Annals of Epidemiology 35 (2019) 53–58

Contents lists available at ScienceDirect

Annals of Epidemiology

Original article

Joint modeling of concurrent binary outcomes in a longitudinal observational study using inverse probability of treatment weighting for treatment effect estimation

George O. Agogo, PhD [a], Terrence E. Murphy, PhD [a, b], Gail J. McAvay, PhD [a], Heather G. Allore, PhD [a, b, *]

[a] Department of Internal Medicine, Section of Geriatrics, Yale School of Medicine, New Haven, CT
[b] Department of Biostatistics, Yale School of Public Health, New Haven, CT

ARTICLE INFO

Article history:
Received 14 October 2018
Accepted 23 April 2019
Available online 2 May 2019

Keywords:
Correlated outcomes
Joint model
Propensity score
Shared-parameter model
Simulation study

ABSTRACT

Purpose: Correlated healthcare utilization outcomes may be encoded as binary outcomes in epidemiologic studies. We demonstrate how to account for correlation between concurrent binary outcomes and confounding by person characteristics when estimating a treatment effect in observational studies.
Methods: We present a joint shared-parameter model, weighted by inverse probability of treatment weights (IPTW) to account for confounding. The model is evaluated in a simulation study that emulates the Medical Expenditure Panel Survey data and compared with a covariate-adjusted joint model and with separate outcome models (IPTW weighted and covariate adjusted).
Results: For the IPTW-weighted joint model, relative bias in the estimated treatment effect on outcome 1 ranged from −0.057 to −0.033 and outcome 2 from −0.077 to −0.043. For the covariate-adjusted joint model, relative bias ranged from −0.010 to −0.083 for outcome 1 and from −0.087 to −0.110 for outcome 2. The covariate-adjusted joint model estimated the effect more closely than the covariate-adjusted separate model. The IPTW-weighted joint model estimated the effect more closely for outcome 1.
Conclusions: The IPTW-weighted joint model handles correlation between binary outcomes, adjusts for confounding, and estimates the treatment effect accurately in observational studies. We illustrate the contribution of person-specific effects in estimating personalized risk.

Introduction

Healthcare utilization (e.g., hospitalization) and patient-centered outcomes (e.g., mobility limitations) are important metrics often represented as binary outcomes. These binary outcomes are usually measured concurrently and are often correlated. For instance, an individual who is frequently hospitalized is also more likely to experience mobility limitations. Modeling these outcomes separately fails to extract shared information regarding a treatment effect, potentially leading to biased estimation [1, 2]. Despite the temporal concurrence of correlated binary outcomes, there is limited research on how to jointly model them [1]. A challenge is the scant information provided by binary outcomes relative to ordinal, continuous, or time-to-event (TTE) outcomes. Joint modeling is commonly used to analyze a TTE outcome and a longitudinal outcome [3,4]. Joint modeling evolved to include shared parameters that can improve estimation for both TTE and longitudinal outcomes [5]. Joint modeling mostly benefits the TTE estimate due to the finer granularity of information contributed by the longitudinal outcome compared with the single TTE outcome [6,7]. A further benefit of the joint shared parameter modeling approach is its ability to account for the measurement error in the observed longitudinal outcomes [8].

Further complexity arises in observational longitudinal studies of concurrent binary outcomes. Observational studies are prone to bias by confounding due to the imbalance in observed baseline characteristics between treatment groups [9–14]. Propensity score (PS) methods are shown to account for confounding bias [15]. Currently, to our knowledge, there is no research using inverse probability of treatment weights (IPTW) to control for confounding in joint modeling of concurrent binary outcomes.

The authors have no conflicts of interest to disclose.

* Corresponding author. Department of Internal Medicine, Section of Geriatrics, Yale School of Medicine, 300 George St, Suite 775, New Haven, CT 06511. Tel: +1 203 737 1892; fax: +1 203 785 4823.
E-mail address: heather.allore@yale.edu (H.G. Allore).

https://doi.org/10.1016/j.annepidem.2019.04.008
1047-2797/

The aim of this study was to demonstrate how to jointly model correlated binary outcomes by using the shared information between longitudinal outcomes while controlling for confounding due to covariate imbalance between treatment groups. We propose a shared-parameter model weighted by IPTW. The shared parameter accounts for the correlation between health outcomes attributable to latent factors [16]. This work was motivated by data from adults with chronic obstructive pulmonary disease (COPD) in the Medical Expenditure Panel Survey (MEPS). We emulated a treatment effect and binary health outcomes based on MEPS dataset and performed a simulation study to estimate the effect of a hypothetical respiratory medication on pairs of concurrent binary outcomes. The proposed model is compared with a joint model adjusted for potential confounders and with the separate outcome models that ignore correlations between binary outcomes. Finally, we illustrate the fundamental contribution of person-specific effects in the estimation of personalized risk.

