**Appendix**

***Developing prediction models when there are systematically missing predictors in an individual patient data meta-analysis***

Michael Seo1,2 \*, Toshi A. Furukawa3, Eirini Karyotaki4,5,6, Orestis Efthimiou1,7

*1 Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland*

*2 Graduate School for Health Sciences, University of Bern, Bern, Switzerland*

*3Departments of Health Promotion and Human Behavior and of Clinical Epidemiology, Kyoto University Graduate School of Medicine/School of Public Health, Kyoto, Japan*

*4Department of Global Health and Social Medicine, Harvard Medical School, Boston, USA*

*5Department of Clinical Neuro- and Developmental Psychology, Vrije Universiteit Amsterdam, the Netherlands*

*6Amsterdam Public Health Research Institute, Amsterdam, the Netherlands*

*7 Department of Psychiatry, University of Oxford, Oxford, United Kingdom*

*\* corresponding author, email:* [*swj8874@gmail.com*](mailto:swj8874@gmail.com)

**Table 1** Percentage of missing data for each study and each predictor in the illustrative example.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Study | PHQ-9 scores | Baseline | Sex | Age | Relationship status | Comorbid anxiety | Previous episodes | Medication | Alcohol |
| De Graaf 2009 | 4 | 0 | 0 | 0 | 2 | 0 | 0 | 100 | 0 |
| Farrer 2011 | 30 | 0 | 0 | 0 | 0 | 1 | 1 | 100 | 1 |
| Geraedts 2014 | 25 | 0 | 0 | 0 | 0 | 0 | 100 | 100 | 100 |
| Gilbody 2015 | 24 | 0 | 0 | 100 | 1 | 100 | 100 | 100 | 100 |
| Johansson 2012 | 6 | 0 | 2 | 0 | 0 | 0 | 100 | 0 | 100 |
| Kivi 2014 | 27 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 4 |
| Klein 2016 high | 24 | 0 | 0 | 0 | 0 | 0 | 0 | 24 | 0 |
| Klein 2016 low | 19 | 0 | 0 | 0 | 0 | 0 | 0 | 19 | 0 |
| Meyer 2015 | 18 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 |
| Montero-Marin 2016 | 21 | 0 | 0 | 0 | 0 | 100 | 22 | 0 | 100 |
| Philips 2014 | 47 | 0 | 2 | 2 | 2 | 4 | 100 | 100 | 100 |
| Rosso 2016 | 22 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 100 |

**Table 2**: Estimated parameters for the prediction model obtained from the dataset in psychotherapies for depression using the naïve method

|  |  |  |
| --- | --- | --- |
| **Parameter** | | **Estimates**  **(Standard Error)** |
| Average Intercept | | 1.93 (0.67) |
| Average treatment effect | | -0.48 (0.30) |
| **Main**  **effects** | Baseline PHQ-9 score | 0.61 (0.04) |
| Sex | -0.10 (0.30) |
| **Effect modifiers**  **(treatment-covariate interaction)** | Baseline PHQ-9 score | -0.07 (0.05) |
| Sex | 0.14 (0.41) |

**Table 3**: Estimated parameters for the prediction model obtained from the dataset in psychotherapies for depression using the imputation method ignoring clustering in studies

|  |  |  |
| --- | --- | --- |
| **Parameter** | | **Estimates**  **(Standard Error)** |
| Average Intercept | | 0.60 (1.06) |
| Average treatment effect | | -0.02 (1.30) |
| **Main**  **effects** | Baseline PHQ-9 score | 0.61 (0.04) |
| Sex | -0.07 (0.31) |
| Age | 0.02 (0.02) |
| Relationship Status | -0.28 (0.30) |
| Comorbid Anxiety | 0.21 (0.37) |
| Number of Previous Episodes | 0.70 (0.46) |
| Medication | 0.30 (0.39) |
| Alcohol | 0.09 (0.77) |
| **Effect modifiers**  **(treatment-covariate interaction)** | Baseline PHQ-9 score | -0.07 (0.05) |
| Sex | 0.13 (0.43) |
| Age | 0.01 (0.02) |
| Relationship Status | -0.38 (0.41) |
| Comorbid Anxiety | 0.05 (0.46) |
| Number of Previous Episodes | -0.22 (0.54) |
| Medication | -0.42 (0.50) |
| Alcohol | -0.56 (1.05) |

