Journal of the Royal Statistical Society Series C: Applied Statistics, 2023, **00**, 1–19
https://doi.org/10.1093/jrsssc/qlad087

ROYAL STATISTICAL SOCIETY DATA | EVIDENCE | DECISIONS
Journal of the Statistics Society Series C Applied Statistics

Original Article

Multivariate longitudinal analysis for the association between brain atrophy and cognitive impairment in prodromal Huntington's disease subjects

Cheng Zheng, Lili Tong and Ying Zhang

Department of Biostatistics, University of Nebraska Medical Center, Omaha, USA

Address for correspondence: Cheng Zheng, Department of Biostatistics, University of Nebraska Medical Center, Omaha, NE 68198, USA. Email: cheng.zheng@unmc.edu

Abstract

Cognitive impairment has been widely accepted as a disease progression measure prior to the onset of Huntington's disease. We propose a sophisticated measurement error correction method that can handle potentially correlated measurement errors in longitudinally collected exposures and multiple outcomes. The asymptotic theory for the proposed method is developed. A simulation study is conducted to demonstrate the satisfactory performance of the proposed two-stage fitting method and shows that the independent working correlation structure outperforms other alternatives. We conduct a comprehensive longitudinal analysis to assess how brain striatal atrophy affects impairment in various cognitive domains for Huntington's disease.

Keywords: disease progression modelling, generalised estimating equation, Huntington's disease, longitudinal data, measurement error

1 Introduction

Huntington's disease (HD) is an autosomal dominant neurodegenerative disease caused by the trinucleotide cytosine–adenine–guanine (CAG) repeats expansion in the first exon of the Huntingtin (HTT) gene. Individuals with CAG repeats 36 or greater are considered HD gene-penetrated and likely to develop HD with a mean disease onset age of around 45 (Ross et al., 2014). The age of HD motor onset is inversely related to the CAG repeats (Zhang et al., 2011). Currently, there is no effective treatment for HD patients.

Though HD manifests as a triad of clinical symptoms in motor, cognition, and psychiatric areas, its onset diagnosis relies primarily on the presence of motor impairment (Paulsen et al., 2008). Many studies have reported cognitive impairment prior to HD motor diagnosis (Duff et al., 2010; Paulsen, 2011; Paulsen et al., 2013; Stout et al., 2012; Zhang et al., 2021). In particular, Zhang et al. (2021) have studied mild cognitive impairment in various cognitive domains studied by Harrington et al. (2012), as an early landmark for the disease progression in prodromal HD individuals.

As is typical in neurodegenerative diseases, brain volume loss occurs in HD subjects. In particular, significant striatal and white-matter atrophy in prodromal-HD subjects was found prior to HD motor onset in some longitudinal studies, and brain volume changes were recommended as outcome measures for potential clinical trials for treating prodromal-HD subjects (Aylward et al., 2011; Tabrizi et al., 2012), because they directly affect daily functionality along with the disease progression. While the association between brain volume loss and cognitive impairment was studied in a limited scope (Ahveninen et al., 2018; Dominguez et al., 2017), a more thorough

Received: August 25, 2022. **Revised:** May 15, 2023. **Accepted:** August 26, 2023

exploration of the impact of brain atrophy on cognitive impairments for prodromal-HD subjects has not been done to the best of our knowledge but is imperative to understanding the disease progression of HD.

The 'Neurobiological Predictors of Huntington's Disease Study' (PREDICT-HD) is a 12-year prospective cohort study conducted through 32 sites in 6 countries (USA, Canada, Germany, Australia, Spain, and the UK) from September 2002 to April 2014. The objective of the study was to enrol prodromal-HD subjects who were not tested positive at the study entry defined by the HD diagnostic confidence level rating of the Unified Huntington's Disease Rating Scale ratings less than 4 for the follow-up observations to determine the clinical and biological markers that are predictive of HD motor onset (Paulsen et al., 2014). A total of 1,155 HD gene-expanded subjects (CAG $\geq$ 36) and 317 healthy controls that were family members of the HD gene-positive subjects were recruited. The study gathered 40 longitudinally collected data that contain Unified Huntington's Disease Rating Scale motor assessment and functional outcome measures of capacity and activity scale, multi-domain cognitive measures, psychiatric measures, and magnetic resonance imaging (MRI) measures of intracranial-corrected volumes. It offers scientific promise to ascertain the associations between brain atrophy and functional impairment for prodromal-HD subjects, and this project aims to study if and how the structure of the brain's volume changes correlate with the impairment in the cognitive domains specific to HD (Harrington et al., 2012).

We note that both MRI and cognitive measures based on neuropsychological tests are subject to potential measurement errors that may be correlated. Standard statistical methods, by ignoring measurement error, could result in substantial bias in estimating association parameters between the latent variables without measurement error (for example, true brain volume loss and cognitive impairment) (Carroll et al., 2006). For measurement error in exposure variables for regression with a univariate outcome, various methods have been developed to handle measurement error in epidemiology studies, including regression calibration (Fraser & Stram, 2001), moment reconstruction (Freedman et al., 2004), corrected score (Nakamura, 1990), conditional score (Stefanski & Carroll, 1987), expected estimating equation (Wang & Pepe, 2000), simulation and extrapolation (Cook & Stefanski, 1994), and multiple imputations (Cole et al., 2006). For time-varying exposure variables with measurement error, methods (Buonaccorsi et al., 2000; Lin et al., 2018; Xiao et al., 2010; Yi et al., 2012) have been developed for the univariate outcome as well. Recently, a few methods (Oh et al., 2021; Shaw et al., 2021) have been proposed to handle the correlated measurement error between exposure and outcome. However, to the best of our knowledge, there is no sophisticated statistical method in the literature that can be directly applied to handle multiple longitudinal outcomes with correlated measurement errors. In this paper, we extend the moment-based method to multivariate longitudinal data settings that allow us to handle the correlated measurement error within exposures, within outcomes, and between exposures and outcomes. The proposed method has the advantage of being computationally efficient and can be shown to have standard asymptotic properties using the generalised estimating equation (GEE) theory. This extension fills the gap in measurement error model literature for dealing with complicated multivariate longitudinal data and permits us to study the associations between structure MRI measures and cognitive impairment in great detail, which is much needed to understand the HD progression.

The rest of this paper is organised as follows: in Section 2, we illustrate our notation, models, and assumptions. In Section 3, we propose the GEE method for estimating the model parameters and state the asymptotic results of the method. In Section 4, we present simulation studies to evaluate the finite sample performance of the proposed method. In Section 5, we apply our method to the PREDICT-HD data to identify the regions of brain volume loss that are related to various types of cognitive impairment. In Section 6, we summarise our findings and discuss potential extensions. The technical details are included in the online supplementary materials.

2 Notation and models

We use $X_i \in R^p$ to denote the baseline covariates for individual i. We use $S_i^*(t) \in R^K$ to denote the true outcome variables of interest (i.e. K cognitive domains in HD) for individual i at time t, and we use $S_i(t) \in R^K$ to denote the error-prone outcome variables measured at time t (i.e. test results for K cognitive domains). We use $I_i^*(t) \in R^Q$ to denote the true exposure variables of interest (i.e. brain volumes in Q regions) for individual i at time t and use $I_i(t) \in R^Q$ to denote the error-prone

exposure variables measured at time t (i.e. MRI measurements for the Q regions). We denote the visit times for individual i as t_{il}, $l = 1, \ldots, L_i$, and we do not require L_i to be the same for all individuals. For the sake of notation simplicity, we use $S_{ikl} = S_{ik}(t_{il})$, $S^*_{ikl} = S^*_{ik}(t_{il})$ for $k = 1, \ldots, K$ and $I_{iql} = I_{iq}(t_{il})$, $I^*_{iql} = I^*_{iq}(t_{il})$ for $q = 1, \ldots, Q$, where $l = 1, \ldots, L_i$. The whole histories of $S^*(t)$, $I^*(t)$, $S(t)$, and $I(t)$ are denoted as $\underline{S}^*_{ikl} = \{S^*_{ik1}, \ldots, S^*_{ikl}\}$, $\underline{I}^*_{iql} = \{I^*_{iq1}, \ldots, I^*_{iql}\}$, $\underline{S}_{ikl} = \{S_{ik1}, \ldots, S_{ikl}\}$, and $\underline{I}_{iql} = \{I_{iq1}, \ldots, I_{iql}\}$, respectively.

Let $\mathcal{MGP}_d$ denote a d-dimensional multivariate Gaussian process, and we propose the following measurement model under the classical measurement error assumption:

$$I_i(t) = I^*_i(t) + e_{Ii}(t)$$
$$S_i(t) = S^*_i(t) + g(t; \gamma) + e_{Si}(t)$$

where

$$(e_{Ii}(\cdot), e_{Si}(\cdot)) \sim \mathcal{MGP}_{Q+K}\left(0, \begin{pmatrix} \Sigma_I(\cdot,\cdot) & \Sigma_{IS}(\cdot,\cdot) \\ \Sigma_{SI}(\cdot,\cdot) & \Sigma_S(\cdot,\cdot) \end{pmatrix}\right)$$

are technical measurement errors assumed to be independent of X_i, $I^*_i(\cdot)$, and $S^*_i(\cdot)$. Here $\Sigma_I(t, s) = Cov(e_{Ii}(t), e_{Ii}(s))$, $\Sigma_S(t, s) = Cov(e_{Si}(t), e_{Si}(s))$, $\Sigma_{IS}(t, s) = Cov(e_{Ii}(t), e_{Si}(s))$, and $\Sigma_{SI}(t, s) = Cov(e_{Si}(t), e_{Ii}(s))$ are the variance–covariance matrices. In the model for cognitive measures, $g(t; \gamma) = (g_1(t; \gamma), \ldots, g_K(t; \gamma))^\top$ represents practice effects associated with repeated neuropsychological tests over time that are commonly considered for developing neuropsychological test batteries in studying neurodegenerative diseases (Falleti et al., 2006; Pereira et al., 2015). Since these variance–covariance parameters $(\Sigma_I(\cdot,\cdot), \Sigma_S(\cdot,\cdot), \Sigma_{IS}(\cdot,\cdot))$ are with infinite dimension, to estimate them we usually need to assume certain parametric models $(\Sigma_I(\cdot,\cdot;\zeta_I), \Sigma_S(\cdot,\cdot;\zeta_S), \Sigma_{IS}(\cdot,\cdot;\zeta_{IS}))$ with finite dimension parameters (ζ_I, ζ_S, ζ_{IS} and γ). In our study, we assume that for $t \neq s$, $\Sigma_I(t, s; \zeta_I) = 0$, $\Sigma_S(t, s; \zeta_S) = 0$, $\Sigma_{IS}(t, s; \zeta_{IS}) = 0$ while for all t,

