

# Supplementary material: Errors-in-variables Modeling of Personalized Treatment-Response Trajectories

Guangyi Zhang, Reza A. Ashrafi, Anne Juuti, Kirsi Pietiläinen, Pekka Marttinen

## I. SPECIFICATION OF MODEL PRIOR DISTRIBUTIONS

### Response function:

$$\tilde{\beta}_l \sim Normal(mu = 0, sd = 5) \quad (1)$$

$$\tilde{\beta}_h \sim Normal(mu = 0, sd = 5) \quad (2)$$

$$\Sigma_l \sim HalfNormal(sd = 1) \quad (3)$$

$$\Sigma_h \sim HalfNormal(sd = 0.5) \quad (4)$$

$$\beta_n^l \sim Normal(mu = \tilde{\beta}_l, sd = \Sigma_l) \quad (5)$$

$$\beta_n^h \sim Normal(mu = \tilde{\beta}_h, sd = \Sigma_h) \quad (6)$$

where the dimension of covariates  $P$  is equal to 1 for notation simplicity. For  $P > 1$ , an independence among different dimensions is assumed.

### Counterfactual trend:

$$ls_n \sim HalfNormal(sd = 10) + 10 \quad (7)$$

$$nu_n \sim HalfNormal(sd = 10) \quad (8)$$

$$c_n \sim HalfNormal(sd = 10) \quad (9)$$

$$K1_n = nu_n * ExpQuad(\cdot|ls_n) \quad (10)$$

$$K2_n = Constant(\cdot|c_n) \quad (11)$$

$$K_n = K1_n + K2_n \quad (12)$$

where  $ExpQuad(\cdot)$  and  $Constant(\cdot)$  are kernel functions. Time is measured in minutes, and the length-scale  $ls_n$  is bounded from below by 10 minutes to increase stability of the estimation (measurements are taken at an interval of 15 minutes in our real-world data set).

### Measurement models:

$$d_n \sim Normal(mu = 0, sd = 10) \quad (13)$$

$$\sigma_n^t \sim HalfNormal(sd = 10) \quad (14)$$

$$t_{nm}^* \sim d_n + \sigma_n^t * Normal(mu = 0, sd = 1) \quad (15)$$

$$\delta_{nm} \sim Lognormal(mu = 0, sd = 0.05) \quad (16)$$

## II. DETAILS OF THE FIRST SIMULATION EXPERIMENT

In total 300 markers consisting of a zero trend and 5 evenly distributed treatments are generated, where 2-dimension input is generated randomly from  $Uniform(0.3, 1)$ . A perturbation rate sampled from  $N(1, 0.2^2)$  are multiplied by each input. The

groundtruth coefficients for the height of treatment response are [5, 3] and [15, 2] for length-scale. The result is shown in Figure 1.

## III. DERIVATION FOR THE MARGINAL INCREMENT OF TREATMENT RESPONSE AREA BY ONE COVARIATE

For simplicity, we focus on a single individual and drop the unnecessary indexing in the notation. The area  $A$  is proportional to length-scale  $l$  and height  $h$  of the response. Hence

$$A = \lambda hl \quad (17)$$

for some constant  $\lambda$  (knowing the shape of the response, solving for  $\lambda$  analytically is straightforward). Denote the amount of one covariate, e.g. sugar, in the  $m$ th meal by  $x_{mi}$  where  $i \in \{1, 2, \dots, P\}$ . Now the length-scale  $l$  depends on  $x$  through

$$l_m(x_{mi}) = g(y_m^l) = g(\beta_i^l x_{mi} + c_m^l), \quad (18)$$

where  $g$  is the *softplus* function and  $c_m^l$  comprises the other parts of the linear predictor that do not depend on the sugar  $x_{mi}$ . Similarly, the height  $h$  depends on  $x$  through

$$h_m(x_{mi}) = g(y_m^h) = g(\beta_i^h x_{mi} + c_m^h). \quad (19)$$

We want to know how area  $A_m$  changes if we change the amount of sugar  $x_{mi}$  by one unit.

$$\frac{dA_m}{dx_{mi}} = \lambda \frac{dl_m}{dx_{mi}} h_m + \lambda l_m \frac{dh_m}{dx_{mi}} \quad (20)$$

$$= \lambda \frac{dl_m}{dy_m^l} \frac{dy_m^l}{dx_{mi}} h_m + \lambda \frac{dh_m}{dy_m^h} \frac{dy_m^h}{dx_{mi}} l_m \quad (21)$$

$$= \lambda(1 + e^{-y_m^l})^{-1} \beta_i^l h_m + \lambda(1 + e^{-y_m^h})^{-1} \beta_i^h l_m. \quad (22)$$