**Table 4**: Estimated parameters for the prediction model obtained from the dataset in psychotherapies for depression using the imputation method accounting for clustering in studies

|  |  |  |
| --- | --- | --- |
| **Parameter** | | **Estimates**  **(Standard Error)** |
| Average Intercept | | 0.93 (1.16) |
| Average treatment effect | | 0.24 (1.38) |
| **Main**  **effects** | Baseline PHQ-9 score | 0.61 (0.04) |
| Sex | -0.08 (0.32) |
| Age | 0.02 (0.02) |
| Relationship Status | -0.25 (0.30) |
| Comorbid Anxiety | 0.30 (0.41) |
| Number of Previous Episodes | 0.52 (0.47) |
| Medication | 0.02 (0.43) |
| Alcohol | -0.09 (1.01) |
| **Effect modifiers**  **(treatment-covariate interaction)** | Baseline PHQ-9 score | -0.09 (0.06) |
| Sex | 0.14 (0.47) |
| Age | 0.00 (0.02) |
| Relationship Status | -0.41 (0.42) |
| Comorbid Anxiety | 0.09 (0.51) |
| Number of Previous Episodes | -0.21 (0.58) |
| Medication | -0.45 (0.49) |
| Alcohol | 0.13 (1.17) |

**Table 5**: Overview of the scenarios explored in the simulation study and the corresponding results. For each scenario, we simulated 100 independent datasets. MSE: mean squared error. For each scenario, the lowest MSE and highest R-squared across the four methods are bolded.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **DESCRIPTION OF SCENARIOS** | | | | | | **SIMULATION RESULTS** | | | |
| **Scenario** | **# of studies** | **# of predictors** | **probability of each predictor to be systematically missing in each study** | **Magnitude of predictor effects** | **Heterogeneity of predictor effects** | **Naïve method**  **(MSE, R-squared, # of failed simulations)** | **Imputation method ignoring heterogeneity**  **(MSE, R-squared, # of failed simulations)** | **Imputation method accounting for heterogeneity**  **(MSE, R-squared, # of failed simulations)** | **Ensemble method**  **(MSE, R-squared, # of failed simulations)** |
| 1 | 2 | 5 | 0.1 | 0.2 | 0.1 | (1.25, 0.07, 0) | (1.25, 0.07, 5) | 48% failure | (**1.24**, **0.08**, 0) |
| 2 | 3 | 5 | 0.1 | 0.2 | 0.1 | (1.22, 0.07, 0) | (1.20, 0.09, 1) | (1.20, 0.09, 9) | (**1.20**, **0.09**, 0) |
| 3 | 5 | 5 | 0.1 | 0.2 | 0.1 | (1.22, 0.08, 0) | (1.17, 0.11, 0) | (1.17, 0.11, 0) | (**0.17**, **0.11**, 0) |