$$\Sigma_I(t, t; \zeta_I) = \begin{pmatrix} \zeta_{I11} & \cdots & \zeta_{I1Q} \\ \cdots & \cdots & \cdots \\ \zeta_{I1Q} & \cdots & \zeta_{IQQ} \end{pmatrix}, \quad \Sigma_S(t, t; \zeta_S) = \begin{pmatrix} \zeta_{S11} & \cdots & \zeta_{S1K} \\ \cdots & \cdots & \cdots \\ \zeta_{S1K} & \cdots & \zeta_{SKK} \end{pmatrix}$$

and

$$\Sigma_{IS}(t, t; \zeta_{IS}) = \begin{pmatrix} \zeta_{IS11} & \cdots & \zeta_{IS1K} \\ \cdots & \cdots & \cdots \\ \zeta_{ISQ1} & \cdots & \zeta_{ISQK} \end{pmatrix}$$

In this article, we are interested in exploring if and how brain volume loss affects cognitive impairment in prodromal-HD subjects based on structural MRI data, so we model the association as follows:

$$S^*_i(t) = b_0 + b_1 t + b_2 X_i + b_3 I^*_i(t) + e_i(t) \tag{1}$$

where $e_i(\cdot) \sim \mathcal{MGP}_K(0, \Sigma_e(\cdot,\cdot))$ has conditional mean of 0 given $I^*_i(\cdot)$, X_i and $b_0 \in R^K$, $b_1 \in R^K$, $b_2 \in R^{K \times p}$, $b_3 \in R^{K \times Q}$ are the regression parameters of interest. Note that we do not require $\Sigma_e(s, t) = 0$ for $s \neq t$ and thus allow for correlation structure from potential random effects. Model (1) adopts the Markov assumption that $S^*_i(t)$ only depends on the current values of $I^*_i(t)$ rather than $\underline{I}^*_i(t)$, the whole history of $I^*_i(t)$.

3 Method

We propose a two-stage method for estimating the regression parameters: $B = (b_0, b_1, b_2, b_3)$. In the first stage, we develop a moment-based estimation method for the practice effects and

variance–covariance components of measurement errors using external data and demonstrate its asymptotic properties. In the second stage, we plug in the estimated practice effects and variance–covariance components of measurement errors to the aforementioned measurement error correction model to obtain the estimate of B and study its asymptotic properties.

3.1 Estimating parameters for practice effects and measurement errors (stage 1)

In this work, we assume that the practice effects and the variance–covariance components of measurement errors are stationary over t and propose to estimate them from external data.

This assumption is reasonable in the PREDICT-HD data, which allows us to get consistent estimates for the parameter γ related to the practice effects $g(t;\gamma)$ as well as parameters $(\zeta_I, \zeta_S, \zeta_{IS})$ for the variance–covariance components $\Sigma_I(\cdot,\cdot)$, $\Sigma_S(\cdot,\cdot)$, $\Sigma_{IS}(\cdot,\cdot)$ of the measurement errors based on the data from the healthy controls in the PREDICT-HD study.

To obtain $\hat{\gamma}$, we propose to fit the following linear mixed effect model: $S_{ikl} = a_i + g_k(t_{il};\gamma)$ and estimate the parameter γ by

$$\hat{\gamma} = \operatorname{argmin} L(\gamma) = \sum_{i=1}^{n_0}\sum_{k=1}^{K}\sum_{l=1}^{L_i}(S_{ikl} - \bar{S}_{ik} - g_k(t_{il};\gamma))^2$$

where $\bar{S}_{ik} = L_i^{-1}\sum_{l=1}^{L_i} S_{ikl}$ and n_0 is the total sample size in the external data (i.e. PREDICT-HD healthy controls).

To obtain $\hat{\Sigma}_I(\cdot,\cdot)$, $\hat{\Sigma}_S(\cdot,\cdot)$, and $\hat{\Sigma}_{IS}(\cdot,\cdot)$ based on our assumed model (i.e. uncorrelated homogeneous measurement errors), we obtain the moment-based estimates by

$$\hat{\zeta}_{Ijk} = \frac{\sum_{i=0}^{n_0}\sum_{l=1}^{L_i} e_{Iijl}e_{Iikl}}{\sum_{i=0}^{n_0}(L_i-1)}, \quad \hat{\zeta}_{Sjk} = \frac{\sum_{i=0}^{n_0}\sum_{l=1}^{L_i} e_{Sijl}e_{Sikl}}{\sum_{i=0}^{n_0}(L_i-1)}$$

and

$$\hat{\zeta}_{ISjk} = \frac{\sum_{i=0}^{n_0}\sum_{l=1}^{L_i} e_{Iijl}e_{Sikl}}{\sum_{i=0}^{n_0}(L_i-1)}$$

For notational simplicity, we denote all the parameters to be estimated in this stage as $\Psi = (\gamma^\top, \zeta_I^\top, \zeta_{IS}^\top, \zeta_S^\top)^\top$ and the corresponding estimator as $\hat{\Psi}$. As shown in the proof of Lemma 2 in Appendix B of the online supplementary material, $\hat{\Psi}$ is consistent and asymptotically normally distributed.

3.2 GEE method (stage 2)

3.2.1 Initial estimator under independent working correlation matrix

Here we first discuss how to estimate regression parameters B for an arbitrarily fixed Ψ. Since all quantities within this section could be functions of Ψ, for notational simplicity, we conceal Ψ in expressions. Denote $\tilde{S}(t) = S(t) - g(t;\gamma)$ and $\tilde{S}^*(t) = S^*(t) - g(t;\gamma)$ and let $\Gamma_{il}^* = (1, t_{il}, X_i, I_{i1l}^*, \ldots, I_{iQl}^*)^\top$ and $\Gamma_{il} = (1, t_{il}, X_i, I_{i1l}, \ldots, I_{iQl})^\top$. Then Model 1 can be rewritten as

$$\tilde{S}_{il}^* = B\Gamma_{il}^* + e_{il} = (\Gamma_{il}^{*\top} \otimes I_K)\mathrm{vec}(B) + e_{il} \tag{2}$$

Let $\tilde{S}_i^* = (\tilde{S}_{i1}^{*\top}, \ldots, \tilde{S}_{iL_i}^{*\top})^\top$, then we have

$$\tilde{S}_i^* = \begin{pmatrix} \Gamma_{i1}^{*\top} \otimes I_K \\ \cdots \\ \Gamma_{iL_i}^{*\top} \otimes I_K \end{pmatrix}\mathrm{vec}(B) + e_i = (\Gamma_i^{*\top} \otimes I_K)\mathrm{vec}(B) + e_i$$

where $\Gamma_i^* = (\Gamma_{i1}, \ldots, \Gamma_{iL_i})$ and $e_i = (e_{i1}^\top, \ldots, e_{iL_i}^\top)^\top \sim N(0, V_{ei})$. In practice, for a simpler form of $\hat{B}$, we can consider a special correlation structure: $V_{ei} = \Sigma_{e.ti} \otimes \Sigma_e$, where $\Sigma_{e.ti}$ can be an autoregressive (AR) matrix with L_i time points or an exponential variance–covariance structure at time $t_{i1}, \ldots, t_{iL_i}$. This correlation structure assumes a homogeneous variance–covariance matrix up to a scalar $\Sigma_{e.t}(t, t)$, i.e. $\Sigma_e = [\Sigma_{e.t}(t, t)]^{-1}\mathrm{Cov}(e_i(t))$ for all t, a homogeneous variance–covariance process $\Sigma_{e.t}(t, s)$ between $e_{ik}(t)$ and $e_{ik}(s)$ up to a scalar $\Sigma_e(k, k)$, i.e. $[\Sigma_e(k, k)]^{-1}\mathrm{Cov}(e_{ik}(t), e_{ik}(s))$ (here $\Sigma_{e.ti}(j, k) = \Sigma_{e.t}(t_{ij}, t_{ik})$) for all $k = 1, \ldots, K$, and the correlation between two residuals $e_{ij}(t)$, $e_{ik}(s)$ is the correlation between $e_{ij}(t)$ and $e_{ij}(s)$ multiplied by the correlation between $e_{ij}(t)$ and $e_{ik}(t)$. Note that using a more complicated variance–covariance structure does not affect getting asymptotically normal estimates. However, for a limited sample size, consideration of this variance–covariance structure can significantly reduce the number of model parameters and hence make the estimation of primary model parameter B more stable.