Material and methods

The MEPS

COPD is characterized by high comorbidity and healthcare utilization providing an illustrative example to evaluate the association of respiratory medications with longitudinal outcomes. We focused on longitudinal pairs of binary outcomes for adults aged 40–85 years with COPD in MEPS from cohort years 2008 through 2012. MEPS is a nationally representative sample of the U.S. civilian population, sponsored by the U.S. Department of Health and Human Services through the Agency for Healthcare Research and Quality and Centers for Disease Control and Prevention [17]. Participants were followed for 2.5 years, with five rounds of interviews occurring approximately every 6 months after the baseline visit. The data for the simulation study came from the Household Component of the survey including respiratory medications, healthcare utilization (outpatient and emergency department visits, and hospitalizations) and measures of function. Whereas medications are typically prescribed based on organ function and mortality, older adults reportedly prefer prescribing decisions based on higher function and relief of symptoms over mortality [18]. Methods that estimate the longitudinal associations of treatment with correlated patient-centered outcomes are needed to improve clinical and patient decision-making. Using the observational-based MEPS dataset to inform the simulation study, we investigated the effects of any respiratory medication on two concurrent patient-centered outcomes, namely, any outpatient visit and any hospital or emergency department visit among adults with COPD. Kendall's tau correlation between these two outcomes is 0.16.

Inverse probability of treatment weights

When estimating a “treatment effect” in an observational study, PS methods are used to control for confounding due to the imbalance of covariates between treatment groups. In MEPS dataset, covariate imbalance between persons taking any respiratory medication and those who do not is evidenced, as quantified by a standardized difference (SD); for instance, the SD for Hispanic race or smoking status, or limitation in mobility is 0.2, SD for any outpatient visit due to respiratory illness is 0.3, and SD for asthma diagnosis is 0.4. The SDs for these covariates are greater than a threshold value 1 for covariance imbalance [19], prompting a need for a statistical technique to correct for the imbalance.

A PS is the probability of receiving a treatment conditional on observed covariates [20]. Focusing on persons with COPD, we denote a binary indicator for respiratory medication for the i-th person by T_i ($T_i = 1$ if treated and $T_i = 0$ if untreated). We denote a vector of observed baseline covariates that may be associated with both the probability of receiving the medication and with the risk of an outcome (Y_i) by $\mathbf{Z}_i$. The PS, $e(\mathbf{Z}_i)$ is given by $e(\mathbf{Z}_i) = \text{Prob}(T_i = 1|\mathbf{Z}_i^{\text{T}})$ and is usually estimated with logistic regression [21]. A PS is a balancing score, meaning that treated and untreated subjects with the same $e(\mathbf{Z}_i)$ will typically have similar distributions of $\mathbf{Z}_i$. There are four assumptions required when using PS methods for causal inference [13]. First, a subject's potential outcome under the treatment received is assumed equal to the subject's observed outcome. Second, it is assumed that there are no unmeasured confounders. Third, each subject is assumed to have a nonzero probability of receiving each treatment, and that there is no combination of the covariates such that the PS equals 1: $0 < \text{Prob}(T_i = t\left|\mathbf{Z}_i^{\text{T}}\right.) < 1$ for all $\mathbf{Z}_i$ and t [22]. Finally, the PS model is assumed to be correctly specified.

In this study, we used PSs as IPTW, $\omega(\mathbf{Z}_i)$. Formally, IPTW is defined as $\omega(\mathbf{Z}_i) = 1/e(\mathbf{Z}_i)$ if treated and $\omega(\mathbf{Z}_i) = 1/(1 - e(\mathbf{Z}_i))$ if untreated. Weighting by $\omega(\mathbf{Z}_i)$ creates pseudo-populations where the distributions of $\mathbf{Z}_i$ between treated and untreated subjects are similar. We explored the distributions of IPTW weights using percentile distributions and graphically using histograms and density plots. The highest IPTW weights were trimmed downward at the 98$^{\text{th}}$ percentile [23].

