Received: Added at production | Revised: Added at production | Accepted: Added at production

DOI: xxx/xxxx

ARTICLE TYPE

Longitudinal Self-Learning of Individualized Treatment Rules in A Nutrient Supplementation Trial with Missing Data

Yiwang Zhou*[1] | Peter X.K. Song[2]

[1]Department of Biostatistics, St. Jude Children's Research Hospital, TN, USA
[2]Department of Biostatistics, University of Michigan, MI, USA

Correspondence
*Corresponding author.
Email: yiwang.zhou@stjude.org

Present Address
262 Danny Thomas Pl, Memphis, TN, 38105

Summary

Longitudinal outcomes are prevalent in clinical studies, where the presence of missing data may make the statistical learning of individualized treatment rules (ITRs) a much more challenging task. We analyzed a longitudinal calcium supplementation trial in the ELEMENT Project and established a novel ITR to reduce the risk of adverse outcomes of lead exposure on child growth and development. Lead exposure, particularly in the form of *in utero* exposure, can seriously impair children's health, especially their cognitive and neurobehavioral development, which necessitates clinical interventions such as calcium supplementation intake during pregnancy. Using the longitudinal outcomes from a randomized clinical trial of calcium supplementation, we developed a new ITR for daily calcium intake during pregnancy to mitigate persistent lead exposure in children at age 3 years. To overcome the technical challenges posed by missing data, we illustrate a new learning approach, termed longitudinal self-learning (LS-learning), that utilizes longitudinal measurements of child's blood lead concentration in the derivation of ITR. Our LS-learning method relies on a temporally weighted self-learning paradigm to synergize serially correlated training data sources. The resulting ITR is the first of this kind in precision nutrition that will contribute to the reduction of expected blood lead concentration in children aged 0-3 years should this ITR be implemented to the entire study population of pregnant women.

KEYWORDS:
DOHaD hypothesis, machine learning, missing endpoint, precision health, support vector machine

1 | INTRODUCTION

This paper presents an applied statistical contribution to the emerging field of precision nutrition. Barker's hypothesis on the "Developmental Origins of Health and Disease (DOHaD)" postulates a key conceptual paradigm for the impact of the perinatal environment on the future health of offspring. That is, prenatal exposure to environmental toxicants may impair the developmental health of children during their infancy and childhood, and even later in their adulthood.[1] A vast literature has unveiled that excessive exposure to lead is detrimental on children's neurobehavioral and cognitive development.[2,3] Both blood and bone lead levels in children have been shown to be inversely associated with their intelligence.[4,5] All these adverse health outcomes make it necessary to control lead exposure to children. The DOHaD hypothesis emphasizes the critical role of lead exposure *in utero* on the health of children, which may be intervened during pregnancy. The two major sources contributing to maternal

blood lead levels include: i) the mother's immediate environmental exposure to lead, and ii) the mobilization of cumulative lead storage from mother's bones into blood circulation.[6,7] Therefore, reducing the amount of lead released into the maternal blood circulation during pregnancy is an important preventive measure to minimize lead exposure to children, since the fetus' prenatal lead exposure comes exclusively from mother. In clinically practice, blood lead control may be achieved with calcium, a lead blocking agent.[8,9,10]

In precision nutrition, a central task is to establish individualized treatment rules (ITRs) for subjects who respond differently to nutrient supplements for benefit. In this study to reduce maternal lead exposure to children, not all pregnant women will benefit from taking calcium supplementation. This may be due to two major reasons: first, some pregnant women already consume enough calcium from their diet, and second, excessive blood calcium may increase the risk of miscarriage.[11] Therefore, an ITR that can guide pregnant women to take calcium supplements is of great importance in precision nutrition. The derivation of such an ITR is usually aimed at maximizing the expectation of the desired outcome. In this paper, we focus on the longitudinal calcium supplementation trial from the Early Life Exposure in Mexico to ENvironmental Toxicants (ELEMENT) Project, an over 25-year cohort designed to study the effect of maternal lead exposure on child health outcomes. This longitudinal trial collected repeated measurements of blood lead concentration for children (PBC; "PB" represents lead and "C" represents children) at month 3, 6, 12, 18, 24, 30 and 36. In particular, the primary outcome for ITR derivation in this study is PBC36 because we are interested in persistent maternal lead exposure to children at 36 months of age. The observations of PBC36 is incomplete (36.7% missing) due to intermittent missing data or dropouts, which would result in a reduced sample size if only the endpoint outcome is used. As a result, the reliability and reproducibility of the resulting ITR would be compromised. Therefore, a natural solution is to utilize longitudinal data of blood concentrations prior to PBC36 to construct our ITR.

The core of ITR is based on certain decision functions that map individual characteristics to treatment choices. Statistical methods for estimating decision functions in the field of precision medicine are abundant, including Q-learning,[12] outcome weighted learning (OWL),[13] and residual weighted learning (RWL),[14] to name a few. Unfortunately, all these current learning methods are only applicable to one-dimensional cross-sectional outcome related to health benefits. The work of Huling[15] focused on the estimation of time-varying treatment effects for closer time points via the fused lasso regularization. The resulting ITR is to maximize the average expected improvement in the benefit outcome throughout the entire treatment period. However, this proposed method did not consider missing data in their work. Therefore, it lacks suitable methods in the literature to exploit temporally correlated outcomes and to deal with missing data in the estimation and optimization of ITRs. To analyze the longitudinal calcium supplementation trial, we propose and demonstrate a new machine learning approach, termed longitudinal self-learning (LS-learning), which not only allows for flexible incorporation of longitudinal outcomes, but also to handles incomplete data in ITR derivation. This new LS-learning approach provides a useful extension to support vector machine (SVM) that serves as the optimization engine in OWL. LS-learning works in a weighted self-learning paradigm through an effective training data augmentation scheme that builds on the assumption that the relevance of longitudinal outcomes increases over time toward the primary endpoint. In addition, estimated subgroup labels learned from longitudinal outcomes measured prior to the primary endpoint are iteratively calibrated with the labels determined by the primary endpoint. The relative contribution of longitudinal outcomes are tuned by minimizing the sum of squared errors (SSE) in predicting treatment benefits. To handle missing data, we follow the concept of pattern mixture model[16] by stratifying subjects into subgroups based on their missing data patterns. To deal with potential computational burden of tuning parameter selection, we employ a scalable tuning procedure that incorporates the nature of longitudinal data collection. We also discuss the problem of algorithmic convergence for LS-learning. Meanwhile, through extensive simulation experiments, we illustrate and confirm that LS-learning can numerically achieve high classification rates for treatment subgroups, high estimated value functions for expected outcomes, and high computational efficiency.

Applying our new LS-learning approach to derive an ITR in the calcium supplementation trial, we found that LS-learning would yield the greatest benefit in reducing children's exposure to blood lead concentration at 36 months if the resulting ITR were implemented for the entire study population of pregnant women compared to other ITRs derived from standard OWL. This improvement is clinically meaningful and was achieved in our analysis by means of longitudinal data augmentation, which demonstrates a useful approach to delivering more informative solutions in precision nutrition. Several biomarkers that play an important role in the formation of ITR were also identified. Interestingly, dietary calcium and fiber intake affects the assignment of calcium supplements, suggesting a complex set of interactions involving different nutrients.

The rest of the paper is organized as follows. Section 2 introduces the calcium supplementation trial. Section 3 presents the LS-learning method, including its algorithmic convergence and parameter tuning method. Data analysis is detailed in Section 4. Section 5 deals with the evaluation of LS-learning through simulation experiments. Section 6 contains some concluding remarks. The proof of the algorithmic convergence of LS-learning is given in Appendix.

FIGURE 1 Trajectory of PBC values in the calcium supplementation and placebo groups. Black summary lines are fitted by the generalized additive model (GAM).

FIGURE 2 Missing PBC values at different visit times. (a) Percentage of missing and observed PBC values at months 3, 6, 12, 18, 24, 30, and 36. (b) Stratification of mother-child pairs into different missing patterns according to individual endpoints.

2 | APPLICATION: CALCIUM SUPPLEMENTATION TRIAL

This paper plans to analyze an important calcium supplementation trial conducted in the third cohort of the ELEMENT study. This clinical trial included a total of 376 women who were recruited during their first trimester of pregnancy in Mexico City between 2001 and 2003, of whom 190 mothers were randomly assigned to receive a daily dose of 1200 mg calcium supplement and 186 were randomly assigned to receive placebo,[17] all of whom had complete baseline information collected. These mothers were followed both prenatally and postnatally. Demographic information and biological samples from both mothers and children (i.e., mother-child pairs) are collected for clinical measurements. The primary outcome of interest is PBC36, an endpoint of clinical importance because of its known effects on neurobehavioral development in children.[18,19] In addition to this primary outcome, six repeated PBC measurements were collected during the follow-up visits as part of the longitudinal study, and for convenience are denoted by PBC3, PBC6, ..., PBC30. Figure 1 shows the longitudinal trajectory of PBC values in the calcium supplementation and placebo groups. Rather than removing mother-child pairs missing measurements of the primary endpoint PBC36, we propose to borrow longitudinal PBC values measured prior to month 36 into ITR derivation, resulting in an improved ITR by avoiding data attrition. We propose to stratify mother-child pairs according to their missing data patterns, resulting in subgroups defined by their subject-specific last observation that is considered as a surrogate or an approximate lead exposure endpoint, measured at an earlier time point than month 36. For example, if PBC is measured in a child at month 3, 12, 18, and 30, then that individual's last PBC observation, PBC30, is used as a surrogate endpoint for ITR derivation. Figure 2 shows the stratification of subgroups with different individual endpoints.

Among many demographic and clinical variables, the following biomarkers were selected for ITR derivation through consultation with collaborators of the ELEMENT study, including baseline measurements of maternal age (year), maternal weight (kg), total years of maternal schooling, marital status (married: yes/no), total number of pregnancies, maternal dietary intake of calcium (mg/day), fiber (g/day), iron (mg/day), zinc (mg/day), and vitamin C (mg/day), and maternal hemoglobin (HgB) concentration (g/dL). Maternal blood lead concentration (μg/dL) measured in the first trimester (PBM; "M" represents mother) is also included as a baseline reference for maternal lead exposure level. Summary statistics of these biomarkers in the calcium supplementation and placebo groups are listed in Table 1. Not surprisingly, none of these biomarkers are marginally significant due to randomization.

3 | FORMULATION

This section presents the details of LS-learning, which will be used to derive an ITR using longitudinal data with missing outcomes.

