914.

[41] THIEBAUT, R., DRYLEWICZ, J., PRAGUE, M., LACABARATZ, C., BEQ, S., JARNE, A., CROUGHS, T., SEKALY, R.-P., LEDERMAN, M. M., SERETI, I., ET AL. Quantifying and predicting the effect of exogenous Interleukin-7 on CD4+ T cells in HIV-1 infection. *PLoS Computational Biology 10*, 5 (2014).

[42] THIÉBAUT, R., JARNE, A., ROUTY, J.-P., SERETI, I., FISCHL, M., IVE, P., SPECK, R. F., D'OFFIZI, G., CASARI, S., COMMENGES, D., ET AL. Repeated cycles of recombinant human interleukin 7 in HIV-infected patients with low CD4 T-cell reconstitution on antiretroviral therapy: results of 2 phase II multicenter studies. *Clinical Infectious Diseases 62*, 9 (2016), 1178–1185.

[43] VILLAIN, L., COMMENGES, D., PASIN, C., PRAGUE, M., AND THIÉBAUT, R. Adaptive protocols based on predictions from a mechanistic model of the effect of IL7 on CD4 counts. *Statistics in Medicine 38*, 2 (2019), 221–235.

[44] WALLACE, M. P., AND MOODIE, E. E. Doubly-robust dynamic treatment regimen estimation via weighted least squares. *Biometrics 71*, 3 (2015), 636–644.

[45] WANG, Y., MURPHY, O., TURGEON, M., WANG, Z., BHATNAGAR, S. R., SCHULZ, J., AND MOODIE, E. E. The perils of quasi-likelihood information criteria. *Stat 4*, 1 (2015), 246–254.

[46] XUE, F., ZHANG, Y., ZHOU, W., FU, H., AND QU, A. Multicategory angle-based learning for estimating optimal dynamic treatment regimes with censored data. *Journal of the American Statistical Association* (2020), 1–110.

[47] YANG, D., AND DALTON, J. E. A unified approach to measuring the effect size between two groups using SAS. In *SAS Global Forum* (2012), vol. 335, pp. 1–6.

Appendix A: INSPIRE Protocol

Patients assigned to the control arm of INSPIRE 2 only have CD4 and Ki67 evaluations in their "Induction" phase as they do not receive IL-7 injections until a year of follow-up.

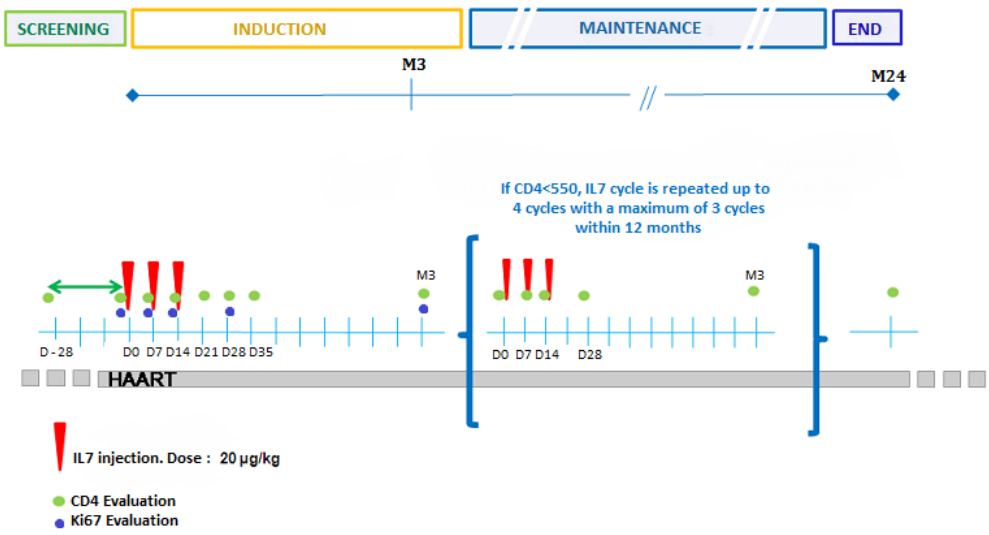

Figure 5: Protocols for INSPIRE 2 and IL-7 arm in INSPIRE 3 ([12, 43]). Three IL-7 injections, each at dose of 20μg/Kg, were provided within the first 3 weeks following enrolment into the study. After 3 months, study participants received another cycle of three injections if their CD4 count fell below 550 cells/μL. CD4 counts were measured at various times points and are displayed by green dots. Although measurement of the Ki67 markers was also taken and depicted by blue dots in the figure above, they were not relevant in our analysis.

Appendix B: Simulation Study

We carry out a simulation study to showcase the theoretical guarantees of the G-dWOLS analysis in devising a myopic treatment regime using longitudinal data. The goal of this simulation study is twofold; firstly, as presented in [44], we want to illustrate the double robustness of blip parameter estimators by conducting simulations with varying sample sizes under different specifications of the treatment-free and treatment model. Secondly, we show that the myopic formulation of the optimization problem is appropriate for longitudinal data under the assumption of no delayed treatment effects. We generate longitudinal data to mimic the evolving nature of the outcome variable and time-varying confounders on which the G-dWOLS algorithm can be applied. Intra-patient correlation due to the contribution of multiple observations to the processed data is also created in the simulated data. The presentation of simulation results is done in two parts: a summary of blip parameter estimates under different simulation designs is first depicted followed by an application of the simulated results on a "test" population to better understand the potential advantages of the G-dWOLS' theoretical guarantees from a practical perspective. That is, we examine bias and variability

of the parameters that together specify the estimated optimal ITR, and then evaluate the impact of that ITR on a hypothetical population of new patients.