An estimation for B can be easily implemented using the moment reconstruction approach (Carroll et al., 2006) for the GEE (Liang & Zeger, 1986) under independent working correlation. When the error-free variables $\tilde{S}_i^*$ and I_i^* are available, we have the following least-squares estimator:

$$\hat{B} = \left(\sum_{i=1}^{n}\sum_{l=1}^{L_i}\Gamma_{il}^{\star\top}\Gamma_{il}^{\star}\right)^{-1}\left(\sum_{i=1}^{n}\sum_{l=1}^{L_i}\Gamma_{il}^{\star\top}\tilde{S}_{il}^{\star}\right)$$

Based on the moment reconstruction, we can replace $\Gamma_{il}^{\star\top}\Gamma_{il}^{\star}$ by $\Gamma_{il}^\top\Gamma_{il} - \Sigma_{\Gamma l}$ and replace $\Gamma_{il}^{\star\top}S_{il}^{\star}$ by $\Gamma_{il}^\top S_{il} - \Sigma_{\Gamma Sl}$, where

$$\Sigma_{\Gamma l} = \begin{pmatrix} 0_{(2+p)\times(2+p)} & 0_{(2+p)\times Q} \\ 0_{Q\times(2+p)} & \Sigma_I(t_{il}, t_{il}) \end{pmatrix} \quad \text{and}$$

$$\Sigma_{\Gamma Sl} = \begin{pmatrix} 0_{(2+p)\times K} \\ \Sigma_{IS}(t_{il}, t_{il}) \end{pmatrix}$$

To avoid non-positivity of the matrix $\sum_{i=1}^n\sum_{l=1}^{L_i}(\Gamma_{il}^\top\Gamma_{il} - \Sigma_{\Gamma l})$, we use Fuller's adjustment (Fuller, 1987). Specifically, let λ be the minimum value such that the smallest eigenvalue of $\sum_{i=1}^n\sum_{l=1}^{L_i}(\Gamma_{il}^\top\Gamma_{il} - \lambda\Sigma_{\Gamma l})$ is 0. If $\lambda \le N/(N-1)$ where $N = \sum_{i=1}^n L_i$, we take $\hat{H} = \sum_{i=1}^n\sum_{l=1}^{L_i}(\Gamma_{il}^\top\Gamma_{il} - \Sigma_{\Gamma l})$; otherwise, we take $\hat{H} = \sum_{i=1}^n\sum_{l=1}^{L_i}(\Gamma_{il}^\top\Gamma_{il} - (\lambda - (6/(N-1)))\Sigma_{\Gamma l})$. Either way, we can express $\hat{H} = \sum_{i=1}^n\hat{H}_i$ with some $\hat{H}_i$. Let $\hat{U} = \sum_{i=1}^n\sum_{l=1}^{L_i}(\Gamma_{il}^\top S_{il} - \Sigma_{\Gamma Sl}) = \sum_{i=1}^n\hat{U}_i$, then we get $\hat{B}_{\mathrm{ind}}(\Psi) = \hat{H}^{-1}\hat{U}$. We use $\mathrm{vec}(\hat{B}_{\mathrm{ind}}) = \hat{B}_{\mathrm{ind}}(\hat{\Psi})$ as our estimator under an independent working correlation, which can be shown as a consistent and asymptotically normally distributed estimator as summarised in the following theorem.

Theorem 1 Under our model assumption with the design such that $n/n_0 \to c$ and the following regularity conditions:

(C.1) $g(t; \gamma)$ is twice differentiable with finite derivative and $\partial\sum_{l=1}^{L_i} g(t_{il}; \gamma)/\partial\gamma$ is of full rank;
(C.2) L_i is bounded above by L with probability 1;
(C.3) $E[\Gamma_{il}^{*\top}\Gamma_{il}^*]$ exists in the sense that for any $j \in \{1, \ldots, p+2\}$, $k \in \{1, \ldots, p+2\}$, $E|e_j^\top\Gamma_{il}^{*\top}\Gamma_{il}^* e_k| < \infty$ and $E[\Gamma_{il}^{*\top}\Gamma_{il}^*]$ is positive definite;

we have

$$\sqrt{n}(\mathrm{vec}(\hat{B}_{\mathrm{ind}}) - \mathrm{vec}(B)) \to N(0, V_{\mathrm{ind}})$$

where V_{ind} can be consistently estimated by a sandwich estimator $\hat{V}_{\mathrm{ind}}$ with the detailed expression given in the proof in Appendix B.1 of the online supplementary material.

Note that if we write $\hat{B}$ as $\hat{B}(\gamma, \Sigma_{\Gamma I}, \Sigma_{\Gamma SI})$, then for a misspecified $(\tilde{\gamma}, \tilde{\Sigma}_{\Gamma I}, \tilde{\Sigma}_{\Gamma SI})$, the difference in $\hat{B}$ can be bounded with high probability by

$$\|\hat{B}(\tilde{\gamma}, \tilde{\Sigma}_{\Gamma I}, \tilde{\Sigma}_{\Gamma SI}) - \hat{B}(\gamma, \Sigma_{\Gamma I}, \Sigma_{\Gamma SI})\|_\infty \leq \|\{E[\Gamma_{il}^{*\top}\Gamma_{il}^{*}]\}^{-1}\|_\infty$$
$$\times \left(\left\| \frac{\partial \Gamma_{il}^{*\top} \sum_{l=1}^{L_i} g(t_{il}; \gamma)}{\partial \gamma} \right\|_\infty \|\tilde{\gamma} - \gamma\|_1 + \|\tilde{\Sigma}_{\Gamma SI} - \Sigma_{\Gamma SI}\|_\infty + \|\tilde{\Sigma}_{\Gamma I} - \Sigma_{\Gamma I}\|_\infty \|\{E[\Gamma_{il}^{*\top}\Gamma_{il}^{*}] + \tilde{\Sigma}_{\Gamma I} - \Sigma_{\Gamma I}\}^{-1}\|_\infty \right)$$

Hence, as $n \to \infty$, the influence of these nuisance parameters on $\hat{B}$ is on a linear scale, leading to the estimation result being relatively robust to small violations of the assumption that case and control have the same parameters (i.e. $\mathbb{E}[\Gamma_{il}^{*\top}\Gamma_{il}^{*}] + \tilde{\Sigma}_{\Gamma I} - \Sigma_{\Gamma I}$ is positive definite). This implies that the inference is insensitive to small violations of the assumption.

3.2.2 *Estimator under non-independent working correlation matrix*

Although $\hat{B}_{\text{ind}}$ is robust against the misspecification of correlation structure as well as potential violations of the Markov assumption (Pepe & Anderson, 1994), modelling variance–covariance matrix correctly could potentially lead to more efficient estimation. We also provide an alternative estimation procedure under a non-independent working correlation structure indexed by the unknown parameter θ. To derive an estimator under a non-independent working correlation, we consider (M1) to estimate θ based on B and Ψ and denote it as $\hat{\theta}(B, \Psi)$; (M2) to estimate B based on the known parameter θ and Ψ and denote it as $\hat{B}(\theta, \Psi)$; and (M3) plug in an initial estimator of B and an estimator Ψ to obtain an estimate of θ using (M1) and plug in the estimates of θ and Ψ to update the estimate of B using (M2). Specifically, (M1) and (M2) can be implemented as follows:

(M1) We estimate $\Sigma_{e.ti}$ and Σ_e with given B and Ψ. Let $\hat{e}_{ij} = \tilde{S}_i^*(t_{ij}) - B\Gamma_{il}$, then we can estimate the variance–covariance process $\Sigma_e(\cdot, \cdot)$ empirically using the kernel estimator $\hat{\Sigma}_e(\cdot, \cdot)$ whose expression is given in online supplementary material, Appendix A. If we assume a parametric model $\Sigma_e(t, s; \theta)$, then we can find $\hat{\theta}(B, \Psi)$ that maximises

$$\sum_{i=1}^{n} \log\det(V_{ei}(\theta)) + tr(V_{ei}(\theta)^{-1}\hat{V}_{ei})$$

where

$$\hat{V}_{ei} = \begin{pmatrix} \hat{\Sigma}_e(t_1, t_1) & \cdots & \hat{\Sigma}_e(t_1, t_{n_i}) \\ \cdots & \cdots & \cdots \\ \hat{\Sigma}_e(t_{n_i}, t_1) & \cdots & \hat{\Sigma}_e(t_{n_i}, t_{n_i}) \end{pmatrix}$$

and

$$\Sigma_{ie.all}(\theta) = \begin{pmatrix} V_{ei}(t_1, t_1; \theta) & \cdots & V_{ei}(t_1, t_{n_i}; \theta) \\ \cdots & \cdots & \cdots \\ V_{ei}(t_{n_i}, t_1; \theta) & \cdots & V_{ei}(t_{n_i}, t_{n_i}; \theta) \end{pmatrix}.$$

(M2) We conduct the moment reconstruction since it works for the GEE-type estimation with any working correlation structure for V_{ei}. In particular, we first provide the estimating equation when $S_i^*(t)$ and $I_i^*(t)$ are available and then use the moment reconstruction approach to find the estimator with observed $S_i(t)$ and $I_i(t)$.
Based on the GEE result without measurement error, we know that the efficient estimation of vec(B) based on the general least-squares method will be using the weight V_{ei}^{-1}, and under the

special correlation structure, this can be written as

$$\text{vec}(B) = \left\{ \left(\sum_{i=1}^{n} \Gamma_i^* \Sigma_{e.ti}^{-1} \Gamma_i^{*\top} \right)^{-1} \otimes \Sigma_e \right\} \left\{ \sum_{i=1}^{n} (\Gamma_i^* \Sigma_{e.ti}^{-1} \otimes \Sigma_e^{-1}) S_i^* \right\}$$

Let w_{ijk} denote the (j, k)-element of $\Sigma_{e.ti}^{-1}$ and $\tilde{w}_{ijl}$ the (j, l)th element of $V_{ei}^{-1} = \Sigma_{e.ti}^{-1} \otimes \Sigma_e^{-1}$, $H = \sum_{i=1}^{n} \Gamma_i^* \Sigma_{e.ti}^{-1} \Gamma_i^{*\top}$ can be estimated by $\hat{H} = \sum_{i=1}^{n} \sum_{j=1}^{L_i} \sum_{k=1}^{L_i} w_{ijk} \{\Gamma_{ij} \Gamma_{ik}^\top - \Omega_I(t_{ij}, t_{ik})\}$ where $\Omega_I(t, s) = \begin{pmatrix} 0_{4\times4} & 0_{4\times Q} \\ 0_{Q\times4} & \Sigma_I(t_{ij}, t_{ik}) \end{pmatrix}$ and $U = \sum_{i=1}^{n} (\Gamma_i^* \Sigma_{e.ti}^{-1} \otimes \Sigma_e^{-1}) S_i^*$ can be estimated by $\hat{U} = \sum_{i=1}^{n} \sum_{j=1}^{L_i \times K} \sum_{l=1}^{L_i \times K} \tilde{w}_{ijl} \{(\Gamma_i \otimes I_K)_j S_{il} - \Omega_{IS}(t_{ij'}, t_{il'}) e_{l''} \otimes e_{j''})\}$ where $j' = [(j-1)/K] + 1$, $l' = [(l-1)/K] + 1$, $j'' = \text{mod}(j, K) + 1$, $l'' = \text{mod}(l, K) + 1$, and $\Omega_{IS}(t, s) = \begin{pmatrix} 0_{4\times K} \\ \Sigma_{IS}(t, s) \end{pmatrix}$. To prevent non-positivity of $\hat{H}$, we can use Fuller's adjustment (Fuller, 1987) as in the previous section. Finally, we obtain the estimator $\hat{B}(\theta, \Psi) = \hat{H}^{-1} \hat{U}$ given known $\Sigma_{e.ti}$ and Σ_e with the detailed derivation provided in online supplementary material, Appendix A.