By replacing  $x_m$  with an average meal, we have

$$\frac{dA}{dx_i} = \lambda(1 + e^{-\bar{y}})^{-1} \beta_i^l \bar{h} + \lambda(1 + e^{-\bar{y}})^{-1} \beta_i^h \bar{l} \quad (23)$$

$$= \lambda(1 + e^{-(\beta^l)^T \bar{x}})^{-1} \beta_i^l \bar{h} + \lambda(1 + e^{-(\beta^h)^T \bar{x}})^{-1} \beta_i^h \bar{l}. \quad (24)$$

Manuscript received April 19, 2005; revised August 26, 2015.

Guangyi Zhang, Reza A. Ashrafi, and Pekka Marttinen are with the Department of Computer Science, Aalto University, Finland. (e-mail:{guangyi.zhang,reza.ashrafi,pekka.marttinen}@aalto.fi)

Anne Juuti and Kirsi Pietiläinen are with the Medical School, University of Helsinki, Finland. (e-mail:anne.juuti@hus.fi,kirsi.pietilainen@helsinki.fi)

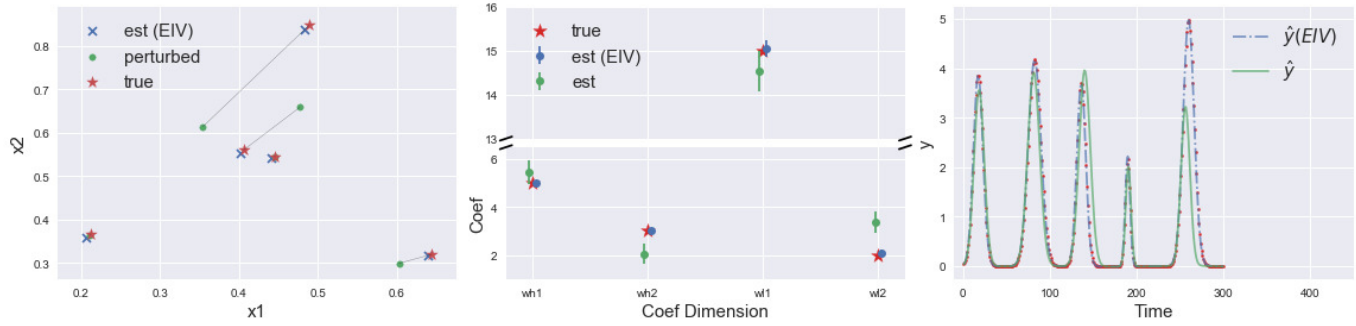


Fig. 1: EIV results in the first simulation experiment. The left plot shows how the perturbations are restored; the middle compares the coefficients estimated with and without EIV; the last one exhibits overall fittings of EIV and baseline models. Red stars and dots stand for groundtruth values and observed points respectively.

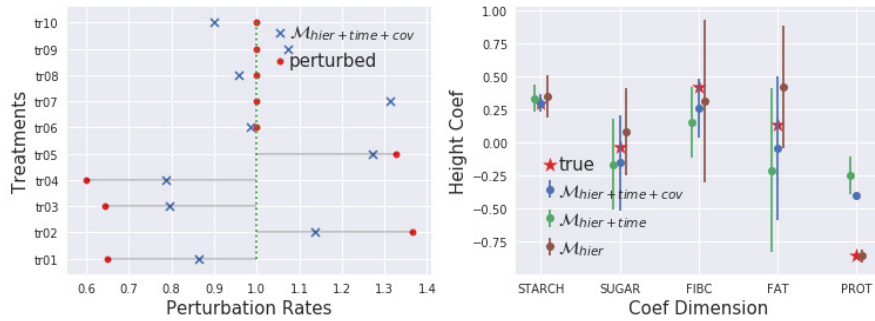


Fig. 2: Results of the simulation experiment in Section IV-A of the main text for the second patient.