3.1 | Notation

Denote by $\mathcal{S} = \{(\boldsymbol{X}_i, A_i, t_{ij}, \boldsymbol{B}_{ij}, R_{ij}), i = 1, \ldots, n, j = 0, \ldots, m_i\}$ the longitudinal data collected from a two-armed randomized clinical trial (e.g., the calcium supplementation trial). Let i and j denote the indexes of subject (i.e., mother-child pair i) and follow-up visit j, respectively. These two indices i or j can be suppressed in the absence of confusion. In the remainder of this paper, "subject" stands for "mother" when referring to baseline biomarker measurements and "child" when referring to outcome measurements (i.e., longitudinal PBC values). There is a total of n mother-child pairs and m_i repeated measurements for subject i in addition to the baseline visit prior to treatment randomization, which is denoted as $j = 0$. For ease of exposition, we consider a longitudinal trial with outcomes measured at common visit times $j = 0, \ldots, m$, where $j = m$ represents the time to measure

the primary endpoint (i.e., PBC36). Consequently, visit times $t_{ij} \equiv t_j, j = 0, \ldots, m$. The treatment randomly assigned at the beginning of the trial is denoted as $A_i \in \mathcal{A} = \{-1, 1\}$, where $A_i = 1$ represents the new treatment (i.e., calcium supplementation) and $A_i = -1$ represents placebo. $\boldsymbol{X}_i \in \mathbb{R}^d$ is the d-dimensional vector of biomarkers measured at baseline. The outcome of interest, denoted by B_{ij}, is repeatedly measured at time t_j for subject i. Note that the baseline measurement B_0 is always collected as a reference line and will not be considered as an outcome benefit in ITR derivation. It is assumed that the larger the value of B, the greater the health benefit. Specifically, the outcome B_m is used as the primary outcome for the derivation of ITR according to the clinical study. All the other intermittent measurements, $B_1, \ldots, B_{m-1}$, are considered as surrogate outcomes, providing relevant and auxiliary information on the benefit trajectory reaching the endpoint B_m. Some individuals in the trial are not measured at the last visit t_m due to dropout or other reasons, resulting in missing data for B_m. For those subjects who do not complete the trial, we do not remove them from the analysis but select certain surrogate endpoints so that they can remain in our ITR derivation. Let $R_{ij} \in \{0, 1\}$ denote the missingness of the outcome B_{ij}, with $R_{ij} = 1$ representing "observed" and $R_{ij} = 0$ representing "missing".

3.2 | Outcome Weighted Learning (OWL)

The central goal of precision nutrition is to derive a decision function $D : \mathcal{X} \to \mathcal{A}$, which is a mapping from the space of prognostic variables to the space of treatments. In the case of a one-dimensional outcome B, OWL[13] is a seminal work enabling the estimation of an optimal ITR D^* that maximizes the expected clinical benefit $E(D) = E\{\frac{I(A=D(\boldsymbol{X}))}{P(A|\boldsymbol{X})}B\}$, where $I(\cdot)$ is an indicator function of whether the randomly assigned treatment A is equal to the estimated treatment $D(\boldsymbol{X})$, and $P(A|\boldsymbol{X})$ is the propensity score for obtaining treatment A in the clinical trial based on feature vector $\boldsymbol{X}$. Although $E(D)$ can generally be computed, the following three assumptions are required in the literature to estimate causal effects: (a) consistency: the observed benefit $B = I(A = 1)B(1) + I(A = -1)B(-1)$, where $B(a)$ is the potential clinical benefit of receiving treatment a; (b) no unmeasured confounding: $A \perp \{B(a)\}_{a\in\mathcal{A}}|\boldsymbol{X}$, indicating that $\boldsymbol{X}$ contains all relevant confounding factors; (c) positivity: $0 < P(A = a|\boldsymbol{X}) < 1, \forall a \in \mathcal{A}$.[20] The maximization problem of OWL becomes a minimization problem if we change the indicator function from equality to inequality, resulting in $D^* \in \underset{D}{\operatorname{argmin}} E\{\frac{I(A\neq D(\boldsymbol{X}))}{P(A|\boldsymbol{X})}B\}$. The term $\frac{I(A\neq D(\boldsymbol{X}))}{P(A|\boldsymbol{X})}$ is actually a weighted classification error, making OWL a weighted classification problem. With a set of $i.i.d$ observations $\{(\boldsymbol{X}_i, A_i, B_i), i = 1, \ldots, n\}$, we can approximate the optimization problem by its empirical value $D^* \in \underset{D}{\operatorname{argmin}} \frac{1}{n}\sum_{i=1}^n \frac{B_i}{P(A_i|\boldsymbol{X}_i)} I(A_i \neq D(\boldsymbol{X}_i))$. Since $D(\boldsymbol{X})$ can always be expressed as $\operatorname{sign}(f(\boldsymbol{X}))$ for some decision function f, where $D(\boldsymbol{X}) = 1$ if $f(\boldsymbol{X}) > 0$ and $D(\boldsymbol{X}) = -1$ otherwise, the optimization problem is equivalent to $f^* \in \underset{f}{\operatorname{argmin}} \frac{1}{n}\sum_{i=1}^n \frac{B_i}{P(A_i|\boldsymbol{X}_i)} I(A_i \neq \operatorname{sign}(f(\boldsymbol{X}_i)))$. The challenge in solving this optimization problem is that it is a weighted sum of 0-1 loss, which is neither convex nor continuous. Therefore, OWL uses a convex surrogate hinge loss $x^+ = \max(0, x)$ to replace the 0-1 loss. To further penalize the complexity of the decision function f to avoid overfitting, OWL adds a l_2 penalty into the optimization problem. The final function that OWL aims to minimize is $\frac{1}{n}\sum_{i=1}^n \frac{B_i}{P(A_i|\boldsymbol{X}_i)}(1 - A_i f(\boldsymbol{X}_i))^+ + \kappa_n||f||_2$, where κ_n is the regularization parameter of the l_2 penalty and $\xi_i = (1 - A_i f(\boldsymbol{X}_i))^+$ is the slack variable. If we assume that the decision rule f is a linear function $f(\boldsymbol{X}) = \boldsymbol{\omega}^\top \boldsymbol{X} + b$, OWL can be formulated as a weighted classification problem in that SVM is invoked to solve the following constrained optimization:

$$\begin{aligned} &\min_{\boldsymbol{\omega},b,\xi} \frac{1}{2}||\boldsymbol{\omega}||^2 + C\sum_{i=1}^n \frac{B_i}{P(A_i|\boldsymbol{X}_i)}\xi_i \\ &\text{s.t. } A_i(\boldsymbol{\omega}^\top \boldsymbol{X}_i + b) \geq 1 - \xi_i \quad \text{and} \quad \xi_i \geq 0,\ i = 1, ..., n \end{aligned} \tag{1}$$

where $C = \frac{1}{2n\kappa_n}$ is a tuning parameter of SVM. The main difference between OWL and standard SVM is that OWL penalizes ξ_i via a propensity score scaled outcome, $\frac{B_i}{P(A_i|\boldsymbol{X}_i)}$. The problem (1) can be generalized by allowing a nonlinear decision rules through a kernel function $\mathcal{K} : \mathcal{X} \times \mathcal{X} \to R$ in the context of *reproducing kernel Hilbert space* (RKHS) $\mathcal{H}_\mathcal{K}$.[21] For simplicity, we focus on linear decision rules in this paper, while LS-learning can be easily extended to nonlinear decision rules.

3.3 | Longitudinal Self-learning (LS-learning)

To overcome the limitation that OWL is only applicable to a single cross-sectional outcome, we propose an extension to address the methodological needs in ITR deviation with longitudinal outcomes. This requires new algorithms to solve a multi-view problem similar to but more difficult than the problem (1) with multiple outcomes. We perform this extension in the context of multiple training datasets obtained by stratifying the missing data patterns in a clinical trial with dropouts. For patients who

miss certain measurements in the follow-up visits of a clinical trial, we use indicator $R_{ij} \in \{0,1\}$ to denote the missingness of the corresponding outcome B_{ij}, where $R_{ij} = 1$ represents "observed" and $R_{ij} = 0$ represents "missing". Define $\mathcal{S}_m = \{(\boldsymbol{X}_i, A_i, t_m, B_{im}), i | I(R_{im} = 1) = 1\}$ as the subset of completers who are measured at the endpoint time for their primary benefit outcome B_m, and $\mathcal{S}_j = \{(\boldsymbol{X}_i, A_i, t_j, B_{ij}), i | I(R_{ij} = 1) \prod_{l=j+1}^{m} I(R_{il} = 0) = 1\}$ as the subset of incompleters who are missing at the endpoint time with their subject-specific last observations $B_{ij}, j \in \mathcal{J}$. In this paper, we primarily focus the benefit outcome B_m measured at the primary endpoint, and it seems clinically sensible to use the surrogate benefit outcome $B_{ij}, j \in \mathcal{J}$ measured at a timepoint closest to the designed endpoint under the assumption that the quality or relevance of benefit outcome increases when it approaches to the endpoint time. To estimate ITR, we plan to apply LS-learning to integrate the subset of completers $\mathcal{S}_m$ with all the other subsets $\mathcal{S}_j, j \in \mathcal{J}$ in a systematic way.

In Section 3.2, we have illustrated the derivation of the objective function (1). When longitudinal outcomes are available, the expected clinical benefit that needs to be maximized for the completers is $E\{\frac{I(A=D(\boldsymbol{X}))}{P(A,R_m=1|\boldsymbol{X})} B_m I(R_m = 1)\}$, resulting in the objective function $\min_{\boldsymbol{\omega},b,\boldsymbol{\xi}} \frac{1}{2}||\boldsymbol{\omega}||^2 + C \sum_{i\in\mathcal{S}_m} \frac{B_{im}}{p_{im}} \xi_i$, s.t. $A_i(\boldsymbol{\omega}^\top \boldsymbol{X}_i + b) \geq 1 - \xi_i$ and $\xi_i \geq 0$, $i \in \mathcal{S}_m$, with $p_{im} = P(A_i, R_{im} = 1 | \boldsymbol{X}_i)$ being the probability of subject i assigned to treatment A_i and contributing benefit B_{im} for ITR derivation. Integrating the other subsets $\mathcal{S}_j, j \in \mathcal{J}$ with auxiliary information B_j into ITR derivation requires the maximization of the expected clinical benefit $E\{\frac{I(A=D(\boldsymbol{X}))}{P(A,R_j=1,R_{j+1}=0,\ldots,R_m=0|\boldsymbol{X})} B_j I(R_j = 1) \prod_{l=j+1}^{m} I(R_l = 0)\}$, leading to the objective function $\min_{\boldsymbol{\omega},b,\boldsymbol{\xi}} \frac{1}{2}||\boldsymbol{\omega}||^2 + C \sum_{i\in\mathcal{S}_j} \frac{B_{ij}}{p_{ij}} \tilde{\xi}_{ij}$, s.t. $A_i(\boldsymbol{\omega}^\top \boldsymbol{X}_i + b) \geq 1 - \tilde{\xi}_{ij}$ and $\tilde{\xi}_{ij} \geq 0$, $i \in \mathcal{S}_j$, with $p_{ij} = P(A_i, R_{ij} = 1, R_{ij+1} = 0, \ldots, R_{im} = 0 | \boldsymbol{X}_i)$ being the probability of subject i assigned to treatment A_i and contributing benefit B_{ij} for ITR derivation. The ITR derivation based on a simple sum of these objective functions is not feasible for two main reasons. First, the quality of information contributed for ITR derivation varies for each benefit outcome. Second, if some outcome B_j is completely untrustworthy, excluding it completely from the derivation of ITR would result in information loss, since the feature vector $\boldsymbol{X}_i, i \in \mathcal{S}_j$ also contains information useful for ITR derivation. To address these challenges, we propose an iterative LS-learning algorithm in a generic manner. Prior to the detailed illustration of the algorithm, we first provide the condition required for LS-learning.