Data Generation Procedure

We let $\mathcal{O}_{ij} = (\text{Sex}_{ij} \equiv \text{Sex}_i, \text{CD4}_{ij}, A_{ij}, Y_{ij})$ denote the jth stage-specific data of individual i as we generate four observations, each of which we refer as a stage, for each participant. We define Y_j to be some clinical outcome to be maximized and $A_j \in \{0, 1, 2\}$ to be a categorical exposure or treatment variable. Two covariates are defined: Sex_i, a time-invariant binary variable which denote a person's biological sex, and CD4_{ij}, a continuous time-varying variable which can represent a person's evolving CD4 cell count. Each individual is associated with four observations $\mathcal{O}_{\cdot j}$, as if these measurements were taken at each stage in a longitudinal study. The Sex variable is sampled from a Bernoulli distribution with $p = 0.7$ whereas the CD4 variable is sampled from a truncated normal distribution denoted by $\text{TRN}_{[50,550]}\,(350, 100)$ (centered around 350 with standard deviation 100 whose domain is bounded by 50 from below and 550 from above). The distribution parameters were chosen either be representative of "real life" or ideal in illustrating the theoretical properties of G-dWOLS; treatment effects and qualitative visualization of simulation results can vary depending on the specified model parameters. For instance, when parameter estimates are theoretically biased, the variation in CD4 counts needs to be somewhat substantial such that the bias is visually evident. The proportion of males in participants from INSPIRE studies 2 and 3 is around 70%, hence the p was specified to be 0.7 in the distribution of the Sex variable. The upper bound value of the truncated normal distribution was selected for context completeness because, in the INSPIRE protocol, patients are only eligible for treatment if their measured CD4 load is below 550 cells/μL. On the other hand, CD4 count is a strictly positive quantity and a lower bound of 50 cells/μL was chosen to reflect a lower limit of detection for testing. Treatment allocation probabilities p_0, p_1 and p_2 depend on parameters $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$, which were chosen such that a reasonable proportion of generated "people" are assigned to each treatment category. The outputs from Algorithm 1 will serve as inputs for the G-dWOLS estimation procedure which was provided in more detail in section 3.2. The assigned stage-specific treatment is sampled from a multinomial distribution whereby the probabilities p_0, p_1 and p_2 for treatment allocation are defined such that they can be estimated using a logistic regression model:

$$p_0 = \frac{1}{1 + \boldsymbol{X}^{\alpha}\boldsymbol{\alpha}_1 + \boldsymbol{X}^{\alpha}\boldsymbol{\alpha}_2}\,, \text{and}$$
$$p_\ell = \frac{\boldsymbol{X}^{\alpha}\boldsymbol{\alpha}_\ell}{1 + \boldsymbol{X}^{\alpha}\boldsymbol{\alpha}_1 + \boldsymbol{X}^{\alpha}\boldsymbol{\alpha}_2} \quad \text{for } \ell = 1, 2\,.$$

where $X_{ij}^{\alpha} = (1, \text{Sex}_{ij}, \text{CD4}_{ij})$. Although the probabilities of receiving assigned treatment a_0, a_1 or a_2 are in fact covariate dependent, the indices i and j are dropped to alleviate the notation. We define our outcome variable Y_{ij} as followed:

$$Y_{ij} = X_{ij}^{\beta}\boldsymbol{\beta} + \mathbb{1}_{A_{ij}=1}X_{ij}^{\psi}\boldsymbol{\psi}_1 + \mathbb{1}_{A_{ij}=2}X_{ij}^{\psi}\boldsymbol{\psi}_2 + b_i + \epsilon_{ij}$$

where $X_{ij}^{\beta} = \left(1, \exp\left\{\text{CD4}_{ij}/200\right\}, \sqrt{\text{CD4}_{ij}}\right)$ and $X_{ij}^{\psi} = (1, \text{Sex}_{ij}, \text{CD4}_{ij})$. Intra-subject correlation is generated by a time-invariant random effect b_i sampled from a normal distribution centered at zero with standard deviation 0.5 and a random effect ϵ_{ij} was sampled from a $\mathcal{N}(0, 3)$ distribution. In a first simulation study, parameters were chosen to be: $\boldsymbol{\alpha}_1$ = (-0.5, -0.2, 0.005), $\boldsymbol{\alpha}_2$ = (-1, -0.4, 0.007), $\boldsymbol{\beta}$ = (45, -10, 1), $\boldsymbol{\psi}_1$ = (-10, 5, 0.02) and $\boldsymbol{\psi}_2$ = (-30, -7, 0.1). The simulations are also conducted under the null, where there is no difference in expected utilities across all treatment categories. In other words, a second set of simulations is also performed with $\boldsymbol{\psi}_1 = \boldsymbol{\psi}_2$ = (0, 0, 0) and same values for $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ and $\boldsymbol{\beta}$ as in the first simulation. A summary of the data generation procedure is available in Algorithm 1.

Algorithm 1: Data Generation Procedure for Simulation Study

Input: Parameters $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\psi}$ and sample size n

for *i in {1, . . . , n}* **do**
  Sample $b_i \sim \mathcal{N}(0, 0.5)$
  for *j in {1, . . . , 4}* **do**
    if *j == 1* **then**
      Sample $\text{Sex}_{ij} \sim \text{Bern}(p)$
      Sample $\text{CD4}_{ij} \sim \text{TRN}_{[50,550]}(350, 100)$
    else
      Set $\text{Sex}_{ij} = \text{Sex}_{i1}$
      Sample $\Delta \sim \mathcal{N}(0, 5)$
      Set $\text{CD4}_{ij} = \text{CD4}_{i(j-1)} + \Delta$
      if $CD4_{ij} > 550$ **then**
        Set $\text{CD4}_{ij} = 550$
      else if $CD4_{ij} < 50$ **then**
        Set $\text{CD4}_{ij} = 50$
    end
    Sample $A_{ij} \sim \text{Multinomial}(p_0, p_1, p_2)$
    Define $X_{ij}^{\alpha} = (1, \text{Sex}_{ij}, \text{CD4}_{ij})$
    Define $X_{ij}^{\beta} = \left(1, \exp\left\{\text{CD4}_{ij}/200\right\}, \sqrt{\text{CD4}_{ij}}\right)$
    Define $X_{ij}^{\psi} = (1, \text{Sex}_{ij}, \text{CD4}_{ij})$
    Sample $\epsilon_{ij} \sim \mathcal{N}(0, 3)$
    Set $Y_{ij} = X_{ij}^{\beta}\boldsymbol{\beta} + \mathbb{1}_{A_{ij}=1} X_{ij}^{\psi}\boldsymbol{\psi}_1 + \mathbb{1}_{A_{ij}=2} X_{ij}^{\psi}\boldsymbol{\psi}_2 + b_i + \epsilon_{ij}$
  end
end
`/* Generating inputs for G-dWOLS estimation process of simulated data */`
Define $\textbf{Sex} = \{\text{Sex}_{ij}\}_{(i,j)\in\{1,\ldots,n\}\times\{1,\ldots,K\}}$
Define $\textbf{CD4} = \{\text{CD4}_{ij}\}_{(i,j)\in\{1,\ldots,n\}\times\{1,\ldots,K\}}$
Define $\boldsymbol{A} = \{A_{ij}\}_{(i,j)\in\{1,\ldots,n\}\times\{1,\ldots,K\}}$
Define $\boldsymbol{Y} = \{Y_{ij}\}_{(i,j)\in\{1,\ldots,n\}\times\{1,\ldots,K\}}$
return $\boldsymbol{Sex}, \boldsymbol{CD4}, \boldsymbol{A}, \boldsymbol{Y}$