An IPTW-weighted joint shared-parameter model

We apply a shared random effects model to jointly estimate the longitudinal treatment effect on concurrent binary outcomes. In this approach, the random effects serve as the conduit for sharing of information between correlated binary outcomes arising from unmeasured variables. We control for confounding of treatment using the IPTW weights. The IPTW-weighted shared-parameter model is defined as follows:

$$\begin{aligned}&\text{Prob}\left(\textit{Outcome } 1_{ij} = 1|T_i, \mu_i\right) = \varphi\left(\beta_{01} + \beta_{T_1} T_i + \mu_i\right)\\ &\text{Prob}\left(\textit{Outcome } 2_{ij} = 1|T_i, \mu_i\right) = \varphi\left(\beta_{02} + \beta_{T_2} T_i + \lambda\mu_i\right),\end{aligned} \tag{1}$$

where $\mu_i \sim N(0, \sigma_u^2)$ is the shared random intercept representing the unobserved factors underlying a person's susceptibility to either outcome, λ is the coefficient of the shared random effect for outcome 2, β_{T_1} and β_{T_2} quantify the treatment effects, and φ is the inverse logit link function. Note that model [1] is weighted by IPTW to control for confounding due to covariate imbalance between treatment groups.

We further fit a joint model where all covariates in the PS model and confounders in the outcome models are adjusted in the joint model (hereafter, direct covariate adjustment). Finally, we fit separate outcome models, where confounding was adjusted using IPTW weighting and direct covariate adjustment methods.

Simulation study

The data were simulated at the person level to know the true effect of a hypothetical respiratory medication use (treatment) on two binary outcomes. We, therefore, used data on a person's characteristics for the 942 persons in MEPS with COPD who had complete outcome data over the 2-year follow-up period.

Simulation of each person's treatment indicator

We simulated each person's hypothetical respiratory medication use as a function of covariates using logistic regression. The covariates from MEPS used to predict a person's probability of

receiving respiratory medication include age, perceived general health (1 = excellent to 5 = poor), asthma (present vs. absent), cardiovascular disease (present vs. absent), and ethnicity (Hispanic vs. non-Hispanic). The coefficients for the covariate effects in the simulation model were obtained from the PS model for respiratory medications from MEPS data, whereas the intercept terms were adjusted such that ~50% receiving treatment. The following logistic regression model was used:

$$\begin{aligned}\text{Prob}\,(T_i = 1 | age_i, genhealth_i,\ asthma_i, CVD_i, HISP_i) \\ = \varphi(\alpha_0 + \alpha_1 age_i + \alpha_2 genhealth_i + \alpha_3 asthma_i + \alpha_4 CVD_i \\ + \alpha_5 HISP_i)\end{aligned}$$

where genhealth denotes perceived general health, CVD denotes cardiovascular disease, and HISP denotes Hispanic ethnicity, $\alpha_0 = -1.7, \alpha_1 = 0.01, \alpha_2 = 0.18, \alpha_3 = 0.70, \alpha_4 = 0.40$, and $\alpha_5 = -0.51$. Each person's covariate values from the MEPS cohort were used to simulate that person's treatment indicator, which was subsequently used to simulate the outcomes.

Simulation of each person's binary outcomes

For each person, we simulated five repeated measures of the pair of binary outcomes that share information through covariates and correlated random intercepts as follows:

$$\begin{aligned}\text{Prob}\,(outcome1_{ij} = 1 | T_i, asthma, \mu_{1i}) \\ = \varphi(\beta_{01} + \beta_{T_1} T_i + \beta_{11} asthma_i + \beta_{21} CVD_{med_i} + \mu_{1i})\end{aligned}$$

$$\begin{aligned}\text{Prob}\,(outcome2_{ij} = 1 | T_i, asthma, \mu_{2i}) \\ = \varphi(\beta_{02} + \beta_{T_2} T_i + \beta_{12} asthma_i + \beta_{22} CVD_{med_i} + \mu_{2i}),\end{aligned} \quad (2)$$

where $j = 1, \ldots, 5$ represent the repeated measures of each outcome for i-th person, $\beta_{T_1} = \beta_{T_2} = -0.3$, $\beta_{11} = 0.2$, $\beta_{12} = 0.1$, $\beta_{21} = \beta_{22} = -0.1$, $\sigma_u = 1$; β_{01} and β_{02} were varied to examine whether the frequency of each outcome influenced results. Note that λ was set to 1 to allow for a nondifferential contribution of the shared random effect on both outcomes, reflecting the same algorithm implemented in the SAS GLIMMIX procedure used to fit these models. Therefore, we examined the following three pairings of frequencies: both outcomes of high frequency (~50%), one outcome of high frequency (~50%) and one of low frequency (~20%), and both outcomes of low frequency (~20%). The chosen treatment effect is within the range of meta-analysis of COPD interventions, including medications and hospital readmission [24]. To obtain high frequency (~50%) for both outcomes, β_{01} was set to 0.2 and β_{02} to 0.32; to obtain the pairing of high-frequency outcome (~50%) and low-frequency outcome (~20%), β_{01} was set to 0.2 and β_{02} to −1.32. For the pairing of two low-frequency outcomes (~20%), β_{01} was set to −1.2 and β_{02} to −1.32. These frequencies are within the range observed for patient-centered outcomes in MEPS. The continuous random effects μ_1 and μ_2 are correlated with a Pearson correlation coefficient, $\rho = 0.9$ to obtain meaningful marginal correlation between pairs of binary outcomes. To study the contribution of random effects in estimating personalized risk, we randomly selected two persons with the same observed characteristics but different random effects and compared their individual probabilities for each outcome with the marginal probabilities.