3.3.1 | Missing Data Mechanisms

To establish LS-learning, we postulate Condition 1: (i) the missing data mechanism follows missing at random (MAR), where $(A, R_j, R_{j+1}, \ldots, R_m) \perp (B_j, B_{j+1}, \ldots, B_m) | \boldsymbol{X}$; and (ii) $A \perp R_j, R_{j+1}, \ldots, R_m | \boldsymbol{X}$. For a specific subgroup $\mathcal{S}_j, j = 1, \ldots, m-1$, the calculation of the benefit expectation that we want to maximize can be formulated as

$$
\begin{aligned}
&E\{\frac{I(A = D(\boldsymbol{X}))}{P(A, R_j = 1, R_{j+1} = 0, \ldots, R_m = 0|\boldsymbol{X})} B_j I(R_j = 1) \prod_{l=j+1}^{m} I(R_l = 0)\} \\
&= \int \int \cdots \int \frac{I(A = D(\boldsymbol{X}))}{P(A, R_j = 1, R_{j+1} = 0, \ldots, R_m = 0|\boldsymbol{X})} B_j I(R_j = 1) \prod_{l=j+1}^{m} I(R_l = 0) \times \\
&\quad f(B_j, B_{j+1}, \ldots, B_m, R_j, R_{j+1}, \ldots, R_m, A | \boldsymbol{X}, \boldsymbol{\theta}, \boldsymbol{\eta}, \boldsymbol{\phi}) d\mu(B_j) d\mu(B_{j+1}) \cdots d\mu(B_m) d\mu(R_j) d\mu(R_{j+1}) \cdots d\mu(R_m) d\mu(A),
\end{aligned}
$$

where $\mu(\cdot)$ is a generic notation for a certain dominated measure (either Lebesgue or counting measure). The parameters $\boldsymbol{\theta}, \boldsymbol{\eta}, \boldsymbol{\phi}$ are those involved in the missing data process, the propensity score modeling, and the outcome data process, respectively. Benefits $B_1, \ldots, B_{j-1}$ are not included in the density function due to the marginalization of integration. Since the treatment A, the missing indicators $R_j, R_{j+1}, \ldots, R_m$, and the benefit B_j are observed and the decision function $D(\boldsymbol{X})$ is a function of the feature $\boldsymbol{X}$, the multiple integration does not involve the missing benefits except for the density $f(B_j, B_{j+1}, \ldots, B_m, R_j, R_{j+1}, \ldots, R_m, A | \boldsymbol{X}, \boldsymbol{\theta}, \boldsymbol{\eta}, \boldsymbol{\phi})$. Based on the conditional probability, we have the density equal to $f(A, R_j, R_{j+1}, \ldots, R_m | B_j, B_{j+1}, \ldots, B_m, \boldsymbol{X}, \boldsymbol{\theta}, \boldsymbol{\eta}) f(B_j, B_{j+1}, \ldots, B_m | \boldsymbol{X}, \boldsymbol{\phi})$. It follows from Condition 1 that

$$
f(A, R_j, R_{j+1}, \ldots, R_m | B_j, B_{j+1}, \ldots, B_m, \boldsymbol{X}, \boldsymbol{\theta}, \boldsymbol{\eta}) = f(A, R_j, R_{j+1}, \ldots, R_m | \boldsymbol{X}, \boldsymbol{\theta}, \boldsymbol{\eta}) = f(A | \boldsymbol{X}, \boldsymbol{\eta}) f(R_j, R_{j+1}, \ldots, R_m | \boldsymbol{X}, \boldsymbol{\theta}).
$$

Therefore, the probability $P(A, R_j = 1, R_{j+1} = 0, \ldots, R_m = 0|\boldsymbol{X}) = P(A|\boldsymbol{X}) P(R_j = 1, R_{j+1} = 0, \ldots, R_m = 0|\boldsymbol{X})$. This leads to an inverse probability weighting scheme in a manner similar to the propensity score. These probabilities can be estimated by means of logistic regression. With these derivations, the only part of the final calculation of the multiple integration that involves the missing benefits is $f(B_j, B_{j+1}, \ldots, B_m | \boldsymbol{X}, \boldsymbol{\phi})$, which is equal to $f(B_{j+1}, \ldots, B_m | B_j, \boldsymbol{X}, \boldsymbol{\phi}) f(B_j | \boldsymbol{X}, \boldsymbol{\phi})$. Since only B_j is involved in the formula for the benefit expectation, the multiple integration $\int \int \cdots \int B_j f(B_{j+1}, \ldots, B_m | B_j, \boldsymbol{X}, \boldsymbol{\phi}) f(B_j | \boldsymbol{X}, \boldsymbol{\phi}) d\mu(B_j) d\mu(B_{j+1}) \cdots d\mu(B_m) = \int B_j f(B_j | \boldsymbol{X}, \boldsymbol{\phi}) d\mu(B_j)$. This implies that

all missing benefit outcomes are integrated out and have no impact on the calculation of the expected benefit. Therefore, deriving ITR by maximizing the expected benefit based on the observed training data produces valid solutions in the sense that no additional model is needed for the missing benefit outcomes.

3.3.2 | Algorithm of LS-learning

Algorithm 1. Suppose we have a set of completers $\mathcal{S}_m = \{(\boldsymbol{X}_i, A_i, t_m, B_{im}), i | I(R_{im} = 1) = 1\}$ with outcome B_m and multiple sets of incompleters $\mathcal{S}_j = \{(\boldsymbol{X}_i, A_i, t_j, B_{ij}), i | I(R_{ij} = 1) \prod_{l=j+1}^{m} I(R_{il} = 0) = 1\}$ with outcome $B_j, j \in \mathcal{J}$. It is assumed that B_m is of higher quality than B_j in ITR derivation. Let $p_{im} = P(A_i, R_{im} = 1 | \boldsymbol{X}_i)$ and $p_{ij} = P(A_i, R_{ij} = 1, R_{ij+1} = 0, \ldots, R_{im} = 0 | \boldsymbol{X}_i)$ be the probabilities of being assigned to treatment A and having B_m or B_j observed for ITR derivation, $i \in \mathcal{S}_m$ and $i \in \mathcal{S}_j$, respectively.

S1 Estimate ITR by OWL using $\mathcal{S}_m$ and get the initial estimates $(\hat{\boldsymbol{\omega}}^{(0)}, \hat{b}^{(0)}, \hat{\boldsymbol{\xi}}^{(0)})$. Then predict labels for subjects in $\bigcup_{j=1}^{m-1} \mathcal{S}_j$. Denote the predicted labels as $\hat{y}_{ij}^{(0)}, i \in \mathcal{S}_j, j \in \mathcal{J}$.

S2 The k-th ($k \geq 1$) iteration runs through the following steps S2.1-S2.3.

S2.1 Define an augmented training dataset $\mathcal{S}^{(k)} = \mathcal{S}_m \cup \mathcal{S}_1^{(k)} \cup \cdots \cup \mathcal{S}_{m-1}^{(k)}$, where individuals in $\mathcal{S}_j^{(k)}$ have their predicted labels from the previous iteration $k-1$, resulting in $\mathcal{S}_j^{(k)} = \{(\boldsymbol{X}_i, A_i, t_j, B_{ij}, \hat{y}_{ij}^{(k-1)}), i = 1, \ldots, n_j\}$. Note that the subjects in $\mathcal{S}_j$ and $\mathcal{S}_j^{(k)}$ are the same. $\mathcal{S}_j^{(k)}$ is created to emphasize the additionally included information $\hat{y}_{ij}^{(k-1)}$.

S2.2 Solve the following optimization problem using $\mathcal{S}^{(k)}$:

$$
\begin{aligned}
\min_{\boldsymbol{\omega}, b, \boldsymbol{\xi}} \quad & \frac{1}{2}||\boldsymbol{\omega}||^2 + C\left[\sum_{i \in \mathcal{S}_m} \frac{B_{im}}{p_{im}} \xi_i + \sum_{j=1}^{m-1} \left\{ \lambda_j \sum_{i \in \mathcal{S}_j^{(k)}} \frac{B_{ij}}{p_{ij}} \tilde{\xi}_{ij} + (1-\lambda_j) \sum_{i \in \mathcal{S}_j^{(k)}} \tau_{B_j} \check{\xi}_{ij} \right\}\right] \\
\text{s.t.} \quad & A_i(\boldsymbol{\omega}^\top \boldsymbol{X}_i + b) \geq 1 - \xi_i \quad \text{and} \quad \xi_i \geq 0, \; i \in \mathcal{S}_m, \\
& A_i(\boldsymbol{\omega}^\top \boldsymbol{X}_i + b) \geq 1 - \tilde{\xi}_{ij} \quad \text{and} \quad \tilde{\xi}_{ij} \geq 0, \; i \in \mathcal{S}_j^{(k)}, \; j \in \mathcal{J}, \\
& \hat{y}_{ij}^{(k-1)}(\boldsymbol{\omega}^\top \boldsymbol{X}_i + b) \geq 1 - \check{\xi}_{ij} \quad \text{and} \quad \check{\xi}_{ij} \geq 0, \; i \in \mathcal{S}_j^{(k)}, \; j \in \mathcal{J}.
\end{aligned} \tag{2}
$$

Denote the estimates of the parameters as $(\hat{\boldsymbol{\omega}}^{(k)}, \hat{b}^{(k)}, \hat{\boldsymbol{\xi}}^{(k)})$ at iteration k. Also, the predicted labels for $\mathcal{S}_j^{(k)}$ are updated as $\hat{y}_{ij}^{(k)}, i \in \mathcal{S}_j^{(k)}, j \in \mathcal{J}$.

S2.3 Calculate the objective function with the estimates $(\hat{\boldsymbol{\omega}}^{(k)}, \hat{b}^{(k)}, \hat{\boldsymbol{\xi}}^{(k)})$ as

$$
h(\hat{\boldsymbol{\omega}}^{(k)}, \hat{\boldsymbol{\xi}}^{(k)}) = \frac{1}{2}||\hat{\boldsymbol{\omega}}^{(k)}||^2 + C\left[\sum_{i \in \mathcal{S}_m} \frac{B_{im}}{p_{im}} \hat{\xi}_i^{(k)} + \sum_{j=1}^{m-1} \left\{ \lambda_j \sum_{i \in \mathcal{S}_j^{(k)}} \frac{B_{ij}}{p_{ij}} \hat{\tilde{\xi}}_{ij}^{(k)} + (1-\lambda_j) \sum_{i \in \mathcal{S}_j^{(k)}} \tau_{B_j} \hat{\check{\xi}}_{ij}^{(k)} \right\}\right]
$$

S3 The algorithm stops if $\frac{|h(\hat{\boldsymbol{\omega}}^{(k)}, \hat{\boldsymbol{\xi}}^{(k)}) - h(\hat{\boldsymbol{\omega}}^{(k-1)}, \hat{\boldsymbol{\xi}}^{(k-1)})|}{h(\hat{\boldsymbol{\omega}}^{(k-1)}, \hat{\boldsymbol{\xi}}^{(k-1)})} < \epsilon$ for a pre-determined precision constant ϵ, say 10^{-4}. The convergence values of $(\hat{\boldsymbol{\omega}}^{(k)}, \hat{b}^{(k)})$ are denoted by $(\hat{\boldsymbol{\omega}}, \hat{b})$, and the predicted labels at convergence for the whole training data $\mathcal{S}$ are denoted as $\hat{y}_i, i = 1, \ldots, n$.