Estimation Under Various Model Specifications

While parameters for the generalized propensity score, the treatment-free and blip models are all estimated in the G-dWOLS algorithm, interest predominantly lies in the blip coefficient estimates $\boldsymbol{\psi}$ as they provide information on the expected difference in utilities compared to baseline treatment $A = 0$. The specification of the set of covariates on which outcome regression and estimation of generalized propensity score are performed will dictate the asymptotic behaviour of $\widehat{\boldsymbol{\psi}}$ estimates. Recall that, in G-dWOLS, the double robustness property allows misspecification of at most one of the nuisance models to ensure consistency of blip parameter estimators.

- Model 1 (both treatment model and treatment-free model incorrectly specified):
  - Regress Y on 1, Sex, CD4 and treatments, with interactions between treatments and each of the two covariates.
  - Fit an intercept only multinomial logistic regression[1].
- Model 2 (correct treatment model, incorrect treatment-free model):
  - Regress Y on 1, Sex, CD4 and treatments, with interactions between treatments and each of the two covariates.
  - Fit a multinomial logistic regression on 1, Sex, CD4.
- Model 3 (incorrect treatment model, correct treatment-free model):
  - Regress Y on 1, $\exp\{\text{CD4}/200\}$, $\sqrt{\text{CD4}}$ and treatments, with interactions between treatments and each of the two covariate, where the interaction between treatment and CD4 is on the untransformed, linear CD4 scale.
  - Fit an intercept only multinomial logistic regression.
- Model 4 (both models correctly specified):
  - Regress Y on 1, $\exp\{\text{CD4}/200\}$, $\sqrt{\text{CD4}}$ and treatments, with interactions between treatments and each of the two covariates, where the interaction between treatment and CD4 is on the untransformed, linear CD4 scale.
  - Fit a multinomial logistic regression on 1, Sex, CD4.

Even though correlation was included in the data generation procedure to induce bias caused by repeated measurements, an independence correlation structure will be used in the G-dWOLS estimation process within the GEE framework. A correct specification of the covariance structure would yield more efficient estimators, but one of the main benefits of the GEE framework in accommodating longitudinal data is its asymptotic unbiasedness of linear estimates ([4]). The standard error estimation can be done via using the sandwich estimator or bootstrapping, but we used the former method to alleviate computational burden.

The simulation study was conducted on sample sizes $n = 100$ and $n = 1000$ to investigate sample properties of estimators. As mentioned in section 3.2, IPT and overlap weights both satisfy the balancing property, a feature of the dWOLS algorithm that ensures the double robustness of $\widehat{\boldsymbol{\psi}}$; both

[1] An intercept only multinomial logistic regression model does not estimate $P(A = a \mid \boldsymbol{X}^{\alpha})$ for each value of $\boldsymbol{X}^{\alpha}$ but rather a marginal or unconditional probability, i.e. $P(A = a \mid \boldsymbol{X}^{\alpha}) = P(A = a) = n^{-1}(\sum_{i=1}^{n} \mathbb{1}_{A_i = a})$.

weights were used in analyzing the simulated data. Each simulation design requires a choice of sample size n, G-dWOLS weight function (IPT or overlap) and one of four possible model specifications. To numerically assess the underlying properties of our statistical methodology, 1000 Monte Carlo replications were carried out for all possible simulation scenario combinations.

Simulation Results

The presentation of the results from the simulation study will be split into two parts: a summary of blip coefficient estimates followed by an application of these estimates on a completely novel or "test" set of 10 000 subject-level data. In a first instance, an overview of the variation and median of the $\widehat{\boldsymbol{\psi}}$ estimated values across the 1000 replications would provide a better idea on the effects of sample sizes, weight function choice and model specification on the blip coefficient estimates. Afterwards, the application of simulation results on an out-of-sample population would shed light on the quantitative benefits of the G-dWOLS estimation procedure in tailoring treatment to individual-level information.

Summary of Blip Coefficient Estimates – Each point estimate of $\boldsymbol{\psi}$ is obtained by applying the G-dWOLS algorithm on a generated dataset consisting of n subjects, for $n = 100$ or $n = 1000$ as mentioned previously. The results of the simulation study are showcased in Figures 6 and 7, which display boxplots summarizing the estimates across the 1000 runs for the non-null and null simulation settings. For purposes of notational simplicity, $\psi_{\cdot 1}$ is the blip model coefficient for Sex whereas $\psi_{\cdot 2}$ is that for CD4, with the $\cdot$ representing treatment level 1 or 2, relative to baseline level of treatment. Because similar conclusions can be drawn from the numerical results presented in both figures, the focus of this discussion will be on Figure 6.