Though (M1) and (M2) may be iteratively implemented until convergence, we propose a one-step algorithm to construct the estimator $\hat{B}$ under a non-independent working correlation matrix, which is fully described as follows:

Step 1: Estimate Ψ using external data to obtain $\hat{\Psi}$.
Step 2: Obtain the initial estimate of $\hat{B}_{\text{ind}}$ using the method given in the previous subsection, $\text{vec}(\hat{B}_{\text{ind}}) = \hat{B}(\hat{\Psi})$.
Step 3: Use $\hat{B}_{\text{ind}}$ and $\hat{\Psi}$ to obtain the estimator of θ, $\hat{\theta} = \hat{\theta}(\hat{B}_{\text{ind}}, \hat{\Psi})$ according to (M1).
Step 4: Use $\hat{\theta}$ and $\hat{\Psi}$ to obtain the final estimator of B, $\text{vec}(\hat{B}) = \hat{B}(\hat{\theta}, \hat{\Psi})$ according to (M2).

The proposed estimation procedure for a non-independent working correlation matrix does not only have a numerical advantage but also leads to a consistent and asymptotically normal estimator $\hat{B}$, which is summarised in the following theorem:

Theorem 2 Under our model assumption and regularity conditions *(C.1)–(C.3)* as listed in Theorem *1*, *(E.1)–(E.5)* as listed in Lemma 3 given in the online supplementary material, and
(C.4) $E[\sum_{j=1}^{n_i} \sum_{k=1}^{n_i} w_{ijk}(\theta) \{\Gamma_{ij} \Gamma_{ik}^\top - \Omega_I(t_{ij}, t_{ik})\}]$ exists and is finite and positive definite,
we have

$$\sqrt{n}(\text{vec}(\hat{B}) - \text{vec}(B)) \to N(0, V)$$

where V can be consistently estimated by a sandwich estimator $\hat{V}$ whose expression is given in the proof for the theorem included in Appendix B.2 of the online supplementary material.

4 Simulation

In this section, we conducted simulation studies to evaluate the performance of the proposed estimation method for regression parameters B. We considered a homogeneous setting where all the variance–covariance matrices are fixed over time. We denote $AR_d(\sigma^2, \rho)$ and $CS_d(\sigma^2, \rho)$ as the d-dimensional AR and compound symmetric variance-covariance matrix with parameters σ^2 and ρ, respectively (i.e. the (i, j)th element of $AR_d(\sigma^2, \rho)$ is $\sigma^2 \rho^{|i-j|}$ and the (i, j)th element of $CS_d(\sigma^2, \rho)$ is $\sigma^2 \rho^{I(i \neq j)}$). In the simulation studies, we set $p = 2$, $Q = 3$, and $K = 4$.

For the diseased subjects, we considered modelling the $I_i^*(t)$ process by the linear mixed effect model

$$I_i^*(t) = \beta_{0i} + \beta_{1i}t + \beta_2 X_i + \epsilon_i(t) \tag{3}$$

in addition to the regression model for the $S_i^*(t)$ process given by model (1).

For the diseased subject i, we generated the multivariate longitudinal data at three time points $t_i = (1, 2, 3)^T$ from models (3) and (1) as follows:

- First, sample the baseline covariates from $X_i \sim N(\mathbf{0}_p, \Sigma_X)$ where $\Sigma_X = CS_p(1, 0.1)$.
- Second, sample the random regression parameters in model (3) from

$$\begin{bmatrix} \beta_{0i} \\ \beta_{1i} \end{bmatrix} \sim \begin{bmatrix} 1_Q \\ (1, 2, \ldots, Q)^T \end{bmatrix} + N(\mathbf{0}_{2\times Q}, \Sigma_\beta)$$

and

$$\beta_2 = \begin{bmatrix} 0.5 & 1 & 1.5 & 2 \\ 0.5 & 1 & 1.5 & 2 \end{bmatrix}^T$$

where $1_Q = (1, 1, \ldots, 1)_{1\times Q}^T$ and $\Sigma_\beta = CS_{2\times Q}(1, -0.1)$; the regression parameters in model (1)

$$b_0 = 1_K,\ b_1 = (1, 2, \ldots, K)^T,\ b_2 = \begin{bmatrix} 0.5 & 1 & 1.5 & 2 \\ 0.5 & 1 & 1.5 & 2 \end{bmatrix}^T,\ b_3 = \begin{bmatrix} 3.2 & 2.4 & 1.6 & 0.8 \\ 3.2 & 2.4 & 1.6 & 0.8 \\ 3.2 & 2.4 & 1.6 & 0.8 \end{bmatrix}^T$$

- Third, sample the random variations of the $I_i^*(t)$ process and the $S_i^*(t)$ process from $\epsilon_i(\cdot) \sim \mathcal{MGP}_Q(\mathbf{0}, \Sigma_Q(\cdot, \cdot))$ and $e_i(\cdot) \sim \mathcal{MGP}_K(\mathbf{0}, \Sigma_K(\cdot, \cdot))$. Here, we have $\Sigma_Q(t, s) = AR_Q(3, 0.2)\Sigma_{\epsilon.ti}(t, s)$ and $\Sigma_K(t, s) = CS_K(2.5, 0.15)\Sigma_{e.ti}(t, s)$ where $\Sigma_{\epsilon.ti}(\cdot, \cdot)$ and $\Sigma_{e.ti}(\cdot, \cdot)$ quantify the correlation between different measurement times. Since we only need to generate three time points, we simply denote the variance–covariance matrix among those three time points derived from the variance–covariance process $\Sigma_{\epsilon.ti}(\cdot, \cdot)$ and $\Sigma_{e.ti}(\cdot, \cdot)$ by $\Sigma_{\epsilon.ti}$ and $\Sigma_{e.ti}$. We applied different settings with $\Sigma_{e.ti} = AR_3(1, \rho_K)$ for $\rho_K = 0.3, 0.5, 0.8$ or $CS_3(1, \rho_K)$ for $\rho_K = 0.2, 0.3, 0.5$, for the variance–covariance matrix for the $S^*(t)$ process. And we set $\Sigma_{\epsilon.t_i} = AR_3(1, \rho_Q)$ or $CS_3(1, \rho_Q)$ for $\rho_Q = 0.3$. We chose the same type of variance–covariance matrix for $\Sigma_{e.ti}$ and $\Sigma_{\epsilon.ti}$, that is, if we chose AR for $\Sigma_{e.ti}$, then AR was chosen for $\Sigma_{\epsilon.ti}$.
- Fourth, sample the measurement error and practice effects to generate $S_i(t)$ and $I_i(t)$. Let $\Sigma_{\text{err}} = \begin{pmatrix} \Sigma_I & \Sigma_{IS} \\ \Sigma_{SI} & \Sigma_S \end{pmatrix} \in R^{(Q+K)\times(Q+K)}$ be the variance–covariance matrix at a specific time for the measurement errors. We randomly sampled one $\tilde{\Sigma}_{\text{err}}$ from the Wishart distribution $W_{K+Q}(df = K + Q, \Sigma = I_{K+Q})$, where I_d is the d-dimensional identity matrix to obtain a positive definitive random matrix. This matrix

$$\tilde{\Sigma}_{\text{err}} = \begin{pmatrix} 13.55 & 1.49 & -0.70 & -5.51 & -3.35 & 4.62 & 2.18 \\ 1.49 & 2.48 & -3.21 & 0.78 & 3.17 & 0.69 & -2.43 \\ -0.70 & -3.21 & 8.18 & -1.87 & -5.20 & 0.87 & 5.58 \\ -5.51 & 0.78 & -1.87 & 5.77 & 4.07 & -2.87 & -3.31 \\ -3.35 & 3.17 & -5.20 & 4.07 & 9.80 & -3.73 & -6.80 \\ 4.62 & 0.69 & 0.87 & -2.87 & -3.73 & 4.28 & 2.75 \\ 2.18 & -2.43 & 5.58 & -3.31 & -6.80 & 2.75 & 6.41 \end{pmatrix}$$

is fixed for all simulations and we let $\Sigma_{\text{err}} = \sigma_{\text{err}}\tilde{\Sigma}_{\text{err}}$ with σ_{err} being a scalar to determine the signal-to-noise ratio (SNR). Then we have $(e_{Ii}, e_{Si}) \sim \mathcal{MGP}_Q(\mathbf{0}, \Sigma_{\text{err}})$. We selected four values

of $\sigma_{\text{err}} = 0.13, 0.34, 0.66$, and 1.33 such that the SNR $\approx 10, 4, 2, 1$. The SNR is computed as the averages of Q values at a specific time $t = 0$ from 50, 000 samples generated following the previous steps, namely,

$$\text{SNR} = \frac{1}{Q}\sum_{q=1}^{Q}\frac{\text{Var}(I^*_{iq}(t))}{\text{Var}(I_{iq}(t)) - \text{Var}(I^*_{iq}(t))}$$

We did not posit correlations between different time points for measurement errors, that is, $\Sigma_{err.ti} = I_3$. We set the practice effects $g(t; \gamma) = \gamma t$ with $\gamma = (0.1, 0.4, 0.5, 0.9)^T$.