Model fitting

In practice, the true distribution of the data is unknown and is approximated with a plausible statistical model; therefore, to mimic reality, we fit several different plausible models in our simulations. We first fit the PS model to estimate the IPTW using the LOGISTIC procedure in SAS. The covariates in the fitted PS model were age, general health status, asthma, CVD, and Hispanic ethnicity. The estimated IPTW weights, trimmed at the 98th percentile, were used in the IPTW-weighted joint model shown in Equation (1) [23]. The covariates in the fitted IPTW-weighted joint model were any respiratory medication and the shared random effects. Within a person, the two concurrent binary outcomes share a common random intercept that permits the sharing of longitudinal information. We fit the joint model using the GLIMMIX procedure in SAS, with a variance component structure for the random intercept and Newton–Raphson with ridging for optimization. The GLIMMIX procedure implicitly assumes a nondifferential contribution of the random effects on both outcomes. The standard errors were estimated using a robust sandwich estimator to account for weighting and correlated repeated measurements within person. We further fit a shared-parameter joint model where the covariates were adjusted directly in the model. The covariates in this model included any respiratory medication, age, general health status, asthma, CVD medication, Hispanic ethnicity, and shared-parameter random effects. Finally, we fit an IPTW-weighted separate model for each outcome and a separate outcome model with direct covariate adjustment to investigate the contribution of outcome correlations in estimating the treatment effect. For comparison purposes, we further fitted the model used to simulate the data (hereafter, the true model). The simulation study and statistical analyses were performed using SAS version 9.4 (SAS Institute, Inc., Cary, NC).

Model evaluation

In this study, we generated and evaluated the first 1000 datasets where convergence was achieved, where each simulation consisted of 942 persons with five repeated measurements per outcome. The accuracy of each estimated treatment effect, β_{T_h}, h = 1, 2, was assessed with the following four measures:

(1) Relative bias in $\widehat{\beta}_{T_h}$: $\text{Rel.bias}(\widehat{\beta}_{T_h}) = \frac{\overline{\widehat{\beta}}_{T_h} - \overline{\beta}_{T_h}}{\overline{\beta}_{T_h}}$,
(2) Empirical standard error: $\text{Emp. SE}\,(\widehat{\beta}_{T_h}) = \frac{1}{(N-1)} \sum_{k=1}^{N} (\widehat{\beta}_{T_{h,k}} - \overline{\widehat{\beta}}_{T_h})^2$, that is, the standard deviation of $\widehat{\beta}_{T_h}$ that quantifies the uncertainty in estimating β_{T_h} over N simulated datasets,
(3) Mean squared error: $\text{MSE}(\widehat{\beta}_{T_h}) = (\overline{\widehat{\beta}}_{T_h} - \overline{\beta}_{T_h})^2 + (\text{Emp.SE}(\widehat{\beta}_{T_h}))^2$,
(4) Coverage, defined as the proportion of times the 95% confidence interval of $\widehat{\beta}_{T_h}$ contains β_{T_h}.

To illustrate personalized risk for the two randomly selected persons with the same observed characteristics who had different random effects, we simulated 500 datasets each consisting of five repeated measures per outcome per person. We compared their predicted probabilities from the IPTW-weighted shared-random effects joint model with the marginal probabilities from the IPTW-weighted generalized estimating equation joint model.

Results

With the simulation specifications described previously, Kendall's tau that quantifies the marginal correlation between the binary outcomes ranged from 0.12 to 0.15 for the three pairwise combinations of outcome frequencies. Simulation results from the joint and separate models (IPTW weighted and covariate adjusted) in comparison with the results from the true model are presented in Table 1.