The first row is parameters corresponding to treatment $A = 1$ whereas the second row corresponds to treatment $A = 2$. In each plot, there are boxplots grouped with respect to the four different combinations of model specifications as described on the previous page (Models 1, 2, 3 and 4). Labelled by their shading and hatch pattern, each set itself showcases the simulation results under different sample sizes and G-dWOLS weight form. The dotted horizontal line in every plot indicates the true value of the parameter used in the data generating process. The most important visual pattern across the plots is the bias of estimated values under Model 1. This result is not surprising, as G-dWOLS does not guarantee consistency when both nuisance models are incorrectly specified. However, it is important to point out that the estimates of ψ_{11} and ψ_{21} under Model 1 do not depart significantly from the true parameter value. In other words, unbiasedness can still be achieved but it is not guaranteed in Model 1. Other noticeable patterns are the decreasing interquartile range from $n = 100$ to $n = 1000$, which is evidence of a gain in efficiency with increasing sample size. The relative closeness between the median of estimates in Models 2, 3 and 4 and the true parameter values depicted by horizontal dotted lines suggests that unbiasedness can be reached in reasonable sample sizes. Although the simulation results show that finite sample performance improves with increasing values of n, $\boldsymbol{\psi}$ estimates at $n = 100$ perform well and data analysis results using similar sample sizes can yield good point estimates with reasonable variability provided that researchers are confident in the modelling choices. The difference in using IPT weights and overlap weights, which can be inferred by com-

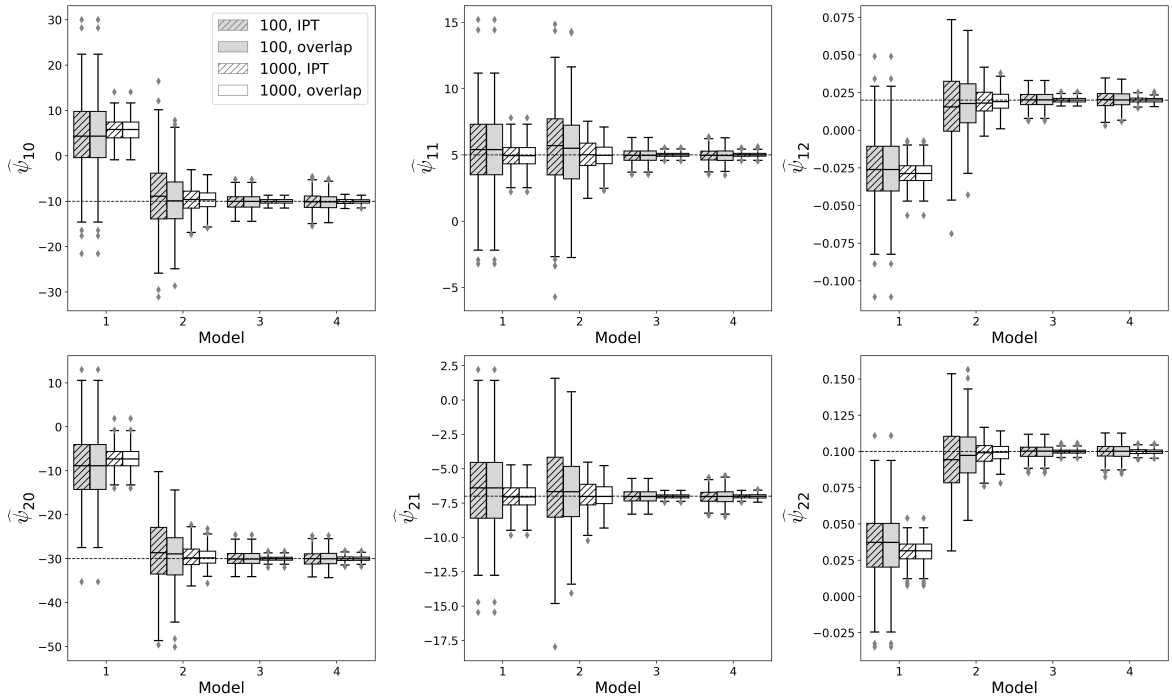

Figure 6: Boxplots of $\widehat{\psi}$ estimates are displayed for four different model specifications, for two different sample sizes ($n = 100, 1000$) and for two forms of balancing weights (IPT and overlap). The top row corresponds to the blip parameters for treatment level 1 (relative to level 0) and the bottom row to treatment level 2. The first column represents the main effect of treatment level, the middle and right columns are, respectively, the interactions of treatment with Sex and CD4.