| 4 | 10 | 5 | 0.1 | 0.2 | 0.1 | (1.23, 0.07, 0) | (1.15, 0.13, 0) | (**1.15**, **0.13**, 0) | (1.15, 0.13, 0) |
| 5 | 2 | 10 | 0.1 | 0.2 | 0.1 | (1.34, 0.13, 0) | (1.31, 0.15, 8) | 77% failure | (**1.30**, **0.16**, 0) |
| 6 | 3 | 10 | 0.1 | 0.2 | 0.1 | (1.33, 0.14, 0) | (1.27, 0.18, 2) | 22% failure | (**1.27**, **0.18**, 0) |
| 7 | 5 | 10 | 0.1 | 0.2 | 0.1 | (1.32, 0.14, 0) | (1.23, 0.20, 0) | (1.23, 0.20, 1) | (**1.23**, **0.21**, 0) |
| 8 | 10 | 10 | 0.1 | 0.2 | 0.1 | (1.40, 0.10, 0) | (1.20, 0.23, 0) | (**1.20**, **0.23**, 0) | (1.21, 0.22, 0) |
| 9 | 2 | 5 | 0.3 | 0.2 | 0.1 | (1.29, 0.03, 0) | (1.24, 0.07, 25) | 91% failure | (**1.24**, **0.08**, 0) |
| 10 | 3 | 5 | 0.3 | 0.2 | 0.1 | (1.27, 0.04, 0) | (1.22, 0.08, 9) | 56% failure | (**1.21**, **0.09**, 0) |
| 11 | 5 | 5 | 0.3 | 0.2 | 0.1 | (1.28, 0.04, 0) | (1.18, 0.11, 0) | (1.18, 0.11, 9) | (**1.18**, **0.11**, 0) |
| 12 | 10 | 5 | 0.3 | 0.2 | 0.1 | (1.28, 0.03, 0) | (1.16, 0.13, 0) | (**1.16**, **0.13**, 0) | (1.17, 0.12, 0) |
| 13 | 2 | 10 | 0.3 | 0.2 | 0.1 | (1.41, 0.09, 0) | (1.35, 0.13, 57) | 100% failure | (**1.30**, **0.16**, 0) |
| 14 | 3 | 10 | 0.3 | 0.2 | 0.1 | (1.45, 0.07, 0) | (1.33, 0.14, 21) | 91% failure | (**1.29**, **0.17**, 0) |
| 15 | 5 | 10 | 0.3 | 0.2 | 0.1 | (1.47, 0.05, 0) | (1.25, 0.19, 1) | 23% failure | (**1.25**, **0.19**, 0) |
| 16 | 10 | 10 | 0.3 | 0.2 | 0.1 | (1.51, 0.03, 0) | (1.21, 0.22, 0) | (**1.21**, **0.22**, 0) | (1.23, 0.21, 0) |
| 17 | 2 | 5 | 0.1 | 0.5 | 0.1 | (1.36, 0.44, 0) | (**1.25**, **0.48**, 5) | 48% failure | (1.26, 0.48, 0) |
| 18 | 3 | 5 | 0.1 | 0.5 | 0.1 | (1.39, 0.42, 0) | (1.21, 0.49, 1) | (**1.21**, **0.49**, 9) | (1.22, 0.49, 0) |
| 19 | 5 | 5 | 0.1 | 0.5 | 0.1 | (1.47, 0.38, 0) | (**1.18**, **0.51**, 0) | (1.18, 0.51, 0) | (1.19, 0.50, 0) |
| 20 | 10 | 5 | 0.1 | 0.5 | 0.1 | (1.70, 0.29, 0) | (1.15, 0.52, 0) | (**1.15**, **0.52**, 0) | (1.16, 0.51, 0) |
| 21 | 2 | 10 | 0.1 | 0.5 | 0.1 | (1.63, 0.54, 0) | (**1.33**, **0.63**, 8) | 77% failure | (1.36, 0.62, 0) |
| 22 | 3 | 10 | 0.1 | 0.5 | 0.1 | (1.70, 0.53, 0) | (**1.28**, **0.64**, 2) | (1.29, 0.64, 20) | (1.30, 0.64, 0) |
| 23 | 5 | 10 | 0.1 | 0.5 | 0.1 | (1.92, 0.46, 0) | (1.24, 0.65, 0) | (**1.24**, **0.65**, 0) | (1.25, 0.65, 0) |