Comparing the results from the joint models, IPTW-weighted joint models yielded lower relative bias, greater coverage but

Table 1
Treatment effect mean, bias, relative bias, empirical standard error, mean squared error, and coverage probability of log odds ratio estimate for effects of treatment on binary concurrent outcomes for varying frequencies of pairings of outcomes in 1000 simulated datasets using true models and joint and separate models

Outcome pairings by frequency	Informational content of outcomes	Models	Outcome 1 (Out1)				Outcome 2 (Out2)			
			Relative bias	Empirical SE	MSE	Coverage	Relative bias	Empirical SE	MSE	Coverage
		True model	0.013	0.085	0.007	0.956	0.023	0.088	0.008	0.945
Out1 (50.6%)	Both high	Joint IPTW	−0.043	0.084	0.007	0.951	−0.043	0.087	0.008	0.948
Out2 (52.4%)	Tau = 0.15	Joint covariates	−0.083	0.079	0.007	0.948	−0.097	0.081	0.007	0.924
		Separate IPTW	−0.053	0.086	0.008	0.944	0.000	0.087	0.008	0.952
		Separate covariates	−0.133	0.079	0.008	0.913	−0.130	0.075	0.007	0.935
		True model	0.017	0.087	0.008	0.952	0.027	0.096	0.009	0.953
Out1 (50.8%)	High and low	Joint IPTW	−0.033	0.088	0.008	0.950	−0.060	0.096	0.010	0.942
Out2 (22.0%)	Tau = 0.13	Joint covariates	−0.100	0.078	0.007	0.940	−0.087	0.089	0.009	0.943
		Separate IPTW	−0.063	0.088	0.008	0.947	−0.047	0.097	0.010	0.939
		Separate covariates	−0.137	0.077	0.008	0.921	−0.133	0.085	0.009	0.932
		True model	0.007	0.096	0.009	0.945	0.053	0.099	0.010	0.949
Out1 (24.2%)	Both low	Joint IPTW	−0.057	0.095	0.009	0.949	−0.077	0.097	0.010	0.941
Out2 (22.2%)	Tau = 0.12	Joint covariates	−0.093	0.088	0.009	0.942	−0.110	0.089	0.009	0.933
		Separate IPTW	−0.093	0.091	0.009	0.950	−0.033	0.095	0.009	0.951
		Separate covariates	−0.113	0.086	0.009	0.923	−0.117	0.085	0.008	0.935

The simulation study design emulates the design of Medical Expenditure Panel Survey, the United States, 2008–2012.
Emp. SE = empirical standard error; MSE = mean squared error; Rel. bias = relative bias; SE = standard error.

higher empirical standard error for the treatment effect estimate than the covariate-adjusted joint model. For example, when the frequency of outcome 1 is 50.6% and that of outcome 2 is 52.4%, with IPTW weighting, the relative bias is reduced by about 48% (from −0.083 to −0.043) for β_{T_1} and by 56% (from −0.097 to −0.043) for β_{T_2}; the coverage increases modestly by 0.3% (from 0.948 to 0.951) for β_{T_1} and by 2.6% (from 0.924 to 0.948) for β_{T_2}. In contrast, the empirical standard error increases modestly by 6% (from 0.079 to 0.084) for β_{T_1} and by 7% (from 0.081 to 0.087) for β_{T_2}. There were no substantial differences in the mean square errors. Similar trends are observed for IPTW-weighted joint model and covariate-adjusted joint model for the other pairings of outcome frequencies.

The treatment effect on the low-frequency outcome is estimated more accurately when jointly modeled with a high-frequency outcome than with a second low-frequency outcome. Using the IPTW-weighted joint model for the pairing of high- and low-frequency outcomes, the relative bias for the treatment effect on outcome 2 is reduced by 22% (from −0.077 to −0.060) relative to the pairing of the two low-frequency outcomes. A similar trend is observed with covariate-adjusted joint model. When both outcomes have low frequencies, the treatment effects exhibit more bias, greater uncertainty, and lower coverage than when both outcomes have high frequencies. Using IPTW-weighted joint model, relative bias increases as the outcome frequency decreases; specifically, by 32.6% (from −0.043 to −0.057) for β_{T_1} and by 79.1% (from −0.043 to −0.077) for β_{T_2} high-frequency outcome pairings compared with low-frequency pairings. The corresponding empirical standard errors increase by 13% (from 0.084 to 0.095) for β_{T_1} and by 10% (from 0.087 to 0.096) for β_{T_2}.

For all outcome pairings, treatment effects estimates from the joint models show lower bias than those from the separate models, except for the treatment effect estimate for outcome 2 when using the IPTW-weighted joint model. As is expected for all outcome pairings, the treatment effects are estimated with the least bias using the true models.

To illustrate the impact of subject-specific random effects in the IPTW-weighted joint model, we randomly selected two 63-year-old, non-Hispanic persons with asthma and cardiovascular disease taking respiratory medication. These two persons had different probabilities of experiencing each outcome due to the differences in their person-specific random intercepts (Fig. 1). This implies that these two persons have different latent factors that underlie their susceptibility to experiencing the two outcomes. There are observable differences in the predicted probabilities of each outcome between the two persons as shown by the mean estimates in the boxplots. Further shown in Figure 2 are the locations of their random effects within the distribution of the study population, that is, −0.86 for person A and −0.11 for person B. It is further evident that both values lie in the lower half of the distribution (<0), explaining why their mean predicted probabilities (shown in Fig. 1) are slightly lower than the mean probabilities of the study population (0.51 for outcome 1 and 0.22 for outcome 2). Using the IPTW-weighted generalized estimating equation joint model (without random effects) resulted in the same probabilities (0.48 for outcome 1 and 0.20 for outcome 2) for the two persons.