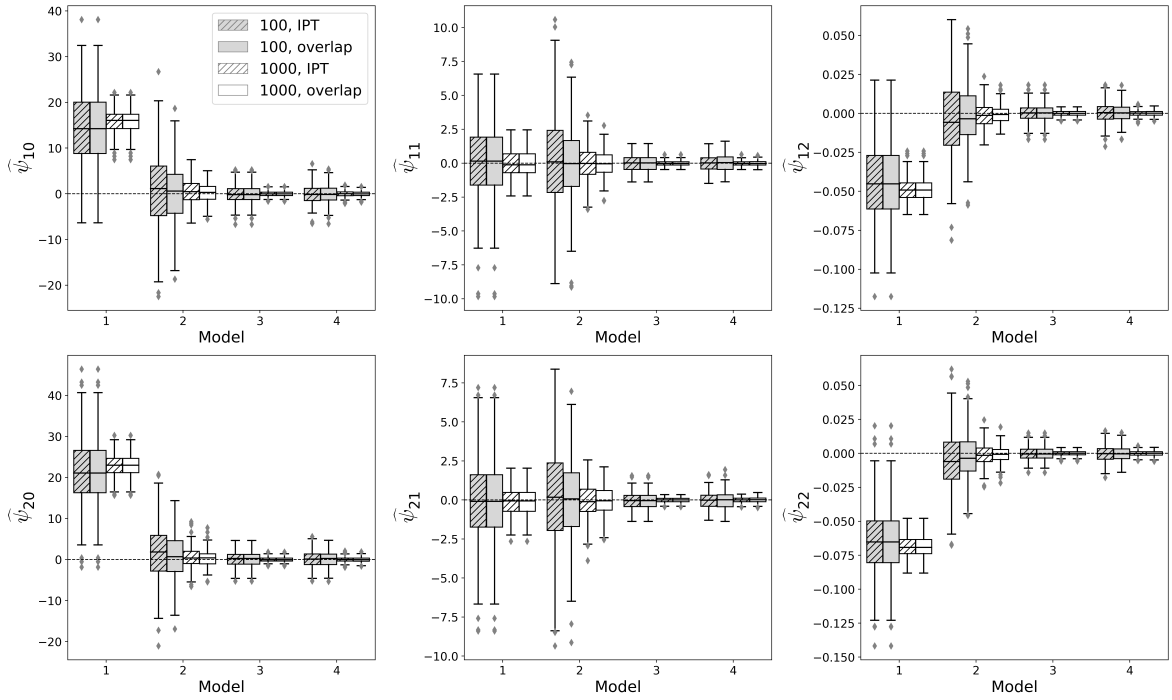

Figure 7: Boxplots of $\widehat{\psi}$ estimates are displayed for four different model specifications, for two different sample sizes and for two forms of balancing weights whereby all blip parameters set to 0 in the data generating process. Rows and columns are as in Figure 6.

paring boxplots with diagonal hatch patterns to the ones without such patterns, does not seem to be dramatic. That being said, the smaller variability in $\widehat{\boldsymbol{\psi}}$ estimates when using overlap weights suggest that having them bounded between 0 and 1 provide a more precise point estimate, since IPT weights are unbounded from above ([44]).

Average Utilities and Agreement Rates – A collection of data comprised of out-of-sample 10 000 subjects, which have not been used in the estimating process, is generated using Algorithm 1 as well to further examine and validate the results of the simulation study. Since this additional dataset was not used in obtaining estimates of blip function parameters, two comparisons which stem from this newly created set of potential participants are made. Firstly, the expected outcome value under uniform treatment (e.g. all 10 000 "test" people receive medical intervention $A = 0$, 1 or 2) are evaluated. Secondly, the utility under the estimated optimal treatment denoted by $\widehat{A}^{\text{opt}}$ depends on the value of $\widehat{\boldsymbol{\psi}}$ of a given iteration. The estimated ideal treatment $\widehat{A}^{\text{opt}}$ in patient-stages will be compared to the true values of A^{opt}, whereby the proportion of A^{opt} agreeing with $\widehat{A}^{\text{opt}}$ will be referred to as the agreement rate. In Figure 8, we show the average outcome value under $\widehat{A}^{\text{opt}}$ for the out-of-sample data which we define as $\overline{Y}^{\text{opt}} = \sum_{i=1}^{10\,000} \mathbb{E}[Y \mid \boldsymbol{X} = \boldsymbol{x}_i,\, A = \widehat{A}^{\text{opt}}; \widehat{\boldsymbol{\psi}}]$. Note that the true values of A^{opt} are known for all patient-stages since blip coefficient parameters are known by design in a simulation study, fixed by the choice of data-generating parameters.

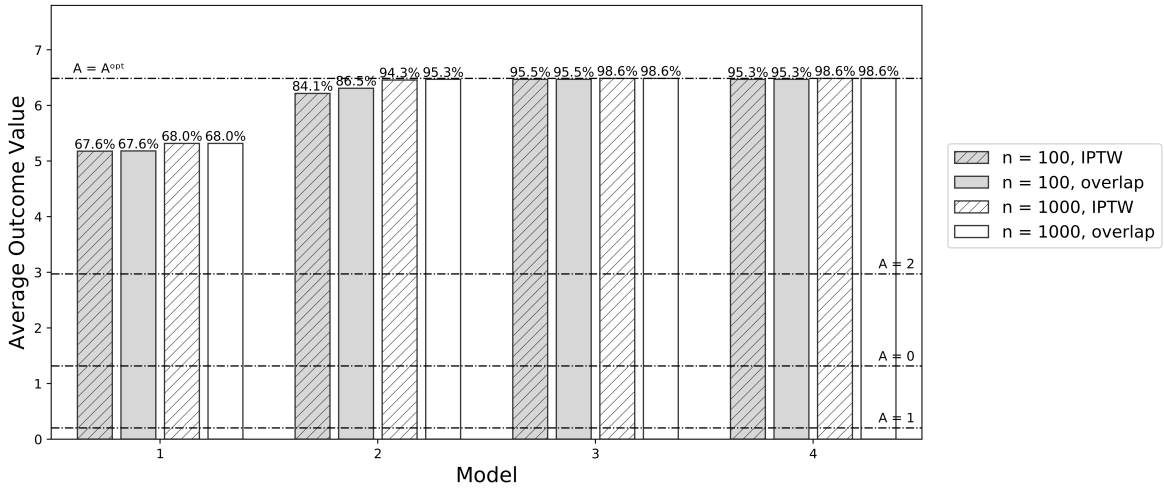

Figure 8: Average outcome under different model specifications using blip coefficients estimated with different sample sizes; the corresponding agreement rate is displayed on top of each bar. Dotted lines are the average utilities under uniform treatment, labelled respectively by the treatment options $A = 0$, 1 or 2; average utility under optimal treatment is indicated by A^{opt}.

In the absence of substantial variability in the outcome variable or treatment effects, a perfect agreement rate can be attained in drawing meaningful conclusions regarding the effect of sample size, forms of G-dWOLS weight functions and model specifications on the estimation of optimal treatment. The values of $\overline{Y}^{\text{opt}}$ are represented by the height of each bar in Figure 8 for various model specifications. The corresponding agreement rate for each simulation design is displayed on top of

the bars in the barplot. Results confirm that the selection of weight function for the G-dWOLS algorithm provides comparable results, although the overlap weights seem slightly better in Model 2 where only the treatment is correctly specified. Models 2, 3 and 4 show that larger outcome values and better agreement rates can be achieved in light of the correct specification of at least one of the treatment-free or treatment model. However, for $n = 100$, Model 3 and 4 seem to outperform Model 2, which suggests that the gain in efficiency can be slightly different in smaller sample sizes depending on which nuisance model is correctly specified. Nonetheless, the main conclusion that can be extracted from Figures 6 and 8 is that desirable results are achievable in relatively larger sample sizes given that at least one of the nuisance models is correctly specified and the chosen weight function adheres to the balancing property.