The predicted label $\hat{y}_{ij}^{(k)}$ is the prediction of the underlying optimal treatment assignment $y_{ij}^{(k)}$ (i.e., $y_{ij}^{(k)} = \text{sign}\{f^{(k)}(\boldsymbol{X}_i)\}$) for each subject i in the subgroup $\mathcal{S}_j$ at iteration k, where $f^{(k)}(\boldsymbol{X}) = \boldsymbol{\omega}^{(k)\top} \boldsymbol{X} + b^{(k)}$ is the ITR derived by weighted SVM at iteration k. The tuning parameter $\lambda_j \in [0, 1]$ is included in the objective function (2) to characterize the relative importance of the auxiliary benefit B_j compared to the primary outcome B_m. When $\lambda_j = 1$, we believe that B_j has the same quality as B_m. When $0 < \lambda_j < 1$, B_j is assumed to have lower quality than B_m, but still provides useful information for ITR derivation. On the contrary, when $\lambda_j = 0$, we consider B_j to be completely untrustworthy and should be excluded from ITR derivation. Selection of the tuning parameter λ_j is discussed in the following Section 3.5. Note that when B_j is completely untrustworthy, its contribution to optimization is reduced to zero under $\lambda_j = 0$. However, the information contained in the feature vector $\boldsymbol{X}$ is good and useful. To use $\boldsymbol{X}$ of subjects in $\mathcal{S}_j$, we introduced the predicted label $\hat{y}_{ij}^{(k-1)}$ to engage $\boldsymbol{X}$ in the optimal decision making. It is known that the predicted label $\hat{y}_{ij}^{(k-1)}$ provides some information about the classification hyperplane, which may be

further sharpened by the valuable information from the feature vector $\boldsymbol{X}$. This sharpening of the classifier is carried out through the added loss function $\sum_{i\in S_j^{(k)}} \tau_{B_j} \check{\xi}_{ij}$ in SVM with a constraint $\hat{y}_{ij}^{(k-1)}(\boldsymbol{\omega}^\top \boldsymbol{X}_i + b) \le 1 - \check{\xi}_{ij}$. In this way, we do not waste any good data $\boldsymbol{X}$ for ITR derivation, even if the associated benefit B_j has no value. Moreover, adding the predicted labels may help the algorithm converge as the tuning can reduce the disagreement between the primary outcomes and the surrogate outcomes with respect to the ITR. The parameter τ_{B_j} is a constant summary statistic (e.g., sample mean) of $\frac{B_{ij}}{p_{ij}}$. Since this is a minimization problem, unbalanced magnitude between $\sum_{i\in S_j^{(k)}} \frac{B_{ij}}{p_{ij}} \widetilde{\xi}_{ij}$ and $\sum_{i\in S_j^{(k)}} \check{\xi}_{ij}$ due to the lack of weights in the latter will result in a certain improper λ_j value (i.e., smaller values of λ_j may be incorrectly favored due to the large weight $\frac{B_{ij}}{p_{ij}}$ for $\widetilde{\xi}_{ij}$), leading to ill-weighted contributions between the second and third terms to the objective function. Therefore, we add a constant τ_{B_j} to balance the magnitude between $\sum_{i\in S_j^{(k)}} \frac{B_{ij}}{p_i} \widetilde{\xi}_{ij}$ and $\sum_{i\in S_j^{(k)}} \check{\xi}_{ij}$ to derive apt tuning of λ_j. We suggest choosing this quantity in advance for the purpose of properly scaling the third term. Since τ_{B_j} is a constant without any information on individual benefits, and $\hat{y}_{ij}^{(k-1)}$ is the estimated underlying optimal label not the randomized treatment from the clinical trial, the third term added in (2) actually represents a standard SVM problem of maximizing the classification accuracy with respect to the optimal treatment assignment of subjects in S_j rather than an OWL problem maximizing the clinical benefit.

3.4 | Algorithmic Convergence

We establish the algorithmic convergence of LS-learning in the following theorem with the proof given in Appendix.

Theorem 1. The objective function $h(\hat{\boldsymbol{\omega}}^{(k)}, \hat{\boldsymbol{\xi}}^{(k)})$ of LS-learning at iteration k satisfies the descending property over iterations, namely $h(\hat{\boldsymbol{\omega}}^{(k)}, \hat{\boldsymbol{\xi}}^{(k)}) \le h(\hat{\boldsymbol{\omega}}^{(k-1)}, \hat{\boldsymbol{\xi}}^{(k-1)})$ for $k \ge 2$. The equality occurs when the algorithm converges.

Note that the problem of maximizing the expected benefit $E\{\frac{I(A=D(\boldsymbol{X}))}{P(A|\boldsymbol{X})} B\}$ is converted to an equivalent problem of minimizing $E\{\frac{I(A\neq D(\boldsymbol{X}))}{P(A|\boldsymbol{X})} B\}$, and the latter being suitable for formulating the SVM algorithm. In other words, the actual optimization implemented in OWL is to minimize objective (1). This same trick is applied to the optimization in LS-learning. Consequently, our algorithmic convergence is established for a consistent decrease of the LS-learning objective function $h(\cdot)$ over iterations.

3.5 | Tuning Parameter Selection

In order to perform LS-learning, we need to determine the tuning parameters C and $\boldsymbol{\lambda} = \{\lambda_j, j = 1, \ldots, m-1\}$. Parameter selection can be done by performing a grid search on the entire dataset $\mathcal{S} = \bigcup_{j=1}^{m} S_j$ for C and $\boldsymbol{\lambda}$. We propose to choose values of C and $\boldsymbol{\lambda}$ at which the sum of squared errors (SSE) of the predicted benefit values is minimized. The reason we tune the parameters according to the minimization of SSE is discussed in Section 6. SSE is defined as $\text{SSE} = \sum_{j=1}^{m} \left\{ \sum_{i=1}^{n_j} (B_{ij} - \hat{B}_{ij})^2 \right\}$, where the predicted benefit value $\hat{B}_{ij}$ is given by a stratum-specific regression model $B_j = \phi_j(\boldsymbol{X}, \boldsymbol{s}(t_j), A\hat{f}(\boldsymbol{X})) + \epsilon_j$, with $\phi_j(a, b, c) = g_j(a) + h_j(b) + \beta_j c$ for some function $g_j(\cdot)$ and $h_j(\cdot)$[22]. $\hat{f}(\cdot)$ is the decision function estimated by LS-learning, and ϵ_j is the error term. Note that inclusion of the third term $A\hat{f}(\boldsymbol{X})$ in the regression model $B_j = \phi(\boldsymbol{X}, \boldsymbol{s}(t_j), A\hat{f}(\boldsymbol{X})) + \epsilon_j$ is motivated by treating $\hat{f}$ as an estimate of the conditional treatment effect (CTE).[23,24,25] Multiple linear regression models or more flexible models, such as the generalized additive model (GAM), can be invoked to build the prediction rule $\phi_j(\cdot)$. In the analysis of the calcium trial discussed in Section 4, and in the simulation experiments described in Section 5, we choose the GAM for ϕ_j to predict the outcomes.

Each auxiliary dataset S_j introduces a λ_j for tuning, and the strategy of greedy search may make parameter tuning computationally expensive and time consuming. To alleviate this computational burden, we propose a scaled tuning method by specifying λ_j as a function of the standardized time $\tilde{t}_j = t_j/t_m$, i.e., $\lambda_j = \psi(\tilde{t}_j|\boldsymbol{\theta})$ with parameter $\boldsymbol{\theta}$. The rationale for this systematic tuning is that the quality of B_j decreases when the time of the current measurement is far from the measuring time t_m of the primary endpoint. As a result, this tuning procedure requires fewer parameters $\boldsymbol{\theta}$ to be tuned compared to the grid search. Various $\psi(\cdot)$ functions can be specified based on some preliminary understanding of the underlying longitudinal relationship between time and the relative quality of outcomes. Some examples of $\psi(\cdot)$ are given as follows. Their time-course relevance to the endpoint are shown in Figure 3:

1. (Linear) $\psi(\tilde{t}_j|\theta) = (1-\theta) + \theta\tilde{t}_j$, with constraint $\theta \in [0, 1]$.
2. (Exponential) $\psi(\tilde{t}_j|\boldsymbol{\theta}) = \theta_1 e^{\theta_2 \tilde{t}_j}$, with constraint $\theta_1 \in (0, 1]$ and $\theta_1 e^{\theta_2} = 1$.

FIGURE 3 Change of the tuning parameter λ_j according to standardized time $\tilde{t}_j$.

3. (Polynomial) $\psi(\tilde{t}_j|\boldsymbol{\theta}) = \theta_1\tilde{t}_j^2 + \theta_2\tilde{t}_j + \theta_3$, with constraint $\theta_3 \in [0,1]$ and $\theta_1 + \theta_2 + \theta_3 = 1$.

Parameter tuning is a crucial part of machine learning methods, including our LS-learning approach, in which the optimal treatment assignment in the training and validation datasets are unknown. In the seminal paper of OWL,[13] cross-validation was used for parameter tuning in SVM, which is not suitable for our case where the data of the primary benefit outcome had a small sample size, resulting in low sensitivity of cross-validation. This is because the low resolution of the estimated value function makes parameter tuning a challenging task, especially in the case of sparse data. Therefore, our experience suggested that cross-validation with limited data in LS-learning did not provide reliable tuning parameter selection. To overcome this challenge, we proposed to minimize the sum of squared errors (SSE) of the predicted benefits in a spirit similar to Q-learning.[12] Since we had little knowledge of the true benefit outcome model, in principle, the prediction model should be specified with a great deal of flexibility. We modeled the longitudinal benefit outcomes with $\boldsymbol{B}_j = \phi_j(\boldsymbol{X}, s(t_j), A\hat{f}(\boldsymbol{X})) + \epsilon_j$, and following the work of Diggle and Zeger,[22] we took $\phi_j(a,b,c) = g_j(a) + h_j(b) + \beta_j c$. The interaction term $A\hat{f}(\boldsymbol{X})$ is used to model the feature-driven treatment effects on outcome $\boldsymbol{B}_j$, adjusting for the major effects through functions $g_j(\cdot)$ and $h_j(\cdot)$ of feature vector $\boldsymbol{X}$ and the time trend t_j, respectively. In the Supporting Information section, we also showed the performance of SSE minimization-based parameter tuning compared to other tuning methods, including F-statistic, robust F-statistic, and truncated F-statistic, which are popular tuning criteria in classification analyses.[26] Their performances were found to be inferior to that of the SSE-based tuning procedure. In addition, we investigated the value information criterion (VIC) and the concordance information criterion (CIC)[27] for parameter tuning and model selection. In short, VIC and CIC are defined as $\text{VIC}(\boldsymbol{\beta}) = n\hat{V}\boldsymbol{\beta} - \kappa_n||\boldsymbol{\beta}||_0$ and $\text{CIC}(\boldsymbol{\beta}) = n\hat{C}(\boldsymbol{\beta}) - \kappa_n(\boldsymbol{\beta})$, where $\hat{V}(\boldsymbol{\beta})$ and $\hat{C}(\boldsymbol{\beta})$ are the estimated value and concordance function with parameter $\boldsymbol{\beta}$. Parameter κ_n is a tuning parameter and $||\boldsymbol{\beta}||_0$ is the number of nonzero elements of $\boldsymbol{\beta}$. If robust learning is used for ITR estimation, cross-validation can be applied to determine the value of the tuning parameter κ_n to select the best model that maximizes VIC and CIC. This additional parameter tuning step is needed when dimension reduction from high-dimensional features is used in ITR derivation, which is not the setting considered in our calcium supplementation trial. In our low-dimensional case, this additional cross-validation tuning step increases the complexity of the algorithm and is unstable due to small training dataset. Therefore, we chose not to use these VIC and CIC information criteria to tune the C and λ parameters in the proposed LS-learning algorithm. Nevertheless, simulation results also showed that the VIC-based tuning produced inferior results compared to the SSE-based tuning procedure (results are included in the Supporting Information section).