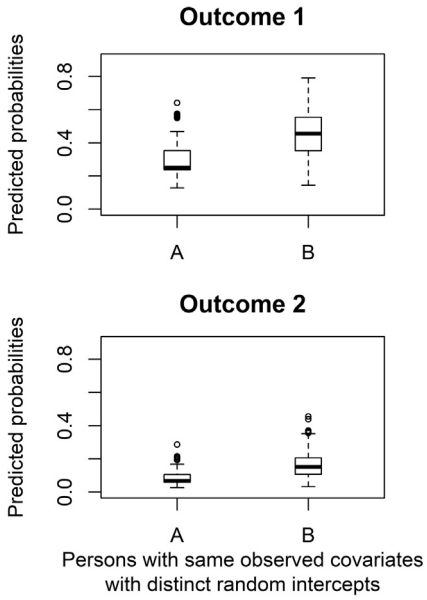

Fig. 1. Predicted probabilities of outcomes 1 and 2 from the IPTW-weighted joint model for two pseudo persons with the same observed characteristics and who took the same medications but have different random intercepts (−0.86 for person A and −0.11 for person B) from 500 simulated datasets, each consisting of five repeated measures per outcome per person, when one outcome has a high frequency ~50% and the other outcome has a low frequency ~20%.

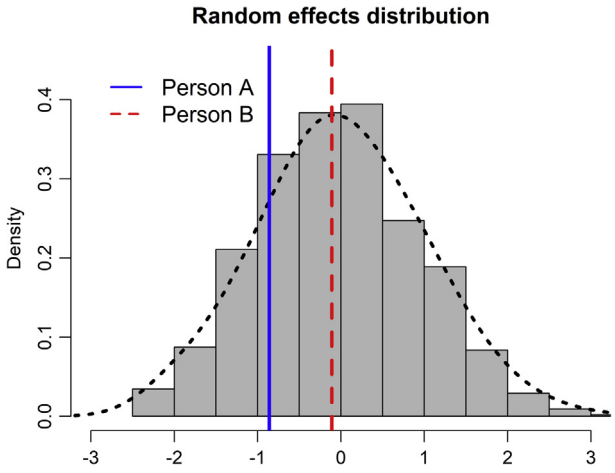

Fig. 2. Location of random intercepts (−0.86 for person A and −0.11 for person B) from the IPTW-weighted joint model for two persons with same observed characteristics and who took same medications in the distribution spectrum of the study population from 500 simulated datasets, each consisting of five repeated measures for each outcome per person, when one outcome has high frequency ~50%, and the other outcome has low frequency ~20%.

Discussion

Observational studies of correlated binary health outcomes may be confounded by a person's characteristics. In this study, we demonstrated how to account for confounding and the longitudinal information shared between pairs of binary health outcomes. The proposed IPTW-weighted joint model accounts for the interdependency of the binary outcomes with a shared random intercept [25,26]. Estimating treatment effects on each binary outcome separately fails to leverage the shared information between outcomes, potentially leading to reduced efficiency, especially when covariates are directly adjusted in the outcome model [1,27]. Similarly, failure to account for confounding by observed baseline covariates leads to biased effect estimation [9].

Using a simulation study, we evaluated the proposed IPTW-weighted joint shared-parameter model by estimating the effect of respiratory medication on two longitudinal binary outcomes. The proposed model approximated the treatment effect more closely than covariate-adjusted joint model and separate model. We found that the treatment effect on a low-frequency outcome was estimated more precisely when jointly modeled with a high-frequency outcome than when modeled with another low-frequency outcome. This finding is similar to other studies that found a more efficient estimation of treatment effect on a survival outcome when jointly modeled with a longitudinal outcome or when multivariate continuous outcomes are jointly modeled [5,7].

A potential benefit of our method is its ability to predict personalized risk of health outcomes as illustrated with two persons with the same baseline characteristics having different probabilities of experiencing an outcome due to different shared person-specific random effects (latent quantities) [16]. These latent factors can have a clinical meaning, reflecting, for instance, a person's unmeasured genetic risk factors or immune system that predisposes the person to the risk of developing the outcomes. Differences in risk predictions among persons with the same observed characteristics but differing person-specific random effects reflect a more personalized estimate of a medication's effect on paired outcomes.