Appendix C: Suitability of the Myopic Strategy for the INSPIRE Setting

Suppose that $Y = Y_1 + Y_2$, i.e. the final, end-of-study outcome of interest in the dynamic treatment strategy is the sum of the stage-specific outcomes, Y_j, and consider the regret functions for a two stage setting:

$$\begin{aligned}
\mu_1(\boldsymbol{x}_1, a_1; \boldsymbol{\psi}) &= \mathbb{E}[Y(a_1^{opt}, a_2^{opt}(a_1^{opt})) - Y(a_1, a_2^{opt}(a_1))|\boldsymbol{X}_1 = \boldsymbol{x}_1, A_1 = a_1] \\
\mu_2(\boldsymbol{x}_2, a_2; \boldsymbol{\psi}) &= \mathbb{E}[Y(a_1, a_2^{opt}(a_1)) - Y(a_1, a_2)|\boldsymbol{X}_1 = \boldsymbol{x}_1, A_1 = a_1, \boldsymbol{X}_2 = \boldsymbol{x}_2, A_2 = a_2].
\end{aligned}$$

We begin with the second-stage optimal treatment. The optimal *dynamic* (rather than myopic) treatment strategy maximizes Y, and hence minimizes the regret:

$$a_{d,2}^{opt} = \arg\max_{a^* \in \mathcal{A}_2} \mathbb{E}[Y(a_1, a^*) - Y(a_1, a_2)|\boldsymbol{X}_1 = \boldsymbol{x}_1, A_1 = a_1, \boldsymbol{X}_2 = \boldsymbol{x}_2, A_2 = a_2].$$

As we condition on all pre-treatment covariates and the first stage treatment, note that Y_1 must be fixed and cannot be affected by choice of the second-stage treatment. Thus the counterfactual first-stage outcome, $Y_1(a_1, a_2)$, is a function of a_1 only. That is, we have that

$$\begin{aligned}
Y(a_1, a_2) &= Y_1(a_1, a_2) + Y_2(a_1, a_2) \\
&= Y_1(a_1) + Y_2(a_1, a_2).
\end{aligned}$$

Hence,

$$\begin{aligned}
&\arg\max_{a^* \in \mathcal{A}_2} \mathbb{E}[Y(a_1, a^*)) - Y(a_1, a_2)|\boldsymbol{X}_1 = \boldsymbol{x}_1, A_1 = a_1, \boldsymbol{X}_2 = \boldsymbol{x}_2, A_2 = a_2] \\
&\quad = \arg\max_{a^* \in \mathcal{A}_2} \mathbb{E}[Y(a_1, a^*))|\boldsymbol{X}_1 = \boldsymbol{x}_1, A_1 = a_1, \boldsymbol{X}_2 = \boldsymbol{x}_2] \\
&\quad = \arg\max_{a^* \in \mathcal{A}_2} \mathbb{E}[Y_2(a_1, a^*))|\boldsymbol{X}_1 = \boldsymbol{x}_1, A_1 = a_1, \boldsymbol{X}_2 = \boldsymbol{x}_2] \\
&\quad = \arg\max_{a^* \in \mathcal{A}_2} \mathbb{E}[Y_2(a_1, a^*)) - Y_2(a_1, a_2)|\boldsymbol{X}_1 = \boldsymbol{x}_1, A_1 = a_1, \boldsymbol{X}_2 = \boldsymbol{x}_2, A_2 = a_2].
\end{aligned}$$

This is sufficient to show that the optimal myopic treatment strategy at the second stage coincides with the optimal dynamic strategy at this stage, as the first expression defines the optimal dynamic strategy and the latter defines the optimal myopic strategy.
Let us turn now to the optimal dynamic treatment strategy in the first stage, defined by

$$\begin{aligned}
&\arg \max_{a^* \in \mathcal{A}_1} \mathbb{E}[Y(a^*, a_2^{opt}(a^*)) - Y(a_1, a_2^{opt}(a_1))|\boldsymbol{X}_1 = \boldsymbol{x}_1, A_1 = a_1] \\
&\quad = \arg \max_{a^* \in \mathcal{A}_1} \mathbb{E}[Y(a^*, a_2^{opt}(a^*))|\boldsymbol{X}_1 = \boldsymbol{x}_1] \\
&\quad = \arg \max_{a^* \in \mathcal{A}_1} \mathbb{E}[Y_1(a^*) + Y_2(a^*, a_2^{opt}(a^*))|\boldsymbol{X}_1 = \boldsymbol{x}_1].
\end{aligned}$$

The optimal myopic treatment strategy at the first stage maximizes $Y_1(a^*)$ over $a^* \in \mathcal{A}_1$; denote this $a_1^{opt,m}$ where the m indicates 'myopic'. If $a_1^{opt,m}$ is the maximizer of $Y_2(a^*, a_2^{opt}(a^*))$ over $a^* \in \mathcal{A}_1$, then clearly the myopic and dynamic treatment strategies at this first stage will also coincide. In what circumstances, then, does this hold? A sufficient set of conditions for the two strategies to coincide are (1) treatment effects are immediate, such that A_j impact Y_k for $k = j$ only, and (2) there are no synergistic or antagonistic effects between treatments at different stages. The latter condition implies that there are no statistical interactions between treatments, and hence that prior treatments do not act as tailoring variables. In the INSPIRE study, these sufficient conditions are biologically plausible because, when including lagged injections as a tailoring variable in the blip function on top of covariates detailed in Section 4.3, their statistical coefficients are not statistically significant at a 5% level for selected values of η, i.e. η =0.7 and 0.9 as we used in our analyses.

Appendix D: Example of Linear Interpolation

Here, we provide an illustrative example on how we construct our outcome variable U^g in the adaptation of the INSPIRE data for the analysis of ITR.