We then generated $n = 200$, 400, and 800 diseased subjects following these steps. The healthy control subjects were generated with a sample size half of the sample size for diseased subjects using $\beta_{1i} = b_1 = \epsilon_i(t) = e_i(t) = 0$ to guarantee that $I^*(\cdot)$ and $S^*(\cdot)$ are constant over time so that we can estimate the learning effect without a subcohort where $I^*(\cdot)$ and $S^*(\cdot)$ are observed. All other settings are set the same as for diseased subjects. The moment-based estimation method for the practice effects and variance–covariance components of measurement errors was implemented using the simulated data for healthy control subjects.

Five model fitting methods were utilised in the simulation study: (1) the ORIND method: oracle (no measurement error) variables were observed and GEE with independent working correlation structure was used to obtain estimates; (2) the OLS method: we did not make measurement error corrections on measurement error data, and GEE with independent working correlation structure was used to obtain estimates; (3) the MCIND method: we made measurement error corrections for the estimating equation based on the independence working correlation structure; (4) the MCCOR method: we made measurement error corrections for the estimating equation based on the true working correlation structure; (5) the MCOS method: we make measurement error corrections with one-step estimation for working correlation.

We performed 1, 000 simulations to estimate the bias, empirical standard deviation (SD), empirical average of estimated standard errors (ESEs) based on the bootstrap method with 100 resamplings of each simulated data, and coverage rate (CR) for 95% confidence interval (CI). We evaluated the performance of our estimators by comparing the different settings mentioned above. The full simulation results are shown in Table 1 and Tables 1–23 in Appendix C of the online supplementary material.

From the results in these tables, we can see that all methods except OLS show little to no bias when the sample size is large ($n = 800$). The OLS method has a bias that does not go away when the sample size increases. Such bias in the OLS method tends to increase with the increase in the scale of measurement error, or, equivalently, the decrease in SNR. When the sample size is small ($n = 200$), MCCOR and MCOS also show substantial bias when the measurement error is strong (i.e. SNR = 1). This is mainly because large measurement errors tend to make the parameter estimation procedure unstable with a larger SD and more outliers in general. MCIND shows a much smaller bias compared to MCCOR and MCOS when the sample size is small. So we recommend that if the measurement error is expected to be large (SNR < 3) in a dataset, a sufficiently large sample size is imperative to yield a more precise estimation, and using MCIND is preferred. All methods except OLS have an appropriate correct CR for their nominal 95% confidence interval, and the CRs are neither affected by the strength of measurement errors nor the sample size.

When we compare the SD and ESE in different settings, we see that the SD for methods using the error-prone data was larger than the SD for the oracle method using data without measurement error, which implies that measurement errors cause a loss of efficiency. When we compare SD and ESE for different measurement error correction estimators (MCIND, MCCOR, and MCOS), we conclude that using various approaches for the correlation structure does substantially change the SD and ESE, and in most settings, MCIND is the most efficient method among these three methods. We conjecture this is due to the fact that measurement error correction methods with non-independent working correlation require estimation of the covariance matrix between measurement errors at different time points, which no doubt introduces additional variability compared to the estimator with independent working correlation. When SNR is small

Table 1. Simulation results on the first outcome in $S^*(t)$: comparison of estimation bias, empirical standard deviation (SD), average estimated standard error (ESE), and coverage rate (CR) for 95% confidence interval with autoregressive (AR) correlation and $\rho_e = 0.5$ under different sample sizes (N) and signal-to-noise ratio (SNR) among the different models: ORIND: oracle (no measurement error) with independent working correlation; OLS: no measurement error correction with independent working correlation; MCIND: measurement error correction with independent working correlation; MCCOR: measurement error correction with true correlation; MCOS: measurement error correction with one-step estimation for correlation

SNR	Method	n	The first component in $I^*(t)$				The second component in $I^*(t)$				The third component in $I^*(t)$			
			Bias	SD	ESE	CR	Bias	SD	ESE	CR	Bias	SD	ESE	CR
10	ORIND	200	0.002	0.029	0.029	0.937	0	0.028	0.029	0.957	0	0.027	0.029	0.945
10	ORIND	400	0.001	0.02	0.02	0.947	0.001	0.02	0.02	0.949	0	0.021	0.02	0.939
10	ORIND	800	0	0.015	0.011	0.943	0	0.014	0.011	0.947	0.001	0.014	0.011	0.939
10	OLS	200	−0.072	0.057	0.056	0.742	−0.152	0.057	0.056	0.22	−0.255	0.054	0.053	0.002
10	OLS	400	−0.072	0.039	0.04	0.552	−0.152	0.04	0.04	0.036	−0.257	0.039	0.038	0
10	OLS	800	−0.072	0.03	0.029	0.268	−0.153	0.028	0.03	0	−0.253	0.027	0.028	0
10	MCIND	200	0.002	0.067	0.066	0.944	0.003	0.065	0.065	0.945	0.005	0.078	0.08	0.965
10	MCIND	400	0.001	0.046	0.046	0.954	0.003	0.045	0.046	0.946	0.002	0.056	0.055	0.953
10	MCIND	800	0	0.034	0.031	0.932	0.001	0.033	0.031	0.944	0.004	0.039	0.039	0.952
10	MCCOR	200	0.001	0.085	0.086	0.95	0.004	0.081	0.083	0.957	0.009	0.114	0.117	0.957
10	MCCOR	400	0.001	0.061	0.06	0.947	0.005	0.058	0.058	0.944	0.002	0.08	0.081	0.948
10	MCCOR	800	0	0.043	0.041	0.937	0.001	0.04	0.04	0.958	0.004	0.056	0.057	0.959
10	MCOS	200	0.001	0.074	0.073	0.943	0.003	0.071	0.072	0.947	0.007	0.095	0.095	0.96
10	MCOS	400	0	0.053	0.052	0.951	0.005	0.051	0.05	0.945	0.002	0.069	0.067	0.946
10	MCOS	800	0	0.039	0.038	0.933	0.001	0.036	0.036	0.958	0.004	0.048	0.049	0.954
4	ORIND	200	0.002	0.029	0.029	0.937	0	0.028	0.029	0.957	0	0.027	0.029	0.945
4	ORIND	400	0.001	0.02	0.02	0.947	0.001	0.02	0.02	0.949	0	0.021	0.02	0.939
4	ORIND	800	0	0.015	0.011	0.943	0	0.014	0.011	0.947	0.001	0.014	0.011	0.939
4	OLS	200	−0.208	0.082	0.079	0.253	−0.38	0.078	0.079	0.005	−0.561	0.07	0.07	0
4	OLS	400	−0.208	0.055	0.055	0.035	−0.378	0.056	0.056	0	−0.564	0.052	0.05	0
4	OLS	800	−0.207	0.042	0.04	0	−0.379	0.039	0.04	0	−0.56	0.036	0.036	0
4	MCIND	200	0.004	0.122	0.124	0.958	0.006	0.114	0.119	0.969	0.016	0.162	0.175	0.968

(*continued*)

Table 1. Continued

SNR	Method	n	The first component in $I^*(t)$				The second component in $I^*(t)$				The third component in $I^*(t)$			
			Bias	SD	ESE	CR	Bias	SD	ESE	CR	Bias	SD	ESE	CR
4	MCIND	400	0.003	0.083	0.083	0.954	0.009	0.083	0.082	0.956	0.005	0.116	0.117	0.954
4	MCIND	800	0.001	0.06	0.058	0.938	0.001	0.057	0.057	0.945	0.007	0.079	0.08	0.952
4	MCCOR	200	0.005	0.175	0.217	0.974	0.01	0.164	0.187	0.966	0.033	0.275	0.358	0.972
4	MCCOR	400	0.004	0.123	0.126	0.959	0.015	0.12	0.118	0.948	0.008	0.192	0.199	0.952
4	MCCOR	800	0.002	0.087	0.085	0.947	0.003	0.079	0.08	0.945	0.009	0.13	0.134	0.952
4	MCOS	200	0.005	0.139	0.147	0.965	0.007	0.131	0.138	0.964	0.021	0.204	0.223	0.972
4	MCOS	400	0.003	0.099	0.098	0.957	0.012	0.098	0.094	0.951	0.006	0.148	0.148	0.954
4	MCOS	800	0.001	0.072	0.069	0.939	0.002	0.067	0.066	0.94	0.008	0.103	0.104	0.948
2	ORIND	200	0.002	0.029	0.029	0.937	0	0.028	0.029	0.957	0	0.027	0.029	0.945
2	ORIND	400	0.001	0.02	0.02	0.947	0.001	0.02	0.02	0.949	0	0.021	0.02	0.939
2	ORIND	800	0	0.015	0.011	0.943	0	0.014	0.011	0.947	0.001	0.014	0.011	0.939
2	OLS	200	−0.412	0.101	0.1	0.017	−0.666	0.094	0.099	0	−0.896	0.081	0.083	0
2	OLS	400	−0.41	0.071	0.07	0.001	−0.669	0.068	0.07	0	−0.899	0.06	0.059	0
2	OLS	800	−0.412	0.05	0.05	0	−0.666	0.049	0.05	0	−0.895	0.042	0.041	0
2	MCIND	200	0.014	0.203	0.525	0.984	0.029	0.19	0.431	0.978	0.042	0.317	0.909	0.967
2	MCIND	400	0.004	0.14	0.15	0.966	0.012	0.13	0.138	0.96	0.026	0.225	0.237	0.961
2	MCIND	800	−0.008	0.094	0.096	0.952	0.003	0.093	0.093	0.961	0.024	0.144	0.148	0.965
2	MCCOR	200	−0.007	0.645	1.447	0.996	0.042	0.397	0.897	0.987	0.151	1.239	2.819	0.978
2	MCCOR	400	−0.001	0.237	0.559	0.987	0.018	0.209	0.36	0.981	0.065	0.425	1.088	0.97
2	MCCOR	800	−0.016	0.147	0.165	0.97	0.003	0.139	0.147	0.966	0.05	0.257	0.283	0.965
2	MCOS	200	0.016	0.252	0.611	0.99	0.035	0.227	0.403	0.981	0.059	0.414	1.114	0.976
2	MCOS	400	0.002	0.17	0.221	0.976	0.013	0.155	0.182	0.968	0.037	0.287	0.381	0.966
2	MCOS	800	−0.012	0.112	0.118	0.958	0.002	0.11	0.111	0.962	0.032	0.188	0.194	0.962
1	ORIND	200	0.002	0.029	0.029	0.937	0	0.028	0.029	0.957	0	0.027	0.029	0.945