The joint modeling framework used here and in recent studies has several strengths [1,2,4,5,26,28,29]. First, jointly modeling pairs of correlated concurrent outcomes account for correlation of the binary data, which is ubiquitous in health outcomes research. Such modeling approach can lead to improved accuracy and efficiency in the estimation of treatment effects [30]. Second, the proposed joint model can be implemented in standard statistical software, such as the GLIMMIX procedure in SAS.

The primary limitation of this study is that the marginal correlation between binary outcomes is strongly bounded, not permitting any exploration between pairs of binary outcomes with marginal correlation greater than 0.2. The observed low marginal correlation results from the simulation of individual values with person-specific random effects, which yield a highly restricted range of marginal correlations between the outcomes [31]. However, given the paucity of literature on the joint modeling of concurrent binary outcomes, its utility in this restricted range is novel and informative. A further limitation involves using the method in a rather simple setting with only two binary outcomes, time-invariant covariates, and complete data. A direction for future research would be to extend the IPTW-weighted joint model to multiple outcomes with different distributional forms (e.g., counts and ordinal), characterized by missing data and time-varying covariates. Multiple outcomes can be modeled jointly by conducting a pairwise modeling approach, where joint models for all pairs of outcomes are fit and combined [25]. Besides investigating the main effects of respiratory treatment, the proposed model can be extended by including treatment by time interaction as covariate to allow for the effect of treatment on the joint outcomes to depend on the follow-up time. For the case of missing data, the proposed method can be extended to include the missing data component. The missingness pattern, denoted by a binary missingness indicator, can be modeled using a logistic or probit model, conditioned on potential predictors of the missing outcome data. This approach allows for nonignorable missingness and can potentially account for unmeasured sources of selection bias [32,33]. When dealing with time-varying covariates and unmeasured confounding, the proposed method can be adapted akin to the modeling approach proposed by Shardell and Ferrucci [32]. Shardell and Ferrucci [32] approach combines a parametric joint mixed-effects model for the study outcome and the exposure with g-computation to identify and estimate effects in the presence of time-varying confounding and unmeasured confounding. The authors further extended their method to handle missing data in the outcomes. Besides IPTW weighting and direct covariate adjustment, other confounding adjustment methods, such as using PS as a covariate, stratification by PS, can be used [34]. For instance, using PS as a covariate method in the joint model resulted in treatment effect estimates with relative bias of −0.097 for a high-frequency (52%) outcome and 0.090 for a low-frequency (22%) outcome and are comparable to the estimates obtained using the fitted joint modeling approaches.

Conclusions

In conclusion, the proposed method exploits the shared temporal information in pairs of correlated binary outcomes, leading to efficient and low-biased joint estimation of treatment effects in observational studies. The method further elucidates the importance of the joint modeling of correlated health outcomes and the contribution of person-specific random effects in the prediction of personalized risk.

Acknowledgments

The authors acknowledge the important contributions of the participants and staff of the Medical Expenditure Panel Survey, including the U.S. Department of Health and Human Services,

Agency for Healthcare Research and Quality, and Centers for Disease Control and Prevention.

This work was supported by U.S. National Institute of Aging at the National Institutes of Health (grant numbers R01 AG047891, P30 AG021342).