Example 5.1. Consider a patient's CD4 dynamics in Figure 9 in which they have three observations. Two injections were provided in this treatment stage and they are represented by the two vertical dotted lines. The three CD4 measurements are 400, 700 and 500 taken at respectively day 0, 24 and 80.

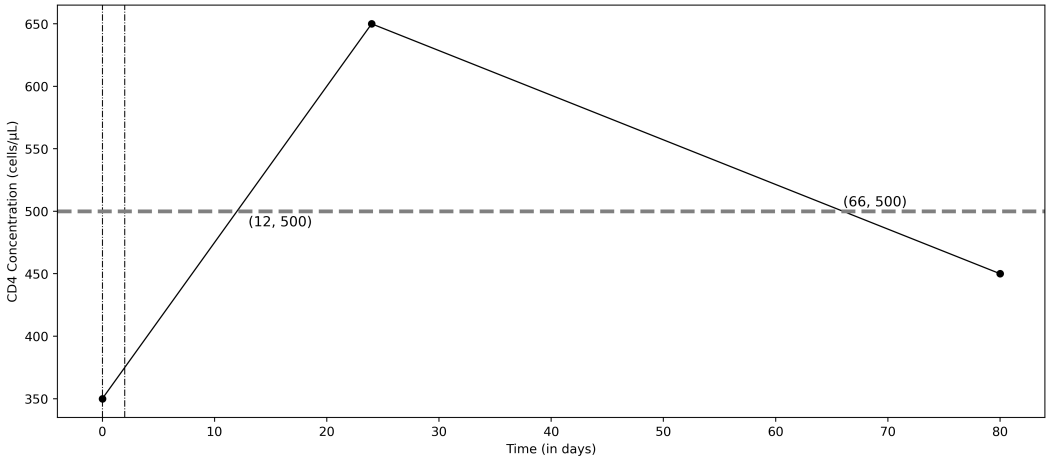

Figure 9: Estimation of CD4 dynamics using linear interpolation.

Because two injections were administered, U^i = -2. The estimated CD4 trajectory in Figure 9 crosses the 500 cells/μL threshold at time points $t \in \{12, 66\}$. The utility associated to immune response can be calculated as follows:

$$U^g = \frac{1}{80}\int_0^{80} \mathbb{1}_{\text{CD4}(t)\geq 500}\,\mathrm{d}t = \frac{1}{80}\int_{12}^{66} 1\,\mathrm{d}t = 0.55$$

Appendix E: Range of A^{opt} with respect to $\eta \in [0, 1]$

Recall that our outcome is defined as $U(\eta) = \eta U^g + (1-\eta)U^{\text{inj}}$, where $U^g \in [0, 1]$ and $U^{\text{inj}} \in \{0, -1, -2, -3\}$. We can evaluate the range of A^{opt} by comparing $U(\eta)$ under $A = k$ for $k = 1, 2, 3$ to 0, the minimal value of $U(\eta)$ under $A = 0$. In other words, because the utility is a non-negative value if no injections are administered, $A = k$ cannot be preferred for certain values of η due to the penalization of U^{inj}.

$$\begin{aligned}
\max_{U^g}\left\{\eta U^g + (1-\eta)U^{\text{inj}}\right\} \leq 0 \qquad & \text{under regime } A = k \\
1 + (1-\eta)(-k) &\leq 0 \\
\eta &\leq \frac{k}{1+k}
\end{aligned}$$

This implies that, for $k = 1$, $U(\eta) \leq 0 \Leftrightarrow A^{\text{opt}} \neq 1$ for $\eta \leq \frac{1}{2}$. Likewise, the same logic can be applied for $k = 2$ and $k = 3$. As a result, we have that:

$$
\begin{aligned}
&A^{\text{opt}} \in \{0\} \quad \text{for } \eta \in \left[0, \frac{1}{2}\right]; \\
&A^{\text{opt}} \in \{0,1\} \quad \text{for } \eta \in \left(1/2, 2/3\right]; \\
&A^{\text{opt}} \in \{0,1,2\} \quad \text{for } \eta \in \left(2/3, 3/4\right]; \\
&A^{\text{opt}} \in \{0,1,2,3\} \quad \text{for } \eta \in \left(3/4, 1\right].
\end{aligned}
$$

Appendix F: Boxplot of Generalized Propensity Scores and Weights

As expected, participants have a higher likelihood to receive 3 or no injections compared to receiving 1 or 2 injections in an injection cycle. This follows immediately from the small sample size of patient-stages where only 1 or 2 injections were administered as a cycle. Likewise, in the boxplots of IPT and overlap weights, it is shown that observations associated with treatment groups $A = 1$ or $A = 2$ receive larger weight values in the dWOLS analysis to accommodate for their underrepresentation in the dataset.

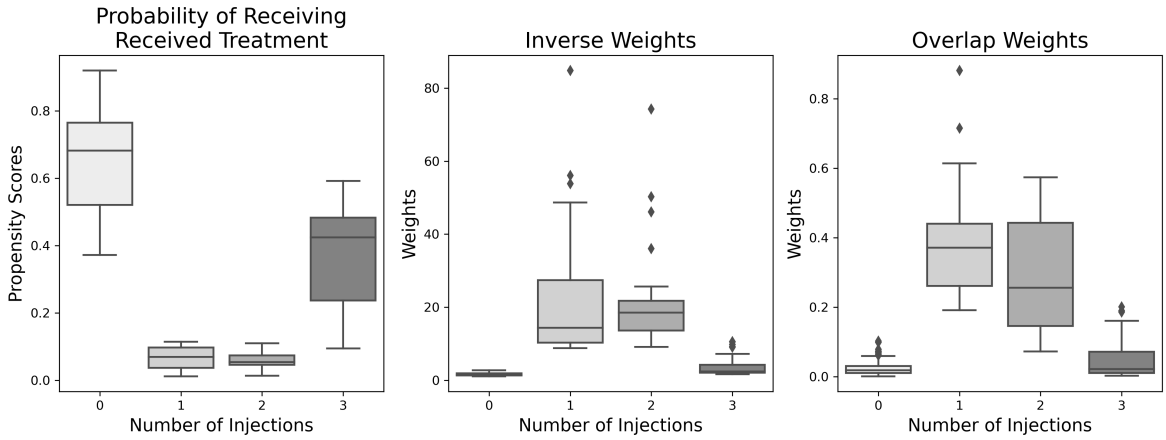

Figure 10: Boxplot of generalized propensity scores, inverse weights and overlap weights