(*continued*)

Table 1. Continued

SNR	Method	n	The first component in $I^*(t)$				The second component in $I^*(t)$				The third component in $I^*(t)$			
			Bias	SD	ESE	CR	Bias	SD	ESE	CR	Bias	SD	ESE	CR
1	ORIND	400	0.001	0.02	0.02	0.947	0.001	0.02	0.02	0.949	0	0.021	0.02	0.939
1	ORIND	800	0	0.015	0.011	0.943	0	0.014	0.011	0.947	0.001	0.014	0.011	0.939
1	OLS	200	−0.774	0.119	0.123	0	−1.097	0.118	0.119	0	−1.363	0.094	0.093	0
1	OLS	400	−0.772	0.089	0.086	0	−1.098	0.084	0.084	0	−1.365	0.066	0.066	0
1	OLS	800	−0.77	0.064	0.061	0	−1.096	0.059	0.06	0	−1.361	0.047	0.047	0
1	MCIND	200	0.043	1.962	16.919	0.998	0.106	1.128	11.192	0.993	0.112	4.092	33.772	0.989
1	MCIND	400	−0.002	0.27	1.748	0.99	0.026	0.248	1.039	0.984	0.079	0.488	3.407	0.979
1	MCIND	800	0.001	0.177	0.205	0.97	0.019	0.177	0.184	0.949	0.048	0.291	0.353	0.961
1	MCCOR	200	−0.383	3.375	4.551	0.989	0.246	2.216	4.245	0.995	1.653	6.567	9.17	0.989
1	MCCOR	400	−0.137	1.703	3.697	0.996	0.023	1.074	2.42	0.993	0.504	3.651	7.697	0.99
1	MCCOR	800	−0.036	0.952	1.969	0.991	0.021	0.629	1.07	0.987	0.174	1.589	3.954	0.984
1	MCOS	200	−0.182	2.07	2.9	0.996	0.107	1.282	2.221	0.993	0.737	4.444	5.977	0.985
1	MCOS	400	−0.052	0.92	1.749	0.995	0.007	0.586	1.041	0.985	0.201	1.777	3.574	0.982
1	MCOS	800	−0.002	0.224	0.566	0.982	0.023	0.208	0.365	0.973	0.07	0.397	1.084	0.971

(SNR ≤ 2), there is some inconsistency between SD and ESE for measurement error correction methods (MCIND, MCCOR, MCOS), and the discrepancy is larger when the SNR is smaller. Again, the main reason for this phenomenon is that large measurement errors cause a loss of efficiency and yield unstable estimators when the sample size is small. But when the sample size increases, the discrepancy between SD and ESE becomes negligible. When the sample size is moderate ($n = 400$), the SD and ESE are closer to each other when using MCIND compared with MCCOR or MCOS. When we compare the results based on various $\Sigma_{e.ti}$ values, all the previous simulation summaries are still valid. All these results suggest that the measurement error correction method with independent working correlation is an efficient approach compared with one-step estimation or even the method using true correlation in the multivariate longitudinal data setting. Given its numerical advantage and well-behaved inferences, MCIND earns our recommendation for real-world applications in analysing error-prone multivariate longitudinal data. Therefore, for the future analysis of model estimation procedures in a similar scenario, we would recommend researchers use independent working correlation.

The sensitivity of performance to the assumption of the same variance–covariance matrix of the measurement error and learning effects between the two cohorts was evaluated by allowing a 10% difference between the parameters from the two cohorts. The results are provided in Appendices C.7 and C.8 of the online supplementary material. We can see from online supplementary material, Tables in Appendix C.7, that when the variance–covariance matrices of the measurement errors between the two cohorts are different, there will be some bias in the estimator, and the CR will be reduced to below 95%, but the magnitude of the bias is small compared to the bias of the OLS method. Since we assume a multiplicative difference between the two matrices, when the SNR is large (10 and 4), which is in the range of our real data, the result is very robust. We can see from online supplementary material, Tables in Appendix C.8, that when the learning effects are different between the two cohorts, the bias is small for our parameters of interest and the coverage is still around 95%. This might be due to the fact that the learning effect is a function of t and the misspecified learning effect contributed mostly to the bias for the estimation of time effects rather than the estimation of imaging domain effects.

The method is implemented using R software, and the code associated with the data generation and parameter estimation for the simulation settings is available at https://github.com/litong8/associationHD.

5 PREDICT-HD analysis

In this section, we applied the proposed method to the PREDICT-HD data, aiming to ascertain how brain atrophy impacts cognitive functions in prodromal-HD subjects. For each study participant, the brain structure MRI data were collected roughly every two years, up to three times. Since Aylward et al. (2011) and Tabrizi et al. (2012) have shown that the brain atrophy manifested in striatum and white-matter regions in prodromal-HD subjects and the fronto striatal atrophy was found to be associated with cognitive dysfunction (Dominguez et al., 2017; Wilkes et al., 2019), we selected three regions in the striatum: caudate, putamen, and globus ($Q = 3$) for $I(t)$ in this analysis to explore their possible association with cognitive impairments. The PREDICT-HD neuropsychological battery contained 18 measures from standardised clinical tests, from which Harrington et al. (2012) developed 6 HD-specific cognitive domains using factor analysis. They are: *speed-inhibition, working memory, motor planning & speed, attention-information integration, sensory-perceptual*, and *verbal learning & memory*. The impairment occurred in all cognitive domains as a part of disease progression in prodromal-HD subjects (Zhang et al., 2021). In this analysis, we chose the most collected test outcome in each domain to represent the cognitive performance in the domain and constituted $K = 6$ cognitive measures for $S(t)$: stroop colour word test (STROOP) for *speed-inhibition*, WAIS-III letter number sequencing (LNS) for *working memory*, paced tapping (PACE) for *motor planning & speed*, symbol digit modalities test (SDMT) for *attention-information integration*, smell identification test (SMELL) for *sensory-perceptual*, and the Hopkins verbal learning test (HVLT) free recall for *verbal learning & memory*. The higher the values on these measures except PACE, the better the cognitive function. PACE measures movement time, hence indicating better cognition in motor planning and speed with a smaller value. We also included information on education (in years), gender, and CAG-age product (CAP)

score (Zhang et al., 2011) at the study entry as the covariates X ($p = 3$). We restricted our analytical sample to those with three repeated measures for structure MRI data, resulting in $N_1 = 325$ prodromal-HD subjects and $N_0 = 111$ healthy controls. The last value carry-forward method was used to impute the missing values to align the measurements from the cognitive domains with those from the structure MRI data. The observed structure MRI data in the caudate, putamen, and globus regions and test scores in the selected 6 cognitive domains for the 111 healthy controls were used to estimate the practice effects and the variance–covariance matrix of the measurement errors in stage 1.

In the primary stage 2 analysis, we analysed the relationships between the three striatal regions ($I^*(t) \in R^3$) and the six cognitive domains ($S^*(t)$) using model (1). The estimation results for $b_0 \in R^6$, $b_1 \in R^6$, $b_2 \in R^{6\times3}$, $b_3 \in R^{6\times3}$ were summarised in Table 2.

The analysis showed that for prodromal-HD subjects, (1) cognition declines over time in all domains, particularly in the domains of *speed-inhibition* ($p < 0.001$) and *motor planning & speed* ($p < 0.001$), which makes a perfect sense as motor impairment is the leading dysfunction for HD; (2) individuals with higher education tend to perform significantly better cognitively almost in all the domains except the domain of *sensory-perceptual*; (3) males appear to be more cognitively impaired in the domains of *sensory-perceptual* ($p < 0.001$) and *verbal learning & memory* ($p = 0.001$) but less impaired in the domain of *motor planning & speed* ($p < 0.001$) than females; (4) as the disease burden index (Zhang et al., 2011), the higher CAP value implies more disease progression in four cognitive domains: *speed-inhibition* ($p < 0.001$), *motor planning & speed* ($p = 0.006$), *attention-information integration* ($p = 0.006$), and *sensory-perceptual* ($p = 0.004$). Adjusted for these risk factors, this analysis found that the striatal volumes in the regions of the caudate and putamen are associated with cognitive functions. In particular, a higher caudate volume is associated with better memory function in both *working memory* ($p < 0.001$) and *verbal learning & memory* ($p < 0.001$) and with better *attention-information integration* ($p = 0.02$); a higher putamen volume may be associated with better *sensory-perceptual* ($p = 0.03$). We also conducted the multivariate linear regression without considering the measurement errors, and the results showed that these associations are also statistically significant and are in the same direction as mentioned above. However, the magnitudes of the estimated associations between caudate and *working memory* ($p < 0.001$), *verbal learning & memory*, and *attention-information integration* are reduced by 1%, 0.2%, and 14%, respectively, and the estimated association between the putamen and *sensory-perceptual* reduced by 24%. These findings are new to the HD research community, and they expand the knowledge of HD progression with respect to the phenotype of cognitive impairments that can be explained by striatal region atrophy and validate the use of striatum volume loss as a useful biomarker for potential intervention trials to treat prodromal-HD subjects.

We looked at the nonlinear time effect by determining whether the quadratic effects of time are significant in order to ascertain whether cognitive impairment over time might be nonlinear, as the reviewers suggested. The results show that the quadratic effects of time are not significant for all the outcomes, and the estimates of the regression coefficients associated with the imaging measures showed almost no change with the addition of such quadratic terms of time.