4 | DERIVATION OF ITR FOR THE CALCIUM SUPPLEMENTATION TRIAL

We applied the LS-learning method with the scaled-tuning scheme to analyze the calcium supplementation trial in the ELEMENT Project. The central objective of our analysis was to derive an ITR that could guide pregnant women to take calcium supplements to minimize their children's persistent lead exposure at age of 3 years. The PBC values were measured repeatedly at a total of seven different follow-up times, denoted as $\boldsymbol{B}_1, \ldots, \boldsymbol{B}_7$, where $\boldsymbol{B}_7$ was designed as the primary endpoint of interest. It is reasonable to assume that the quality of these longitudinal outcomes increases as the PBC was measured closer to month 36. LS-learning is proposed based on ranked benefit values, where the larger the magnitude of the outcome, the greater the clinically desirable benefit. Therefore, to fit this framework, we reversed the direction of lead concentration by a transformation, $\max_i(\boldsymbol{B}_{ij}) - \boldsymbol{B}_{ij}, i \in \mathcal{S}_j, j = 1, \ldots, 7$. A total of 376 mother-child pairs with no missing data at the baseline visit were used for our ITR derivation, with 190 mothers taking calcium supplements and 186 mothers taking placebo.

The mother-child pairs on the two treatment groups are not perfectly balanced due to subgroup stratification, so propensity scores were estimated by a logistic regression model with all the major effects of the biomarkers. After weighting the PBC values with the estimated propensity scores, we used standardized mean difference (SMD) to access the balance of the propensity scores and, more importantly the balance of the distribution of the individual biomarkers. Figure 4 shows that the distribution balance of biomarkers is satisfactorily achieved.

We now derive the ITR for the daily intake of calcium supplementation. We perform (i) a standard OWL on the completer group $\mathcal{S}_7$ with $\boldsymbol{B}_7$; (ii) a standard OWL on the entire training data $\mathcal{S}$ with missing information imputed by multiple imputation using predictive mean matching (PMM) via the R package MICE[28]. A total of 50 imputed datasets were created, from which

FIGURE 4 (a) Distributional balance of the propensity scores before and after weighting. (b) Balance of individual biomarkers before and after weighting adjusted by SMD at a threshold of 0.1, which are indicated by the vertical dashed lines.

the average statistics of OWL performances were calculated; (iii) a standard OWL on the entire training data S with missing information replaced by the subject-specific last observed PBC under the last observation carried forward (LOCF) strategy.[29] In effect, LOCF implies that the subject-specific last observed PBC is used as endpoint outcome for ITR derivation. Apparently, this approach ignores the different relevance of longitudinal outcomes in the reference to the primary endpoint $\boldsymbol{B}_7$. From the perspective of tuning parameter, this is equivalent to naively setting $\lambda_j = 1, j = 1, \ldots, 6$; (iv) LS-learning with the scaled tuning scheme on the entire training data S, which takes into account differences among longitudinal outcomes $\boldsymbol{B}_j, j = 1, \ldots, 6$ in terms of their relevance to $\boldsymbol{B}_7$. In LS-learning, we use the linear kernel in the optimization function (2) to train the ITR, tuned with the utility of the exponential scaling function (the middle panel of Figure 3). We set $C \in \{1/128, 1/64, \ldots, 1/2, 1\}$ and $\theta_1 \in \{0.1, 0.2, \ldots, 0.9, 1.0\}$ as the tuning parameter choices. The 3-fold cross-validation method was used to select the optimal tuning parameters in the derivation of ITRs. The loading coefficients for the biomarkers in the four estimated ITRs are listed in Table 2.

According to the exponential scaling procedure, we obtain a tuning parameter $\theta_1 = 0.4$, which, when ITR derivation is performed by LS-learning, generates $\lambda_1 = 0.432, \lambda_2 = 0.466, \lambda_3 = 0.543, \lambda_4 = 0.632, \lambda_5 = 0.737$ and $\lambda_6 = 0.858$. The tuning parameter θ_2 is not tuned because it satisfies a constrain $\theta_1 e^{\theta_2} = 1$ due to the reason that at the standardized time $\tilde{t}_j = 1$, the measured benefit is actually the primary benefit $\boldsymbol{B}_m$, leading to $\theta_1 e^{\theta_2 \tilde{t}_j} = \theta_1 e^{\theta_2} = 1$. As a result, when θ_1 is selected, θ_2 is given by $\theta_2 = \ln(1/\theta_1)$. In this LS-learning process, we use two tuning parameters C and θ_1, each with 8 and 10 different values, resulting in a total of 80 different pairs of C and θ_1. In contrast, if we use the greedy tuning strategy under $\lambda_j \in \{0, 0.1, \ldots, 1.0\}, j = 1, \ldots, 6$, we would have to deal with a total of $8 \times 11^6 = 14,172,488$ pairs in the tuning. Thus, our scaling procedure has saved 177,156 folds of computational runs.

Applying each of the four estimated decision rules above, we allocate a pregnant woman to take calcium supplement if $\hat{f} > 0$; otherwise, not to take the calcium supplement. Among the 376 mothers, the ITRs derived from OWL on S_m, OWL using PMM imputation, OWL using LOCF imputation and LS-learning would designate 239 (63.6%), 236 (62.8%), 235 (62.5%) and 251 (66.8%) pregnant women taking calcium supplements, respectively, all higher than the 50% randomly assigned. In particular, comparing between ITRs derived from complete randomization and by LS-learning, we see significant differences in the reallocation of calcium supplementation (See Table 3). Our McNemar test for the hypothesis of homogeneous allocation yields a p-value of 1.39×10^{-5}, indicating that there is a significant difference between the two allocation rules; in other words, these different treatment allocations do not occur by chance.

Furthermore, to compare the performance of the four estimated ITRs based on the different algorithms for minimizing continuous lead exposure at different time points, we calculate the estimated value functions $\hat{V}(\boldsymbol{B}_j)$ of $\boldsymbol{B}_j, j = 1, \ldots, 7$ using $\mathbb{E}_n^*[I(A = D(\boldsymbol{X}))\boldsymbol{B}/P(A, R = 1|\boldsymbol{X})]/\mathbb{E}_n^*[I(A = D(\boldsymbol{X}))/P(A, R = 1|\boldsymbol{X})]$.[30] The estimated value functions of the longitudinal outcomes for each visit time and their average values are summarized in Table 4. Clearly, the LS-learning method gives the highest average estimated value function, yielding the lowest continuous lead exposure compared to the other three decision rules derived by standard OWL. This suggests that the utility of longitudinal auxiliary data via surrogate outcomes $\boldsymbol{B}_j, j = 1, \ldots, 6$ arguably improves the estimated ITR to maximize the benefit to children's growth and development. This improvement is rooted in the successful use on the ordering of time-course relevance among longitudinal outcomes in the multi-view extension of OWL with multiple training datasets.

We also compare the distribution of individual biomarkers between the two resulting treatment groups by using the ITR derived from LS-learning; see Table 5, by which we identify those biomarkers that are significantly different between the two groups. We found that mothers assigned to take calcium supplementation, as specified by the LS-learning-derived ITR, are significantly older. That is, children of older mother are more likely to benefit from maternal calcium supplementation to reduce persistent lead exposure at age 3. In addition, pregnant women assigned to calcium supplement via the LS-learning-derived ITR have significantly longer years in school, lower dietary intake of calcium, higher dietary intake of fiber, and lower maternal blood lead concentration (PBM) at the first trimester. Notably, mothers assigned to the calcium supplementation group via the LS-learning-derived ITR do have a significant difference in their calcium intake from food compared to mothers assigned to the placebo group (1059.6 (521.2) versus 1172.9 (514.9)). This suggests that the maternal dietary calcium intake at baseline does influence child's persistent lead exposure at age 3 years. This finding has clinical value and is consistent with our common sense that calcium supplementation should be specified for pregnant women who do not receive adequate calcium intake from their

daily diet. In contrast, no significant differences are identified between the two resulting allocation groups in terms of maternal weight, marital status, total number of pregnancies, HgB concentration, dietary intake of iron, zinc, and vitamin C for the ITR derived by the LS-learning method.

5 | SIMULATION EXPERIMENT

5.1 | Simulation Setting

This section reports a simulation experiment to illustrate the finite sample performance of scaled-tuned LS-learning in the derivation of ITR using longitudinal trial data with missing information. We randomly generate a 10-dimensional feature vector $\boldsymbol{X} = (X_1, \ldots, X_{10})^\top \in \mathbb{R}^{10}$ for each subject, with each feature $X_v \sim U(0,1), v = 1, \ldots, 10$ and $Corr(X_v, X_{v'}) = 0.2, v \neq v'$. Treatment $A \in \{-1, 1\}$ is randomly assigned to each subject with equal probability, $P(A = 1|\boldsymbol{X}) = P(A = -1|\boldsymbol{X}) = 0.5$. In addition to a baseline visit, subjects are measured at $m = 5$ follow-up times, with a scaled time of $t_j = j/m$ and a standardized time $\tilde{t}_j = t_j/t_m, j = 1, \ldots, m$. A set of cubic spline plus basis function $\boldsymbol{s}(t_j)$ with three knots is specified at $k_1 = 0.4, k_2 = 0.6$ and $k_3 = 0.8$, resulting in seven basis functions $\boldsymbol{s}(t_j) = (1, t_j, t_j^2, t_j^3, (t_j - k_1)_+^3, (t_j - k_2)_+^3, (t_j - k_3)_+^3)^\top$ at visit j to represent the time trajectory. The term $\boldsymbol{s}(t_j)$ is created to model the nonparametric effect of time on the influence of benefit outcome $\boldsymbol{B}_j$. The underlying true decision function f is specified as $f(\boldsymbol{X}) = 1 + X_1 - \log(X_2 + 1) + 2X_3^3 - \exp(X_4)$. The outcomes measured of subject i at time t_j is generated by an equation: $B_{ij} = 0.01 + 0.02X_{i,1} + \boldsymbol{s}(t_j)^\top \boldsymbol{\beta}_3 + \{\boldsymbol{s}(t_j)^\top \boldsymbol{\beta}_4\}\{0.1(0.4X_{i,5} + 0.6X_{i,6} - X_{i,7})\} + 3A_i f(\boldsymbol{X}_i) + \gamma_i + \epsilon_{ij}$, where $\boldsymbol{\beta}_3 = \boldsymbol{\beta}_4 = (3, 0.5, 0.5, -3.5, -2, -2, -0.1)^\top$, random intercepts $\gamma_i \overset{i.i.d}{\sim} N(0, 0.5)$ and random errors $\epsilon_i \overset{i.i.d}{\sim} MVN(\boldsymbol{0}, \boldsymbol{\Sigma})$ with $\boldsymbol{\Sigma}(\rho)$ being an AR(1) correlation matrix and correlation coefficient $\rho = 0.5$. The missing outcome indicator $\boldsymbol{R}_{ij}$ for each outcome $\boldsymbol{B}_{ij}$ is generated independently according to a Bernoulli distribution with probability (a) $P(R_{i1} = 0) = 0.01$ for the first measurement $\boldsymbol{B}_{i1}$ and (b) logit $P(R_{ij} = 0) = 1 + 2\tilde{t}_j - 1$ for $j = 2, \ldots, m$. Based on the missing patterns, subjects can be stratified into subgroups $\mathcal{S}_j, j = 1, \ldots, m$ based on the subject-specific last observation. The propensity score of treatment assignment for each subject is estimated by a logistic regression model with all the majors effects of the covariates $\boldsymbol{X}$. The total sample size for the training data $\mathcal{S} = \mathcal{S}_1 \cup \cdots \cup \mathcal{S}_m$ is set to $n \in \{500, 800, 1000, 1200\}$. In addition, an independent test dataset with a sample size of 10,000, with the optimal treatment assignment determined by $f(\boldsymbol{X}) > 0$ or not, is generated as external validation data. An ITR is derived using the LS-learning method under linear kernel. Scaled tuning was performed via an exponential function $\psi(\tilde{t}_j|\boldsymbol{\theta}) = \theta_1 e^{\theta_2 \tilde{t}_j}$ with $\theta_2 = \log(1/\theta_1)$ to characterize the relevant time course of the outcomes to the primary endpoint benefit measured at visit m. We set two tuning parameter pools $C \in \{1/128, 1/64, \ldots, 1/2, 1\}$ and $\theta_1 \in \{0.1, 0.2, \ldots, 0.9, 1.0\}$ for tuning parameter selection, respectively. GAM regression with B-spline basis functions (four degrees of freedom for each variable) was used to estimate the mean prediction functions $g_j(\boldsymbol{X})$ and $h_j(\boldsymbol{s}(t_j))$ for parameter tuning. The simulation is repeated for 200 times to produce summary statistics.