| 24 | 10 | 10 | 0.1 | 0.5 | 0.1 | (2.49, 0.30, 0) | (1.21, 0.66, 0) | (**1.21**, **0.66**, 0) | (1.23, 0.66, 0) |
| 25 | 2 | 5 | 0.3 | 0.5 | 0.1 | (1.65, 0.32, 0) | (**1.26**, **0.48**, 25) | 91% failure | (1.31, 0.46, 0) |
| 26 | 3 | 5 | 0.3 | 0.5 | 0.1 | (1.74, 0.27, 0) | (**1.24**, **0.48**, 9) | 53% failure | (1.28, 0.46, 0) |
| 27 | 5 | 5 | 0.3 | 0.5 | 0.1 | (1.88, 0.22, 0) | (1.19, 0.50, 0) | (**1.19**, **0.50**, 7) | (1.23, 0.49, 0) |
| 28 | 10 | 5 | 0.3 | 0.5 | 0.1 | (2.00, 0.16, 0) | (1.16, 0.51, 0) | (**1.16**, **0.51**, 0) | (1.21, 0.49, 0) |
| 29 | 2 | 10 | 0.3 | 0.5 | 0.1 | (2.15, 0.41, 0) | (**1.39**, **0.61**, 57) | 100% failure | (1.46, 0.59, 0) |
| 30 | 3 | 10 | 0.3 | 0.5 | 0.1 | (2.51, 0.30, 0) | (**1.40**, **0.61**, 21) | 92% failure | (1.46, 0.59, 0) |
| 31 | 5 | 10 | 0.3 | 0.5 | 0.1 | (2.85, 0.20, 0) | (**1.29**, **0.64**, 1) | 20% failure | (1.37, 0.61, 0) |
| 32 | 10 | 10 | 0.3 | 0.5 | 0.1 | (3.17, 0.12, 0) | (1.24, 0.65, 0) | (**1.23**, **0.66**, 0) | (1.34, 0.63, 0) |
| 33 | 2 | 5 | 0.1 | 0.2 | 0.3 | (1.88, -0.09, 0) | (1.93, -0.12, 5) | 48% failure | (**1.83**, **-0.07**, 0) |
| 34 | 3 | 5 | 0.1 | 0.2 | 0.3 | (1.76, -0.04, 0) | (1.77, -0.05, 1) | (1.76, -0.04, 9) | (**1.74**, **-0.03**, 0) |
| 35 | 5 | 5 | 0.1 | 0.2 | 0.3 | (1.71, -0.01, 0) | (1.68, 0.00, 0) | (1.68, 0.00, 0) | (**1.65**, **0.02**, 0) |
| 36 | 10 | 5 | 0.1 | 0.2 | 0.3 | (1.67, 0.02, 0) | (1.60, 0.06, 0) | (1.60, 0.06, 0) | (**1.59**, **0.06**, 0) |
| 37 | 2 | 10 | 0.1 | 0.2 | 0.3 | (2.39, -0.09, 0) | (2.43, -0.11, 8) | 77% failure | (**2.31**, **-0.06**, 0) |
| 38 | 3 | 10 | 0.1 | 0.2 | 0.3 | (2.29, -0.02, 0) | (2.32, -0.03, 2) | 21% failure | (**2.20**, **0.02**, 0) |
| 39 | 5 | 10 | 0.1 | 0.2 | 0.3 | (2.15, 0.03, 0) | (2.10, 0.05, 0) | (2.10, 0.05, 0) | (**2.04**, **0.07**, 0) |
| 40 | 10 | 10 | 0.1 | 0.2 | 0.3 | (2.19, 0.04, 0) | (2.01, 0.11, 0) | (2.01, 0.11, 0) | (**2.00**, **0.12**, 0) |
| 41 | 2 | 5 | 0.3 | 0.2 | 0.3 | (1.90, -0.10, 0) | (1.93, -0.13, 25) | 91% failure | (**1.81**, **-0.05**, 0) |
| 42 | 3 | 5 | 0.3 | 0.2 | 0.3 | (1.78, -0.05, 0) | (1.81, -0.07, 9) | 54% failure | (**1.73**, **-0.02**, 0) |
| 43 | 5 | 5 | 0.3 | 0.2 | 0.3 | (1.74, -0.03, 0) | (1.69, -0.00, 0) | (1.70, -0.00, 8) | (**1.64**, **0.03**, 0) |