One important assumption we made in this analysis is that the measurement error distribution and learning effects are the same between the HD subjects and healthy controls. We believe that this assumption is reasonable because the subjects in the HD cohort are still in the prodromal stage of the disease; therefore, we do not expect a large difference from the healthy controls regarding measurement error mechanisms and learning effects in the cognitive tests. We also empirically estimated the measurement error distribution and learning effects for the HD subjects using the same approach as for the healthy controls. For the measurement error distribution, we obtained comparable measurement errors except for the caudate, putamen, and SDMT, where the estimated variance components of the measurement errors are 1.4–3.3 times higher. The individual heterogeneity in the longitudinal impairment for the HD cohort, while the estimation was under the assumption of homogeneity regarding the impairment in the study cohort, may help to explain the relatively large discrepancy in estimating these quantities. The overestimation of variance components of the measurement errors in the disease cohort cannot be removed without additional parametric assumptions on the variance–covariance matrix of the measurement errors for identifiability. Alternative untestable parametric assumptions for the individual longitudinal impairment over time could be used to separate the variation due to measurement error from the heterogeneity in individual longitudinal impairment.

Table 2. Analysis results for PREDICT-HD data using measurement error correction with independent working correlation (MCIND)

Variable	Outcome	Estimate	L95% CI	U95% CI	*p* value
Intercept	STROOP	4.77	3.35	6.20	< 0.0001
Intercept	LNS	1.30	0.51	2.09	0.001
Intercept	PACE	11.92	9.56	14.28	< 0.0001
Intercept	SDMT	2.53	0.45	4.61	0.017
Intercept	SMELL	9.66	8.06	11.25	< 0.0001
Intercept	HVLT	2.74	1.58	3.89	< 0.0001
Time	STROOP	−0.12	−0.17	−0.07	< 0.0001
Time	LNS	0.04	−0.03	0.11	0.28
Time	PACE	0.17	0.77	0.26	0.0003
Time	SDMT	−0.03	−0.12	0.06	0.50
Time	SMELL	−0.05	−0.14	0.04	0.29
Time	HVLT	−0.01	−0.09	0.08	0.90
Education (years)	STROOP	0.10	0.06	0.14	< 0.0001
Education (years)	LNS	0.07	0.03	0.10	0.0002
Education (years)	PACE	−0.08	−0.15	−0.01	0.02
Education (years)	SDMT	0.11	0.05	0.16	< 0.0001
Education (years)	SMELL	0.01	−0.05	0.06	0.80
Education (years)	HVLT	0.10	0.05	0.1	< 0.0001
Male	STROOP	−0.13	−0.38	0.13	0.32
Male	LNS	0.11	−0.04	0.27	0.15
Male	PACE	−0.64	−0.99	−0.30	0.0002
Male	SDMT	−0.30	−0.58	−0.02	0.04
Male	SMELL	−0.61	−0.91	−0.31	< 0.0001
Male	HVLT	−0.33	−0.53	−0.13	0.001
CAP × 100	STROOP	−0.29	−0.46	−0.13	0.0005
CAP × 100	LNS	−0.02	−0.10	0.05	0.55
CAP × 100	PACE	0.51	0.15	0.87	0.006
CAP × 100	SDMT	−0.29	−0.50	−0.08	0.006
CAP × 100	SMELL	−0.32	−0.53	−0.10	0.004
CAP × 100	HVLT	−0.09	−0.23	0.05	0.23
Caudate	STROOP	0.06	−0.09	0.21	0.41
Caudate	LNS	0.22	0.11	0.33	0.0001
Caudate	PACE	−0.09	−0.35	0.17	0.50
Caudate	SDMT	0.24	0.33	0.47	0.02
Caudate	SMELL	−0.10	−0.30	0.09	0.31
Caudate	HVLT	0.36	0.21	0.52	< 0.0001
Putamen	STROOP	0.07	−0.15	0.29	0.53
Putamen	LNS	0.07	−0.12	0.26	0.46
Putamen	PACE	−0.18	−0.54	0.17	0.31
Putamen	SDMT	0.18	−0.11	0.46	0.23
Putamen	SMELL	0.40	0.04	0.76	0.03
Putamen	HVLT	0.02	−0.19	0.23	0.86

(*continued*)

Table 2. Continued

Variable	Outcome	Estimate	L95% CI	U95% CI	*p* value
Globus	STROOP	0.05	−0.20	0.30	0.67
Globus	LNS	−0.09	−0.33	0.15	0.47
Globus	PACE	−0.30	−0.71	0.11	0.16
Globus	SDMT	0.07	−0.31	0.44	0.73
Globus	SMELL	−0.06	−0.42	0.31	0.77
Globus	HVLT	−0.10	−0.35	0.14	0.41

Note. PREDICT-HD = Neurobiological Predictors of Huntingtons Disease Study; CI = confidence interval; STROOP = stroop colour word test; LNS = letter number sequencing; PACE = paced tapping; SDMT = symbol digit modalities test; SMELL = smell identification test; HVLT = Hopkins verbal learning test.

Repeated measurements within a short period of time for the disease cohort would be ideal for future designs to allow for weaker assumptions. For the learning effects, the empirical results also showed that they were also statistically significantly different between the two cohorts. However, the learning effects are possibly confounded with the disease effect on cognitive impairment; making the learning effects the same between the HD cohort and controls will somehow separate the disease effect from the learning effects. The different estimated learning effects between the disease cohort and health controls, as observed in the analysis, are more likely to reflect the disease effect on the cognitive impairment than different learning effects. Even if there are small differences in the learning effects, the analytical results will not be substantially changed based on the theoretical note given in the method section.

In addition, we conducted a series of sensitivity analyses to assess the robustness of our findings. First, we ran an analysis restricted to the white population only, and the results can be found in the online supplementary material (Appendix D.1). Second, we ran an analysis using only the four significant cognitive domains identified from the primary analysis, and the results can be found in the online supplementary material (Appendix D.2). Third, we performed multiple imputations by chained equation (van Buuren & Groothuis-Oudshoorn, 2011) and the results can be found in the online supplementary material (Appendix D.3). From the results, we can see that except for the association between the caudate and SDMT, which becomes non-significant at level 0.05 when multiple imputations were used, all other detected associations between the striatal region volumes and cognitive functions in the primary analysis remained statistically significant at level 0.05, and the sensitivity analyses resulted in the same direction and similar magnitude for the associations, which suggested the robustness of our findings. Lastly, we performed sensitivity analyses, expanding the study sample to those with at least two repeated measurements, and the results can be found in the online supplementary material (Appendix D.4). The results confirmed that the significant associations we identified between brain volume in the striatal region and cognitive impairment are robust.

6 Conclusion and discussion

In summary, we made the following contributions to the field: (1) we proposed a sophisticated measurement error correction method that can handle correlated measurement errors in multiple exposures and multiple outcomes that are both measured longitudinally; (2) we derived the asymptotic results for our proposed estimator from the two-stage fitting approach and derive the bias bound for such an estimator when the estimation of the variance–covariance matrix of measurement error is biased in the first stage; (3) we illustrated through simulation that the proposed methods perform well, and the method with an independent working correlation outperforms other alternatives, which is thus recommended in general; (4) we applied our method to the PREDICT-HD dataset and found interesting associations between brain volume in the striatal region and cognitive measures in specific domains for the prodromal-HD subjects, which are new discoveries and will enhance the knowledge of HD disease progression in the HD research community.

In this work, we conducted the association analysis using a marginal approach to the raw measurements. Technically, we can also carry out the analysis in an alternative manner. For example, we may use $I^*(t)$ and $I(t)$ to denote the measurement from the MRI without and with error, and we are interested in directly exploring the effect of brain atrophy on cognitive measures adjusting for the baseline cognitive measures, which gives rise to the following model:

$$S_i^*(t) = b_0 + b_1 t + b_2 X_i + b_3 (I_i^*(t) - I_i^*(0)) + b_4 S_i^*(0) + e_i(t) \tag{4}$$

If we use the transformation $\tilde{I}^*(\cdot) = (I_i^*(\cdot) - I_i^*(0), S_i^*(0))$ and $\tilde{b}_3 = (b_3, b_4)$, then we can rewrite this model as

$$S_i^*(t) = b_0 + b_1 t + b_2 X_i + \tilde{b}_3 \tilde{I}^*(t) + e_i(t)$$

while

$$\Sigma_{\tilde{I}}(t, s) = \begin{pmatrix} \Sigma_I(t, s) - \Sigma_I(t, 0) - \Sigma_I(0, s) + \Sigma_I(0, 0) & \Sigma_{IS}(t, 0) - \Sigma_{IS}(0, 0) \\ \Sigma_{SI}(0, s) - \Sigma_{SI}(0, 0) & \Sigma_S(0, 0) \end{pmatrix}$$

$$\Sigma_{\tilde{I}S}(t, s) = \begin{pmatrix} \Sigma_{IS}(t, s) - \Sigma_{IS}(0, s) \\ \Sigma_S(0, s) \end{pmatrix}$$

Then we can apply the proposed model using this new estimated variance–covariance matrix for the error term. However, due to the limited number of longitudinal observations, this approach will lead to one fewer observation point, resulting in unstable inferences (large variation) in our PREDICT-HD analysis.

The requirement of the existence of external data to help estimate nuisance quantities such as variance–covariance components in measurement errors and learning effects limits the generalizability of the proposed methods to more general longitudinal data. Our proposed method is specifically tailored to the study design in neurodegenerative diseases, like the PREDICT-HD study that has collected parallel information from healthy controls as the external data, and we do not intend to propose a method that is generalised to analyse any multivariate longitudinal data with measurement errors. We do not think there is a general template for analysing multivariate longitudinal data with measurement errors, and the method needs to take into account the features of the study designs and use corresponding approaches to identify those nuisance parameters (for example, using repeated measures within a short period for the disease group). Another limitation of our analysis is that we assume the missing at random for missing data. This assumption, while commonly adopted in practice, cannot be directly tested based on the observed data unless certain parametric models are proposed for the missing mechanism under the not missing at random setting, which is worth future exploration but is beyond the scope of this paper.

Acknowledgments

This research is partly supported by U54GM115458 from NIGMS and R01NS103475 from NIH-NINDS.

Conflict of interests: None declared.

Funding statement

NIH, NIGMS, Grant/Award Number: U54GM115458; NIH, NINDS, Grant/Award Number: R01NS103475.

Data availability statement

Publicly available datasets were analysed in this study. These data can be found at National Center for Biotechnology Information (NCBI) dbGaP, https://www.ncbi.nlm.nih.gov/gap/, PREDICT-HD Huntington Disease Study (phs000222.v6.p2).