Appendix G: Residual Plots for Outcome Variable $U(0.7)$

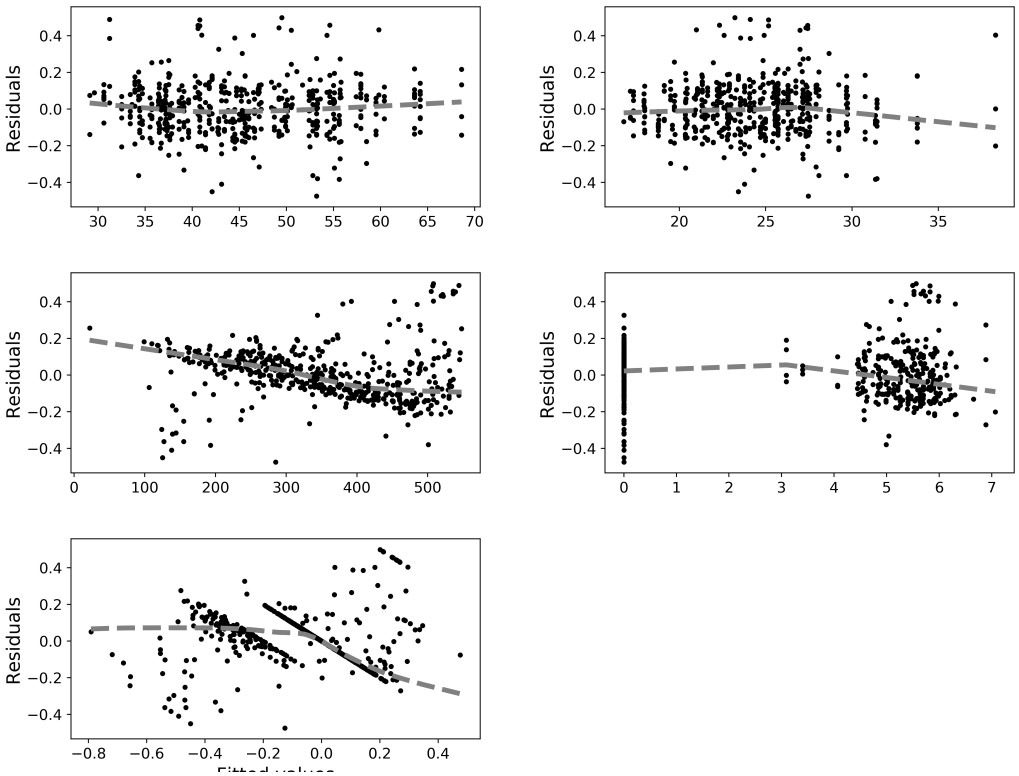

Figure 11: Residuals plotted versus Age, BMI, $CD4_{init}$, logResp and fitted values for outcome variable $U(0.7)$

Appendix H: Residual Plots for Outcome Variable $U(0.9)$

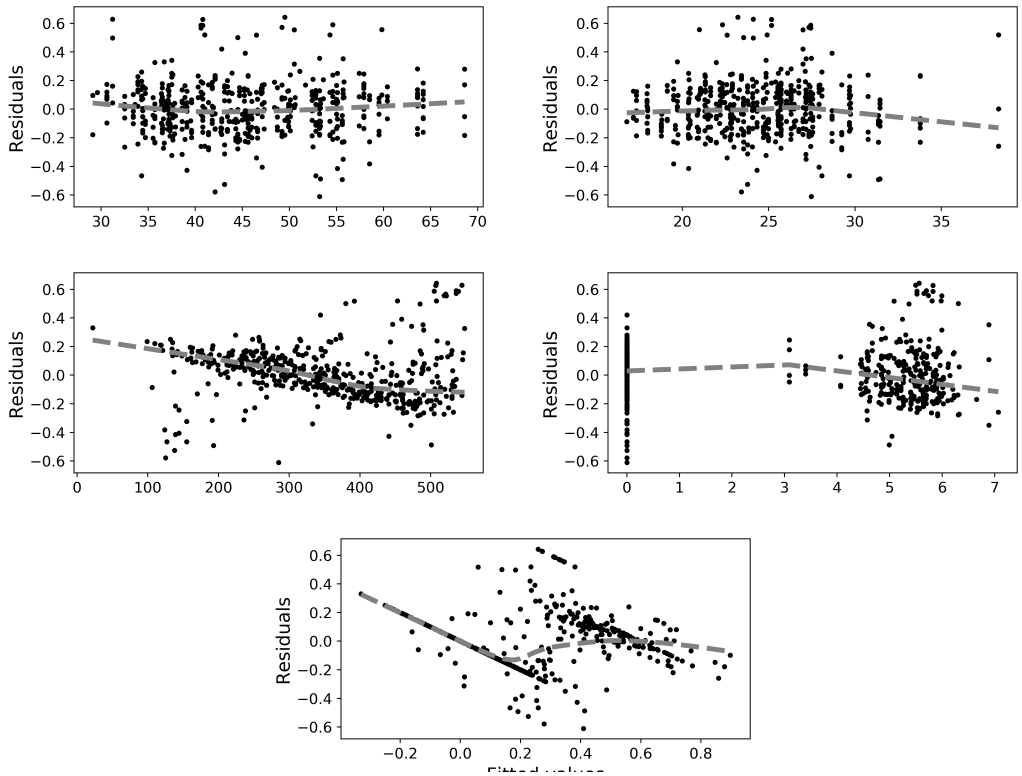

Figure 12: Residuals plotted versus Age, BMI, $CD4_1$, logResp and fitted values for outcome variable $U(0.9)$