5.2 | Simulation Results

We compared the predictive accuracy of the five methods on the validation dataset for the underlying optimal treatment assignment. They are (M1) OWL on the entire training data $\mathcal{S}$ with no missing information. The endpoint $\boldsymbol{B}_m$ is fully observed and used for ITR derivation based on the standard OWL. This serves as the gold standard and is referred to as the super oracle method for comparison purposes. Note that endpoint benefit $\boldsymbol{B}_m$ is often not available in practice due to missing data. (M2) LS-learning with scaled tuning on the entire training data $\mathcal{S}$ with missing information. It incorporates auxiliary benefits $\boldsymbol{B}_j, j = 1, \ldots, m-1$ prior to the endpoint $\boldsymbol{B}_m$ using subject stratification. We derive ITR using the LS-learning method that takes into account the relevance of the time-course outcomes. (M3) OWL on the entire training data $\mathcal{S}$ with missing data imputed by the method of LOCF. It differs from M2 in that it ignores the incorporation of time course outcomes. (M4) OWL on the whole training data with missing data imputed by the method of PMM. The imputed endpoint $\boldsymbol{B}_m$ is used for ITR derivation via the standard OWL. A total of 50 imputed datasets were created using all available data for $\boldsymbol{X}$, A and the observed benefits, from which the average statistics of OWL performances were calculated. (M5) OWL on the training data of available endpoint benefits from those who are the completers of the clinical trial. All standard OWL-based methods, including M1, M3, M4, and M5, use 5-fold cross-validation with respect to the maximization of estimated value function for tuning parameter selection as performed in.[13] Tuning of the proposed LS-learning method M2 was performed according to the minimization of SSE as discussed in Section 3.5.

Table 6 lists the average prediction accuracy and the standard deviation (sd) of the underlying optimal treatment assignment on the validation dataset obtained by the five different ITR derivation methods. The results show that, as expected, the super

oracle M1 method always gives the highest prediction accuracy because it uses the complete information to derive the ITR. However, M1 is not feasible for use in practice in the presence of missing data. The LS-learning method M2, developed with missing information, gives a slightly lower prediction accuracy compared to M1. Obviously, M3, M4, and M5 give much lower prediction accuracy than M2. In addition to the prediction accuracy, we also calculated the estimated value functions evaluated at the five outcomes $\hat{V}(B_j)$ on the validation dataset (see Table 7). It is shown that the super oracle M1 method always gives the highest estimated value function as expected. The LS-learning method M2 with scaled tuning method gives higher estimated value functions compared to the other three OWL methods M3, M4, and M5.

In LS-learning, the mean and (sd) values of the selected tuning parameters θ_1 and θ_2 in the exponential function, as well as the resulting $\lambda_j, j = 1, 2, 3, 4$ corresponding to B_1 to B_4 are shown in Table 8. There is no tuning parameter for B_5 since it is the primary benefit outcome used for ITR derivation. Table 8 shows a decreasing trend for the relevance of outcomes from the endpoint to the time of the first visit. But the relevance remains strong since $\mathrm{x}\lambda_1 > 0.7$, which is a relatively large weight of overlap for B_1 with B_5. It is meaningful to compare the computation time of LS-learning with greedy tuning and LS-learning with the proposed scaled tuning. Scaled tuning is proposed to save the computation and cost of LS-learning by reducing the number of tuning parameters. In this simulation, with greedy tuning, we need to tune a total of four λ values ($\lambda_j, j = 1, \ldots, 4$), which are reduced to one (θ_1) by scaled tuning. The latter takes an average of 267.60 seconds (i.e., 4.46 minutes) when the sample size of the training data equals 1,000, while greedy tuning is estimated to take 391,793.2 seconds (i.e., 4.53 days) under four grid sequences $\lambda_j \in \{0, 0.1,' \ldots, 1.0\}, j = 1, \ldots, 4$. It is clear that the computational efficiency of the scaled tuning method makes LS-learning practically feasible.

6 | CONCLUDING REMARKS

In this article, we utilize a new learning approach, called longitudinal self-learning (LS-learning), to establish ITRs in that we integrate longitudinal data sources with time-varying relevance to the primary clinical endpoint as well as with dropouts. The term "self-learning" is used to describe the process whereby the algorithm uses the primary benefit outcome B_m to predict the optimal treatment assignment $\hat{y}_{ij}^{(k-1)}$, which is then used as part of the objective function to exercise a self-examination. Our method provides a useful extension to the existing one-dimensional cross-sectional OWL by allowing for temporally correlated outcomes in the search for the optimal ITR. To synergize different training data sources with different degrees of relevance to the primary data source, we introduce additional tuning parameters in SVM. The interpretation of the new tuning parameters is intuitive: the closer the tuning parameter is to 1, the higher the relevance of the corresponding auxiliary data source with the primary one. The proposed approach is particularly attractive in practical clinical studies where repeated measurements are often collected, which can lead to data attrition and loss of statistical power if appropriate strategies are not employed to handle this problem.

We applied LS-learning to derive an ITR for pregnant women in their daily calcium supplement intake to reduce persistent lead exposure in their children at age 3 years. The main difficulty in analyzing the data from this clinical trial is the large number of missing values for PBC during the three years of the study. LS-learning stratifies subjects into subgroups of different outcomes according to their respective missing patterns in a manner similar to that considered in the pattern mixture model proposed by.[16] It is shown that the new ITR derived by LS-learning would lead to a greater reduction of persistent lead exposure to children compared to other ITRs given by the standard OWL learning if it were implemented for the whole population of pregnant women throughout the study. In addition to the real data application, comprehensive simulation experiments in Section 5 show that LS-learning outperforms standard OWL in ITR estimation with respect to the prediction accuracy as well as the estimated value function. In the meantime, we have identified several baseline biomarkers that may play an important role in the allocating of calcium supplements, including maternal age, total years in school, and maternal blood lead concentration at the first trimester. In addition, dietary intake of calcium and fiber may affect ITR for the allocation of calcium supplementation.

A limitation of our data analysis is the use of the linear kernel for ITR derivation because of its ease in illustrating the estimated ITRs. The nonlinear kernel can also be applied to derive more flexible decision rules, which may be further explored as a future extension. Another challenge relates to the imbalance of treatment assignment in the cleaned data, mainly caused by missing baseline biomarkers. We used the means of propensity score via inverse probability weighting (IPW) to adjust for treatment allocation bias. Other methods, like augmented inverse probability weighting (AIPW)[31] and overlap weighting,[32] are worth further exploration to overcome treatment assignment bias in the ITR analysis of calcium supplementation clinical trial.

ACKNOWLEDGMENTS

The authors would like to thank the co-editor, the associate editor, and two anonymous referees for their valuable comments that helped improve this article significantly.

Conflict of interest

The authors declare no potential conflict of interests.

Data Availability Statement

The data used in this paper cannot be freely shared with the public since restrictions apply to the availability of these data, which were used under license for this study. Readers who want to analyze the data should contact the ELEMENT project manager via the corresponding author.

Supporting Information

Codes for simulations described in Section 5 are available as part of the online article. Simulation results showing the prediction accuracy of the optimal treatment assignment based on different tuning procedures are also included.

□

APPENDIX

PROOF OF THEOREM 1

Proof of Theorem 1. According to **Algorithm 1**, at iteration k ($k \geq 2$), $(\hat{\boldsymbol{\omega}}^{(k)}, \hat{b}^{(k)}, \hat{\boldsymbol{\xi}}^{(k)})$ is the optimal solution of the optimization problem (2). First we prove that $(\hat{\boldsymbol{\omega}}^{(k-1)}, \hat{b}^{(k-1)}, \hat{\boldsymbol{\xi}}^{(k-1)})$ is a feasible solution of (2). It immediately follows that $h(\hat{\boldsymbol{\omega}}^{(k)}, \hat{\boldsymbol{\xi}}^{(k)}) \leq h(\hat{\boldsymbol{\omega}}^{(k-1)}, \hat{\boldsymbol{\xi}}^{(k-1)})$ due to the fact that $(\hat{\boldsymbol{\omega}}^{(k)}, \hat{b}^{(k)}, \hat{\boldsymbol{\xi}}^{(k)})$ is the optimal solution for (2). The equality holds when $h(\hat{\boldsymbol{\omega}}^{(k-1)}, \hat{\boldsymbol{\xi}}^{(k-1)})$ reaches the minimum. To proceed, it is sufficient to show that the following conditions hold,

$$\begin{aligned}
&A_i(\hat{\boldsymbol{\omega}}^{(k-1)\top}\boldsymbol{X}_i + \hat{b}^{(k-1)}) \geq 1 - \hat{\xi}_i^{(k-1)}, \quad \text{and} \quad \hat{\xi}_i^{(k-1)} \geq 0,\ i \in \mathcal{S}_m,\\
&A_i(\hat{\boldsymbol{\omega}}^{(k-1)\top}\boldsymbol{X}_i + \hat{b}^{(k-1)}) \geq 1 - \tilde{\tilde{\xi}}_{ij}^{(k-1)}, \quad \text{and} \quad \tilde{\tilde{\xi}}_{ij}^{(k-1)} \geq 0,\ i \in \mathcal{S}_j^{(k)}, j \in \mathcal{J},\\
&\hat{y}_{ij}^{(k-1)}(\hat{\boldsymbol{\omega}}^{(k-1)\top}\boldsymbol{X}_i + \hat{b}^{(k-1)}) \geq 1 - \hat{\tilde{\xi}}_{ij}^{(k-1)}, \quad \text{and} \quad \hat{\tilde{\xi}}_{ij}^{(k-1)} \geq 0,\ i \in \mathcal{S}_j^{(k)}, j \in \mathcal{J}.
\end{aligned} \tag{1}$$