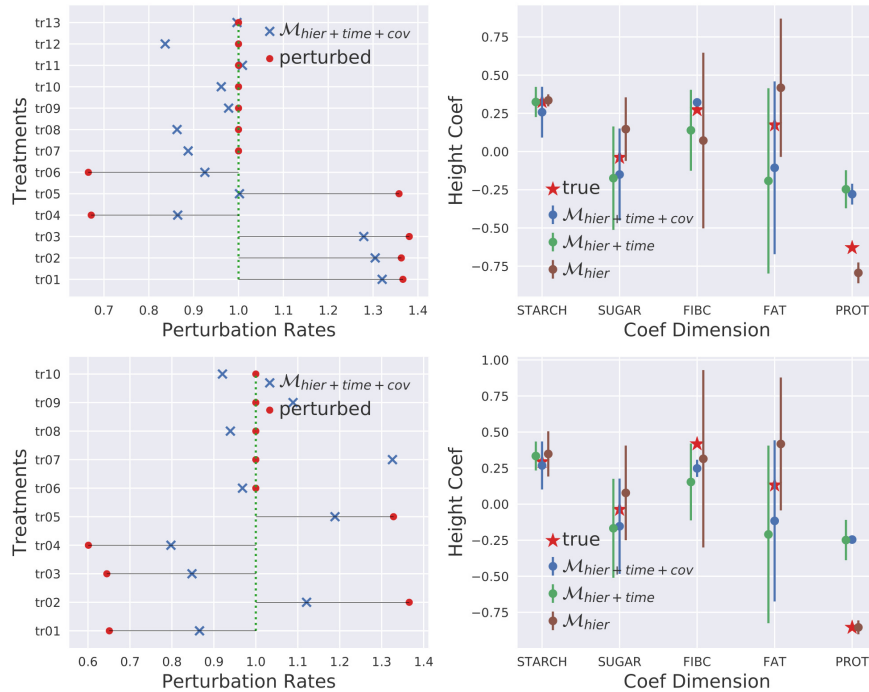


Fig. 3: Replication of the results of the simulation experiment in Section IV-A of the main text, initialized with a different seed. Results for both patients are shown.

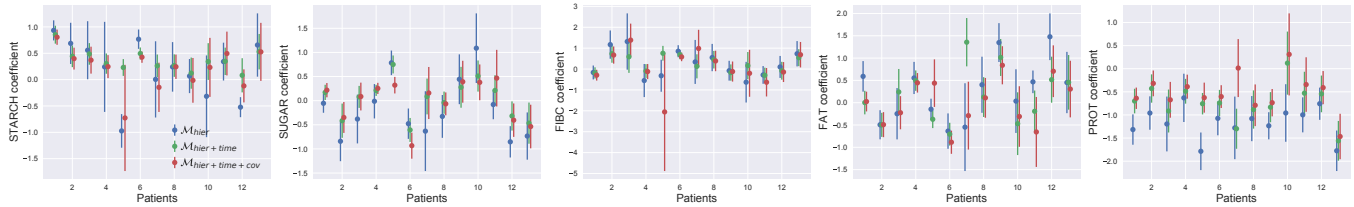


Fig. 4: Posterior uncertainty in the personalized regression coefficients, estimated with different models from the real-world data.  $\mathcal{M}_{ind}$  model is excluded because of its much wider and distracting uncertainty.

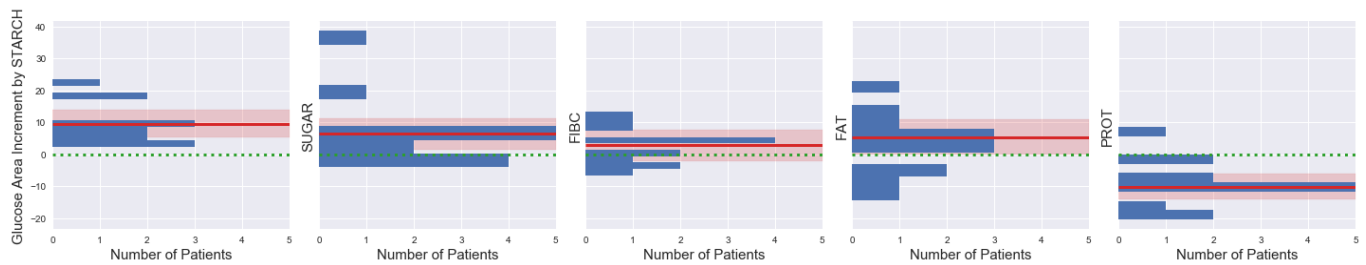


Fig. 5: Histogram of personalized coefficients for different covariates, estimated with the real-world data. Red line and the surrounding region coloured red show mean +/- one SD.

Fig. 6: Visualization of the observed and estimated trajectories for all patients.