References

[1] Lipsitz SR, Fitzmaurice GM, Ibrahim JG, Sinha D, Parzen M, Lipshultz S. Joint generalized estimating equations for multivariate longitudinal binary outcomes with missing data: an application to AIDS data. J R Stat Soc Ser A Stat Soc 2009;172(1):3–20.
[2] Gebregziabher M, Zhao Y, Dismuke C, Axon N, Hunt KJ, Egede LE. Joint modeling of multiple longitudinal cost outcomes using multivariate generalized linear mixed models. Health Serv Outcomes Res Methodol 2013;13(1):39–57.
[3] Tsiatis AA, DeGruttola V, Wulfsohn MS. Modeling the relationship of survival to longitudinal data measured with error: applications to survival and CD4 counts in patients with AIDS. J Am Stat Assoc 1995;90:27–37.
[4] Wulfsohn MS, Tsiatis AA. A joint model for survival and longitudinal data measured with error. Biometrics 1997;53(1):330–9.
[5] Henderson R, Diggle P, Dobson A. Joint modelling of longitudinal measurements and event time data. Biostatistics 2000;1(4):465–80.
[6] Guo X, Carlin BP. Separate and joint modeling of longitudinal and event time data using standard computer packages. Am Stat 2004;58(1):16–24.
[7] Murphy TE, Allore HG, Han L, Peduzzi PN, Gill TM, Xu X, et al. A longitudinal, observational study with many repeated measures demonstrated improved precision of individual survival curves using Bayesian joint modeling of disability and survival. Exp Aging Res 2015;41(3):221–39.
[8] Rizopoulos D. Joint models for longitudinal and time-to-event data: with applications in R. Chapman and Hall/CRC: Boca Raton, FL; 2012.
[9] D'Agostino Jr RB. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. Stat Med 1998;17(19):2265–81.
[10] Austin PC, Mamdani MM. A comparison of propensity score methods: a case-study estimating the effectiveness of post-AMI statin use. Stat Med 2006;25(12):2084–106.
[11] Austin PC, Stuart EA. Estimating the effect of treatment on binary outcomes using full matching on the propensity score. Stat Methods Med Res 2017;26(6):2505–25.
[12] Austin PC, Stuart EA. The performance of inverse probability of treatment weighting and full matching on the propensity score in the presence of model misspecification when estimating the effect of treatment on survival outcomes. Stat Methods Med Res 2017;26(4):1654–70.
[13] Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. Stat Med 2015;34(28):3661–79.
[14] Thorlund K, Mills E. Stability of additive treatment effects in multiple treatment comparison meta-analysis: a simulation study. Clin Epidemiol 2012;4:75.
15 Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika 1983;70(1):41–55.
[16] Gao SJ. A shared random effect parameter approach for longitudinal dementia data with non-ignorable missing data. Stat Med 2004;23(2):211–9.
[17] MEPS survey background. Rockville, MD: Agency for Healthcare Research and Quality, https://meps.ahrq.gov/mepsweb/about_meps/survey_back.jsp. [Accessed 5 February 2018].
[18] Fried TR, Fragoso CAV, Rabow MW. Caring for the older person with chronic obstructive pulmonary disease. JAMA 2012;308(12):1254–63.
[19] Austin PC. Some methods of propensity-score matching had superior performance to others: results of an empirical investigation and Monte Carlo simulations. Biom J 2009;51(1):171–84.
[20] Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika 1983;70(1):41–55.
[21] Rosenbaum PR, Rubin DB. Constructing a control-group using multivariate matched sampling methods that incorporate the propensity score. Am Stat 1985;39(1):33–8.
[22] McCaffrey DF, Griffin BA, Almirall D, Slaughter ME, Ramchand R, Burgette LF. A tutorial on propensity score estimation for multiple treatments using generalized boosted models. Stat Med 2013;32(19):3388–414.
[23] Lee BK, Lessler J, Stuart EA. Weight trimming and propensity score weighting. PLoS One 2011;6(3):e18174.
[24] Pedersen PU, Ersgard KB, Soerensen TB, Larsen P. Effectiveness of structured planned post discharge support to patients with chronic obstructive pulmonary disease for reducing readmission rates: a systematic review. JBI Database Syst Rev Implement Rep 2017;15(8):2060–86.
[25] Fieuws S, Verbeke G. Pairwise fitting of mixed models for the joint modeling of multivariate longitudinal profiles. Biometrics 2006;62(2):424–31.
[26] Fieuws S, Verbeke G. Joint models for high-dimensional longitudinal data. Boca Raton: Crc Press-Taylor & Francis Group; 2009.
[27] Roy J, Lin X. Latent variable models for longitudinal data with multiple continuous outcomes. Biometrics 2000;56(4):1047–54.
[28] Proust-Lima C, Letenneur L, Jacqmin-Gadda H. A nonlinear latent class model for joint analysis of multivariate longitudinal data and a binary outcome. Stat Med 2007;26(10):2229–45.
[29] Proust-Lima C, Joly P, Dartigues JF, Gadda HJ. Joint modelling of multivariate longitudinal outcomes and a time-to-event: a nonlinear latent class approach. Comput Stat Data Anal 2009;53(4):1142–54.
[30] Berridge DM, Crouchley R. Multivariate generalized linear mixed models using R. CRC Press: Boca Raton, FL; 2011.
[31] Demirtas H, Hedeker D. A practical way for computing approximate lower and upper correlation bounds. Am Stat 2011;65(2):104–9.
[32] Shardell M, Ferrucci L. Joint mixed-effects models for causal inference with longitudinal data. Stat Med 2018;37(5):829–46.
[33] Agogo GO, Ramsey CM, Gnjidic D, Moga DC, Allore H. Longitudinal associations between different dementia diagnoses and medication use jointly accounting for dropout. Int Psychogeriatr 2018;30:1477–87.
[34] Schmoor C, Gall C, Stampf S, Graf E. Correction of confounding bias in non-randomized studies by appropriate weighting. Biometrical J 2011;53(2):369–87.