Note that at iteration $k-1$, $(\hat{\boldsymbol{\omega}}^{(k-1)}, \hat{b}^{(k-1)}, \hat{\boldsymbol{\xi}}^{(k-1)})$ is the solution of problem (2) with $k-1$ replaced by $k-2$.

$$\begin{aligned}
\min_{\boldsymbol{\omega},b,\boldsymbol{\xi}} \quad & \frac{1}{2}||\boldsymbol{\omega}||^2 + C\left[\sum_{i\in\mathcal{S}_m}\frac{B_{im}}{p_i}\xi_i + \sum_{j=1}^{m-1}\left\{\lambda_j \sum_{i\in\mathcal{S}_j^{(k-1)}}\frac{B_{ij}}{p_i}\tilde{\xi}_{ij} + (1-\lambda_j)\sum_{i\in\mathcal{S}_j^{(k-1)}}\tau_{B_j}\check{\xi}_{ij}\right\}\right]\\
\text{s.t.} \quad & A_i(\boldsymbol{\omega}^\top\boldsymbol{X}_i + b) \geq 1 - \xi_i \quad \text{and} \quad \xi_i \geq 0,\ i \in \mathcal{S}_m,\\
& A_i(\boldsymbol{\omega}^\top\boldsymbol{X}_i + b) \geq 1 - \tilde{\xi}_{ij} \quad \text{and} \quad \tilde{\xi}_{ij} \geq 0,\ i \in \mathcal{S}_j^{(k-1)},\ j \in \mathcal{J},\\
& \hat{y}_{ij}^{(k-2)}(\boldsymbol{\omega}^\top\boldsymbol{X}_i + b) \geq 1 - \check{\xi}_{ij} \quad \text{and} \quad \check{\xi}_{ij} \geq 0,\ i \in \mathcal{S}_j^{(k-1)},\ j \in \mathcal{J}.
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
&A_i(\hat{\boldsymbol{\omega}}^{(k-1)\top}\boldsymbol{X}_i + \hat{b}^{(k-1)}) \geq 1 - \hat{\xi}_i^{(k-1)}, \quad \text{and} \quad \hat{\xi}_i^{(k-1)} \geq 0,\ i \in \mathcal{S}_m\\
&A_i(\hat{\boldsymbol{\omega}}^{(k-1)\top}\boldsymbol{X}_j + \hat{b}^{(k-1)}) \geq 1 - \tilde{\tilde{\xi}}_{ij}^{(k-1)}, \quad \text{and} \quad \tilde{\tilde{\xi}}_{ij}^{(k-1)} \geq 0,\ i \in \mathcal{S}_j^{(k-1)}, j \in \mathcal{J}.
\end{aligned}$$

That is, the constraints involving treatment A in (1) are satisfied by $(\hat{\boldsymbol{\omega}}^{(k-1)}, \hat{b}^{(k-1)}, \hat{\boldsymbol{\xi}}^{(k-1)})$. In addition, we have

$$\hat{y}_{ij}^{(k-2)}(\hat{\boldsymbol{\omega}}^{(k-1)\top}\boldsymbol{X}_i + \hat{b}^{(k-1)}) \geq 1 - \hat{\tilde{\xi}}_{ij}^{(k-1)}, \quad \text{and} \quad \hat{\tilde{\xi}}_{ij}^{(k-1)} \geq 0,\ i \in \mathcal{S}_j^{(k-1)}, j \in \mathcal{J}. \tag{2}$$

To prove the constraints involving $\hat{y}_{ij}^{(k-1)}$ in (1), we consider two scenarios between $\hat{y}_{ij}^{(k-1)}$ and $\hat{y}_{ij}^{(k-2)}$ for subject $i \in S_j$. First, if $\hat{y}_{ij}^{(k-1)} = \hat{y}_{ij}^{(k-2)}$ is true, we have

$$\hat{y}_{ij}^{(k-1)}(\hat{\boldsymbol{\omega}}^{(k-1)T}\boldsymbol{X}_i + \hat{b}^{(k-1)}) = \hat{y}_{ij}^{(k-2)}(\hat{\boldsymbol{\omega}}^{(k-1)T}\boldsymbol{X}_i + \hat{b}^{(k-1)}) \geq 1 - \hat{\xi}_{ij}^{(k-1)}.$$

Second, if $\hat{y}_{ij}^{(k-1)} \neq \hat{y}_{ij}^{(k-2)}$ is true, then

$$\hat{y}_{ij}^{(k-1)}(\hat{\boldsymbol{\omega}}^{(k-1)T}\boldsymbol{X}_i + \hat{b}^{(k-1)}) > \hat{y}_{ij}^{(k-2)}(\hat{\boldsymbol{\omega}}^{(k-1)T}\boldsymbol{X}_i + \hat{b}^{(k-1)}) \geq 1 - \hat{\xi}_{ij}^{(k-1)}.$$

This is because labels are binary and can only take values from $\{-1, 1\}$. Therefore, we must have $\hat{y}_{ij}^{(k-1)} = -\hat{y}_{ij}^{(k-2)}$ if $\hat{y}_{ij}^{(k-1)} \neq \hat{y}_{ij}^{(k-2)}$. Also, from the definition of $\hat{y}_{ij}^{(k-1)}$, we have $\hat{y}_{ij}^{(k-1)}(\hat{\boldsymbol{\omega}}^{(k-1)T}\boldsymbol{X}_i + \hat{b}^{(k-1)}) > 0$. Therefore, $\hat{y}_{ij}^{(k-2)}(\hat{\boldsymbol{\omega}}^{(k-1)T}\boldsymbol{X}_i + \hat{b}^{(k-1)}) < 0$.

Combines the above two cases, we obtain

$$\hat{y}_{ij}^{(k-1)}(\hat{\boldsymbol{\omega}}^{(k-1)T}\boldsymbol{X}_i + \hat{b}^{(k-1)}) \geq 1 - \hat{\xi}_{ij}^{(k-1)}, i \in S_j^{(k)}, j \in \mathcal{J},$$

where the equality only occurs when the algorithm converges with the predicted labels $\hat{y}_{ij}$ stopping changing between two adjacent iterations $k-1$ and $k-2$.

Thus, all constraints in (2) have been proved to hold. This implies that all constraints in (1) hold. This completes the proof of the descending property $h(\hat{\boldsymbol{\omega}}^{(k)}, \hat{\boldsymbol{\xi}}^{(k)}) < h(\hat{\boldsymbol{\omega}}^{(k-1)}, \hat{\boldsymbol{\xi}}^{(k-1)})$ when $k \geq 2$. □ □

References

1. Wadhwa PD, Buss C, Entringer S, Swanson JM. Developmental origins of health and disease: brief history of the approach and current focus on epigenetic mechanisms. In: . 27. © Thieme Medical Publishers. ; 2009: 358–368.

2. Hornung RW, Lanphear BP, Dietrich KN. Age of greatest susceptibility to childhood lead exposure: a new statistical approach. *Environ Health Perspect* 2009; 117(8): 1309–1312.

3. Téllez-Rojo MM, Bellinger DC, Arroyo-Quiroz C, et al. Longitudinal associations between blood lead concentrations lower than 10 μg/dL and neurobehavioral development in environmentally exposed children in Mexico City. *Pediatrics* 2006; 118(2): e323–e330.

4. Jusko TA, Henderson Jr CR, Lanphear BP, Cory-Slechta DA, Parsons PJ, Canfield RL. Blood lead concentrations< 10 μg/dL and child intelligence at 6 years of age. *Environ Health Perspect* 2008; 116(2): 243–248.

5. Canfield RL, Henderson Jr CR, Cory-Slechta DA, Cox C, Jusko TA, Lanphear BP. Intellectual impairment in children with blood lead concentrations below 10 μg per deciliter. *N Engl J Med* 2003; 348(16): 1517–1526.

6. Gulson B, Mahaffey K, Mizon K, Korsch M, Cameron M, Vimpani G. Contribution of tissue lead to blood lead in adult female subjects based on stable lead isotope methods.. *J Lab Clin Med* 1995; 125(6): 703–712.

7. Gulson BL, Mizon KJ, Korsch MJ, Palmer JM, Donnelly JB. Mobilization of lead from human bone tissue during pregnancy and lactation—a summary of long-term research. *Sci Total Environ* 2003; 303(1-2): 79–104.

8. Ettinger AS, Hu H, Hernandez-Avila M. Dietary calcium supplementation to lower blood lead levels in pregnancy and lactation. *J Nutr Biochem* 2007; 18(3): 172–178.

9. Gulson BL, Mizon KJ, Palmer JM, Korsch MJ, Taylor AJ, Mahaffey KR. Blood lead changes during pregnancy and postpartum with calcium supplementation. *Environ Health Perspect* 2004; 112(15): 1499–1507.

10. Téllez-Rojo M, Lamadrid-Figueroa H, Mercado-García A, et al. A randomized controlled trial of calcium supplementation to reduce blood lead levels (and fetal lead exposure) in pregnant women. *Epidemiology* 2006; 17(6): S123.

11. Norman J, Politz D, Politz L. Hyperparathyroidism during pregnancy and the effect of rising calcium on pregnancy loss: a call for earlier intervention. *Clin Endocrinol (Oxf)* 2009; 71(1): 104–109.

12. Qian M, Murphy SA. Performance guarantees for individualized treatment rules. *Ann Stat* 2011; 39(2): 1180.

13. Zhao Y, Zeng D, Rush AJ, Kosorok MR. Estimating individualized treatment rules using outcome weighted learning. *J Am Stat Assoc* 2012; 107(499): 1106–1118.

14. Zhou X, Mayer-Hamblett N, Khan U, Kosorok MR. Residual weighted learning for estimating individualized treatment rules. *J Am Stat Assoc* 2017; 112(517): 169–187.

15. Huling JD, Yu M, Smith M. Fused comparative intervention scoring for heterogeneity of longitudinal intervention effects. *The Annals of Applied Statistics* 2019; 13(2): 824–847.

16. Little RJ. Selection and pattern-mixture models. *Longitudinal data analysis* 2008: 409–431.

17. Zhang A, Hu H, Sánchez BN, et al. Association between prenatal lead exposure and blood pressure in children. *Environ Health Perspect* 2012; 120(3): 445–450.

18. Wasserman GA, Staghezza-Jaramillo B, Shrout P, Popovac D, Graziano J. The effect of lead exposure on behavior problems in preschool children.. *Am J Public Health* 1998; 88(3): 481–486.

19. Mendelsohn AL, Dreyer BP, Fierman AH, et al. Low-level lead exposure and behavior in early childhood. *Pediatrics* 1998; 101(3): e10–e10.