Supplementary material

Supplementary material is available online at *Journal of the Royal Statistical Society: Series C* (http://mtp.oxfordjournals.org/).

References

Ahveninen L. M., Stout J. C., Georgiou-Karistianis N., Lorenzetti V., & Glikmann-Johnston Y. (2018). Reduced amygdala volumes are related to motor and cognitive signs in Huntington's disease: The IMAGE-HD study. *NeuroImage: Clinical*, *18*, 881–887. https://doi.org/10.1016/j.nicl.2018.03.027

Aylward E. H., Nopoulos P. C., Ross C. A., Langbehn D. R., Pierson R. K., Mills J. A., Johnson H. J., Magnotta V. A., Juhl A. R., Paulsen J. S., & PREDICT-HD Investigators (2011). Longitudinal change in regional brain volumes in prodromal Huntington disease. *Journal of Neurology, Neurosurgery & Psychiatry*, *82*(4), 405–410. https://doi.org/10.1136/jnnp.2010.208264

Buonaccorsi J., Demidenko E., & Tosteson T. (2000). Estimation in longitudinal random effects models with measurement error. *Statistica Sinica*, *10*(3), 885–903.

Carroll R., Ruppert D., Stefanski L., & Crainiceanu C. (2006). *Measurement error in nonlinear models: A modern perspective*. CRC Press.

Cole S., Chu H., & Greenland S. (2006). Multiple-imputation for measurement-error correction. *International Journal of Epidemiology*, *35*(4), 1074–1081. https://doi.org/10.1093/ije/dyl097

Cook J., & Stefanski L. (1994). Simulation-extrapolation estimation in parametric measurement error models. *Journal of the American Statistical Association*, *89*(428), 1314–1328. https://doi.org/10.1080/01621459.1994.10476871

Dominguez J., Poudel G., Stout J. C., Gray M., Chua P., Borowsky B., Egan G. F., & Georgiou-Karistianis N. (2017). Longitudinal changes in the fronto-striatal network are associated with executive dysfunction and behavioral dysregulation in Huntington's disease: 30 months IMAGE-HD data. *Cortex*, *92*, 139–149. https://doi.org/10.1016/j.cortex.2017.04.001

Duff K., Paulsen J., Mills J., Beglinger L. J., Moser D. J., Smith M. M., Langbehn D., Stout J., Queller S., & Harrington D. L. (2010). Mild cognitive impairment in prediagnosed Huntington disease. *Neurology*, *75*(6), 500–507. https://doi.org/10.1212/WNL.0b013e3181eccfa2

Falleti M. G., Maruff P., Collie A., & Darby D. G. (2006). Practice effects associated with repeated assessment of cognitive function using the CogState battery at 10-minute, one week and one month test-retest intervals. *Journal of Clinical and Experimental Neuropsychology*, *28*(7), 1095–1112. https://doi.org/10.1080/13803390500205718

Fraser G., & Stram D. (2001). Regression calibration in studies with correlated variables measured with error. *Biometrika*, *154*(9), 836–844. https://doi.org/10.1093/aje/154.9.836

Freedman L. S., Fainberg V., Kipnis V., Midthune D., & Carroll R. (2004). A new method for dealing with measurement error in explanatory variables of regression models. *Biometrics*, *60*(1), 172–181. https://doi.org/10.1111/j.0006-341X.2004.00164.x

Fuller W. (1987). *Measurement error models*. Wiley.

Harrington D. L., Smith M. M., Zhang Y., Carlozzi N. E., Paulsen J. S., & PREDICT-HD Investigators of the Huntington Study Group (2012). Cognitive domains that predict time to diagnosis in prodromal Huntington disease. *Journal of Neurology, Neurosurgery & Psychiatry*, *83*(6), 612–619. https://doi.org/10.1136/jnnp-2011-301732

Liang K., & Zeger S. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, *73*(1), 13–22. https://doi.org/10.1093/biomet/73.1.13

Lin H., Qin G., Zhang J., & Zhu Z. (2018). Analysis of longitudinal data with covariate measurement error and missing responses: An improved unbiased estimating equation. *Computational Statistics and Data Analysis*, *121*, 104–112. https://doi.org/10.1016/j.csda.2017.11.010

Nakamura T. (1990). Corrected score function for errors in variables models: Methodology and application to generalized linear models. *Biometrika*, *77*(1), 127–137. https://doi.org/10.1093/biomet/77.1.127

Oh E., Shepherd B., Lumley T., & Shaw P. (2021). Raking and regression calibration: Methods to address bias from correlated covariate and time-to-event error. *Statistics in Medicine*, *40*(3), 631–649. https://doi.org/10.1002/sim.v40.3

Paulsen J. S. (2011). Cognitive impairment in Huntington disease: Diagnosis and treatment. *Current Neurology and Neuroscience Reports*, *11*(5), 474–483. https://doi.org/10.1007/s11910-011-0215-x

Paulsen J. S., Langbehn D. R., Stout J. C., Aylward E., Ross C. A., Nance M., Guttman M., Johnson S., MacDonald M., Beglinger L. J., & Duff K. (2008). Detection of Huntington's disease decades before diagnosis: Predict-HD study. *Journal of Neurology, Neurosurgery & Psychiatry*, *79*(8), 874–880. https://doi.org/10.1136/jnnp.2007.128728

Paulsen J. S., Long J. D., Ross C. A., Harrington D. L., Erwin C. J., Williams J. K., Westervelt H. J., Johnson H. J., Aylward E. H., Zhang Y., & Bockholt H. J. (2014). Prediction of manifest Huntington's disease with clinical

and imaging measures: A prospective observation study. *The Lancet Neurology*, *13*(12), 1193–1201. https://doi.org/10.1016/S1474-4422(14)70238-8

Paulsen J. S., Smith M., & Long J. (2013). Cognitive decline in prodromal Huntington disease: Implications and clinical trials. *Journal of Neurology, Neurosurgery & Psychiatry*, *84*(11), 1233–1239. https://doi.org/10.1136/jnnp-2013-305114

Pepe M., & Anderson G. (1994). A cautionary note on inference for marginal regression models with longitudinal data and general correlated response data. *Communications in Statistics-Simulation and Computation*, *23*(4), 939–951. https://doi.org/10.1080/03610919408813210

Pereira D., Costa P., & Cerquerira J. (2015). Repeated assessment and practice effects of the written symbol digit modalities test using a short-inter test interval. *Archives of Clinical Neuropsychology*, *30*(5), 424–434. https://doi.org/10.1093/arclin/acv028

Ross C. A., Aylward E. H., Wild E. J., Langbehn D. R., Long J. D., Warner J. H., Scahill R. I., Leavitt B. R., Stout J. C., Paulsen J. S., & Reilmann R. (2014). Huntington disease natural history, biomarkers and prospects for therapeutics. *Nature Reviews Neurology*, *10*(4), 204–216. https://doi.org/10.1038/nrneurol.2014.24

Shaw P., He J., & Shepherd B. (2021). Regression calibration to correct correlated errors in outcome and exposure. *Statistics in Medicine*, *40*(2), 271–286. https://doi.org/10.1002/sim.v40.2

Stefanski L., & Carroll R. (1987). Conditional scores and optimal scores for generalized linear measurement-error models. *Biometrika*, *74*(4), 703–716. https://doi.org/10.1093/biomet/74.4.703

Stout J. C., Jones R., Labuschagne I., O'Regan A. M., Say M. J., Dumas E. M., Queller S., Justo D., Santos R. D., Coleman A., & Hart E. P. (2012). Evaluation of longitudinal 12 and 24 month cognitive outcomes in premanifest and early Huntington's disease. *Journal of Neurology, Neurosurgery & Psychiatry*, *83*(7), 687–694. https://doi.org/10.1136/jnnp-2011-301940

Tabrizi S. J., Reilmann R., Roos R. A., Durr A., Leavitt B., Owen G., Jones R., Johnson H., Craufurd D., Hicks S. L., & Kennard C. (2012). Potential endpoints for clinical trials in premanifest and early Huntington's disease in the TRACK-HD study: Analysis of 24 month observational data. *The Lancet Neurology*, *11*(1), 42–53. https://doi.org/10.1016/S1474-4422(11)70263-0

van Buuren S., & Groothuis-Oudshoorn K. (2011). MICE: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, *45*(3), 1–67. https://doi.org/10.18637/jss.v045.i03

Wang C., & Pepe M. (2000). Expected estimating equations to accommodate covariate measurement error. *Journal of the Royal Statistical Society. Series B*, *62*(3), 509–524. https://doi.org/10.1111/1467-9868.00247

Wilkes F. A., Abaryan Z., Ching C. R., Gutman B. A., Madsen S. K., Walterfang M., Velakoulis D., Stout J. C., Chua P., Egan G. F., & Thompson P. M. (2019). Striatal morphology and neurocognitive dysfunction in Huntington disease: The IMAGE-HD study. *Psychiatry Research: Neuroimaging*, *291*, 1–8. https://doi.org/10.1016/j.pscychresns.2019.07.003

Xiao Z., Shao J., & Palta M. (2010). GMM in linear regression for longitudinal data with multiple covariates measured with error. *Journal of Applied Statistics*, *37*(5), 791–805. https://doi.org/10.1080/02664760902890005

Yi G., Ma Y., & Carroll R. (2012). A functional generalized method of moments approach for longitudinal studies with missing responses and covariate measurement error. *Biometrika*, *99*(1), 151–165. https://doi.org/10.1093/biomet/asr076

Zhang Y., Long J. D., Mills J. A., Warner J. H., Lu W., Paulsen J. S., & PREDICT-HD Investigators and Coordinators of the Huntington Study Group (2011). Indexing disease progression at study entry with individuals at-risk for Huntington disease. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *156*(7), 751–763. https://doi.org/10.1002/ajmg.b.v156.7

Zhang Y., Zhou J., Gehl C., Long J., Johnson H., Magnotta V., Sewell D., Shannon K., & Paulsen J. (2021). Mild cognitive impairment as an early landmark in Huntington's disease. *Frontiers in Neurology*, *12*, 678652. https://doi.org/10.3389/fneur.2021.678652