20. Robins JM. Marginal structural models. 1997.

21. Cortes C, Vapnik V. Support-vector networks. *Mach Learn* 1995; 20(3): 273–297.

22. Zeger SL, Diggle PJ. Semiparametric models for longitudinal data with application to CD4 cell numbers in HIV seroconverters. *Biometrics* 1994: 689–699.

23. Imai K, Ratkovic M. Estimating treatment effect heterogeneity in randomized program evaluation. 2013.

24. Lin Y. Support vector machines and the Bayes rule in classification. *Data Mining and Knowledge Discovery* 2002; 6: 259–275.

25. Zhang T. Statistical behavior and consistency of classification methods based on convex risk minimization. *The Annals of Statistics* 2004; 32(1): 56–85.

26. Fisher RA. The use of multiple measurements in taxonomic problems. *Annals of eugenics* 1936; 7(2): 179–188.

27. Shi C, Song R, Lu W. Concordance and value information criteria for optimal treatment decision. *Ann Stat* 2021; 49(1): 49–75.

28. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. *J Stat Softw* 2011; 45(3): 1-67.

29. Molnar FJ, Hutton B, Fergusson D. Does analysis using "last observation carried forward" introduce bias in dementia research?. *CMAJ* 2008; 179(8): 751–753.

30. Murphy SA, Laan v. dMJ, Robins JM, Group CPPR. Marginal mean models for dynamic regimes. *J Am Stat Assoc* 2001; 96(456): 1410–1423.

31. Glynn AN, Quinn KM. An introduction to the augmented inverse propensity weighted estimator. *Polit Anal* 2010: 36–56.

32. Thomas LE, Li F, Pencina MJ. Overlap weighting: a propensity score method that mimics attributes of a randomized clinical trial. *JAMA* 2020; 323(23): 2417–2418.

TABLE 1 Summary statistics of the biomarkers included for ITR derivation from the calcium supplementation trial. Mean (sd) and percentage values are shown for numeric and categorical variables, where p-values are obtained from Wilcoxon rank-sum test and chi-aquare test for numeric and categorical variables, respectively.

Biomarker	Calcium	Placebo	p-value
age	26.9 (5.7)	25.9 (5.4)	7.10×10^{-2}
weight	62.0 (11.3)	61.6 (9.6)	7.17×10^{-1}
years in school	10.8 (2.9)	10.7 (2.9)	6.00×10^{-1}
marital status	0.695	0.677	8.02×10^{-1}
number of pregnancies	2.1 (1.0)	2.1 (1.1)	8.67×10^{-1}
HgB concentration	13.2 (1.0)	13.2 (1.0)	8.00×10^{-1}
dietary intake of calcium	1115.5 (489.3)	1078.6 (552.5)	4.93×10^{-1}
dietary intake of fiber	24.4 (10.0)	22.8 (9.0)	1.25×10^{-1}
dietary intake of iron	13.7 (5.9)	12.9 (5.5)	2.23×10^{-1}
dietary intake of zinc	9.8 (3.5)	9.3 (3.6)	1.73×10^{-1}
dietary intake of vitamin C	178.1 (92.1)	172.7 (80.4)	5.48×10^{-1}
PBM concentration	4.7 (2.7)	5.3 (3.7)	7.00×10^{-2}

TABLE 2 Estimated intercepts and coefficients of the biomarkers in the estimated ITRs derived by the four different methods.

Biomarker	OWL on S_m	OWL (PMM)	OWL (LOCF)	LS-learning
intercept	0.46	0.05	-0.05	0.39
age	0.33	0.25	0.24	0.29
weight	0.02	0.08	0.05	0.01
years in school	-0.08	-0.06	-0.09	-0.05
marital status	0.24	0.13	0.17	0.19
number of pregnancies	-0.57	0.15	0.23	-0.23
HgB concentration	0.10	0.08	0.08	0.04
dietary intake of calcium	-0.47	-0.25	-0.34	-0.37
dietary intake of fiber	0.28	0.23	0.26	0.30
dietary intake of iron	-0.15	0.10	0.02	-0.13
dietary intake of zinc	0.38	0.27	0.39	0.41
dietary intake of vitamin C	-0.02	-0.26	-0.28	-0.14
PBM concentration	-0.37	-0.33	-0.33	-0.20

TABLE 3 Treatment assignment comparison between complete randomization and the ITR derived by LS-learning.

		LS-learning		
		placebo	calcium	total
randomization	placebo	57	129	186
	calcium	68	122	190
	total	125	251	376

TABLE 4 Estimated value functions at each visit and the average based on the estimated ITRs derived by the four different methods in the calcium supplementation trial.

ITR	$\hat{V}(B_1)$	$\hat{V}(B_2)$	$\hat{V}(B_3)$	$\hat{V}(B_4)$	$\hat{V}(B_5)$	$\hat{V}(B_6)$	$\hat{V}(B_7)$	Average
OWL on S_m	6.79	3.39	5.49	7.33	7.84	7.92	8.24	6.714
OWL (PMM)	6.79	3.39	5.49	7.37	8.09	7.90	8.23	6.751
OWL (LOCF)	6.79	7.16	5.49	7.33	8.46	8.01	8.24	7.354
LS-learning	9.19	7.19	6.36	7.59	8.55	8.06	8.33	7.896

TABLE 5 Summary statistics of the biomarkers based on treatment allocation according to the derived ITR by LS-learning. Mean (sd) and percentage values are shown, where p-values are obtained from Wilcoxon rank-sum test and chi-aquare test for numeric and categorical variables, respectively.

Biomarker	Calcium	Placebo	p-value
age	27.5 (5.4)	24.2 (5.1)	2.94×10^{-8}
weight	62.6 (10.7)	60.4 (10.1)	5.19×10^{-2}
years in school	11.0 (3.0)	10.3 (2.7)	2.72×10^{-3}
marital status	0.721	0.616	5.10×10^{-2}
number of pregnancies	2.1 (1.1)	2.0 (1.0)	6.20×10^{-1}
HgB concentration	13.3 (0.9)	13.1 (1.2)	9.37×10^{-1}
dietary intake of calcium	1059.6 (521.2)	1172.9 (514.9)	2.37×10^{-2}
dietary intake of fiber	24.5 (9.8)	21.8 (8.7)	1.10×10^{-2}
dietary intake of iron	13.6 (5.8)	12.7 (5.5)	1.45×10^{-1}
dietary intake of zinc	9.7 (3.7)	9.0 (3.3)	1.15×10^{-1}
dietary intake of vitamin C	178.2 (87.5)	169.8 (84.4)	4.55×10^{-1}
PBM concentration	4.7 (3.1)	5.5 (3.5)	3.70×10^{-3}

TABLE 6 Average prediction accuracy (mean (sd)) calculated using ITRs derived by the five different methods.

n	M1: super oracle	M2: LS-learning	M3: OWL (LOCF)	M4: OWL (PMM)	M5: OWL on S_m
500	0.783 (0.049)	0.769 (0.057)	0.740 (0.059)	0.710 (0.064)	0.605 (0.092)
800	0.816 (0.037)	0.799 (0.045)	0.774 (0.051)	0.751 (0.057)	0.641 (0.096)
1000	0.832 (0.032)	0.812 (0.036)	0.794 (0.046)	0.776 (0.052)	0.658 (0.087)
1200	0.838 (0.032)	0.818 (0.037)	0.803 (0.041)	0.785 (0.047)	0.670 (0.086)

TABLE 7 Estimated value functions $\hat{V}(B_1)$ to $\hat{V}(B_5)$ calculated using ITRs derived by the five different methods.

	n	M1: super oracle	M2: LS-learning	M3: OWL (LOCF)	M4: OWL (PMM)	M5: OWL on $\mathcal{S}_m$
$\hat{V}(B_1)$	500	11.00 (0.45)	10.97 (0.47)	10.89 (0.47)	10.80 (0.49)	10.45 (0.54)
	800	11.09 (0.44)	11.05 (0.45)	10.99 (0.46)	10.93 (0.47)	10.58 (0.51)
	1000	11.09 (0.44)	11.05 (0.45)	11.01 (0.45)	10.96 (0.46)	10.61 (0.53)
	1200	11.11 (0.43)	11.07 (0.43)	11.03 (0.43)	10.99 (0.43)	10.65 (0.50)
$\hat{V}(B_2)$	500	10.97 (0.45)	10.93 (0.47)	10.85 (0.47)	10.76 (0.48)	10.41 (0.54)
	800	11.06 (0.44)	11.02 (0.45)	10.96 (0.46)	10.89 (0.47)	10.54 (0.51)
	1000	11.06 (0.44)	11.02 (0.45)	10.97 (0.45)	10.92 (0.46)	10.57 (0.53)
	1200	11.07 (0.43)	11.03 (0.43)	11.00 (0.43)	10.95 (0.43)	10.62 (0.50)
$\hat{V}(B_3)$	500	10.62 (0.45)	10.58 (0.47)	10.50 (0.47)	10.41 (0.49)	10.07 (0.54)
	800	10.71 (0.44)	10.67 (0.45)	10.61 (0.46)	10.54 (0.47)	10.20 (0.51)
	1000	10.71 (0.44)	10.67 (0.45)	10.62 (0.45)	10.57 (0.46)	10.23 (0.53)
	1200	10.72 (0.43)	10.68 (0.43)	10.65 (0.43)	10.60 (0.43)	10.27 (0.50)
$\hat{V}(B_4)$	500	9.70 (0.45)	9.66 (0.47)	9.58 (0.47)	9.49 (0.49)	9.14 (0.54)
	800	9.79 (0.44)	9.75 (0.45)	9.69 (0.46)	9.62 (0.47)	9.27 (0.51)
	1000	9.79 (0.44)	9.74 (0.45)	9.70 (0.45)	9.65 (0.46)	9.30 (0.53)
	1200	9.80 (0.43)	9.76 (0.43)	9.73 (0.43)	9.68 (0.43)	9.34 (0.50)
$\hat{V}(B_5)$	500	7.85 (0.46)	7.81 (0.47)	7.74 (0.47)	7.64 (0.49)	7.30 (0.54)
	800	7.94 (0.44)	7.90 (0.45)	7.84 (0.46)	7.78 (0.47)	7.43 (0.51)
	1000	7.94 (0.44)	7.90 (0.45)	7.86 (0.45)	7.81 (0.46)	7.46 (0.53)
	1200	7.95 (0.43)	7.91 (0.43)	7.88 (0.43)	7.83 (0.43)	7.50 (0.50)

TABLE 8 The selected tuning parameter θ_1 and θ_2 in LS-learning and the resulting $\lambda_j, j = 1, 2, 3, 4$ values corresponding to the surrogate outcomes B_1 to B_4.

n	θ_1	θ_2	λ_1	λ_2	λ_3	λ_4
500	0.663 (0.314)	0.583 (0.669)	0.705 (0.284)	0.755 (0.244)	0.818 (0.189)	0.898 (0.111)
800	0.730 (0.290)	0.446 (0.598)	0.765 (0.261)	0.807 (0.222)	0.858 (0.171)	0.921 (0.100)
1000	0.727 (0.294)	0.453 (0.596)	0.762 (0.264)	0.804 (0.225)	0.856 (0.173)	0.920 (0.101)
1200	0.783 (0.277)	0.357 (0.568)	0.812 (0.249)	0.845 (0.212)	0.886 (0.163)	0.937 (0.095)