| 44 | 10 | 5 | 0.3 | 0.2 | 0.3 | (1.71, 0.00, 0) | (1.62, 0.05, 0) | (1.62, 0.05, 0) | (**1.60**, **0.06**, 0) |
| 45 | 2 | 10 | 0.3 | 0.2 | 0.3 | (2.39, -0.08, 0) | (2.49, -0.12, 57) | 100% failure | (**2.29**, **-0.03**, 0) |
| 46 | 3 | 10 | 0.3 | 0.2 | 0.3 | (2.36, -0.05, 0) | (2.43, -0.08, 21) | 88% failure | (**2.19**, **0.03**, 0) |
| 47 | 5 | 10 | 0.3 | 0.2 | 0.3 | (2.24, -0.01, 0) | (2.18, 0.01, 1) | 21% failure | (**2.04**, **0.08**, 0) |
| 48 | 10 | 10 | 0.3 | 0.2 | 0.3 | (2.29, -0.01, 0) | (2.04, 0.10, 0) | (2.04, 0.10, 0) | (**2.01**, **0.11**, 0) |
| 49 | 2 | 5 | 0.1 | 0.5 | 0.3 | (1.99, 0.29, 0) | (1.92, 0.32, 5) | 48% failure | (**1.86**, **0.34**, 0) |
| 50 | 3 | 5 | 0.1 | 0.5 | 0.3 | (1.93, 0.29, 0) | (1.77, 0.35, 1) | (1.77, 0.35, 9) | (**1.75**, **0.36**, 0) |
| 51 | 5 | 5 | 0.1 | 0.5 | 0.3 | (1.97, 0.28, 0) | (1.68, 0.39, 0) | (1.68, 0.38, 0) | (**1.67**, **0.39**, 0) |
| 52 | 10 | 5 | 0.1 | 0.5 | 0.3 | (2.15, 0.22, 0) | (1.60, 0.42, 0) | (**1.60**, **0.42**, 0) | (1.60, 0.42, 0) |
| 53 | 2 | 10 | 0.1 | 0.5 | 0.3 | (2.70, 0.36, 0) | (2.42, 0.42, 8) | 77% failure | (**2.38**, **0.43**, 0) |
| 54 | 3 | 10 | 0.1 | 0.5 | 0.3 | (2.68, 0.38, 0) | (2.33, 0.46, 2) | 21% failure | (**2.25**, **0.48**, 0) |
| 55 | 5 | 10 | 0.1 | 0.5 | 0.3 | (2.76, 0.35, 0) | (2.11, 0.50, 0) | (2.11, 0.50, 0) | (**2.07**, **0.51**, 0) |
| 56 | 10 | 10 | 0.1 | 0.5 | 0.3 | (3.30, 0.24, 0) | (2.02, 0.53, 0) | (**2.01**, **0.53**, 0) | (2.02, 0.53, 0) |
| 57 | 2 | 5 | 0.3 | 0.5 | 0.3 | (1.99, 0.00, 0) | (1.93, 0.02, 25) | 91% failure | (**1.83**, **0.08**, 0) |
| 58 | 3 | 5 | 0.3 | 0.5 | 0.3 | (1.89, 0.03, 0) | (1.81, 0.06, 9) | 54% failure | (**1.75**, **0.09**, 0) |
| 59 | 5 | 5 | 0.3 | 0.5 | 0.3 | (2.35, 0.15, 0) | (1.70, 0.38, 0) | (**1.70**, 0.38, 7) | (1.70, **0.38**, 0) |
| 60 | 10 | 5 | 0.3 | 0.5 | 0.3 | (2.43, 0.12, 0) | (1.62, **0.41**, 0) | (**1.62**, 0.41, 0) | (1.66, 0.40, 0) |
| 61 | 2 | 10 | 0.3 | 0.5 | 0.3 | (3.12, 0.27, 0) | (2.46, 0.42, 57) | 100% failure | (**2.44**, **0.43**, 0) |
| 62 | 3 | 10 | 0.3 | 0.5 | 0.3 | (3.41, 0.21, 0) | (2.48, 0.42, 21) | 89% failure | (**2.36**, **0.45**, 0) |
| 63 | 5 | 10 | 0.3 | 0.5 | 0.3 | (3.63, 0.15, 0) | (2.20, 0.48, 1) | (**2.16**, 0.49, 19) | (2.16, **0.49**, 0) |
| 64 | 10 | 10 | 0.3 | 0.5 | 0.3 | (3.98, 0.08, 0) | (2.06, 0.52, 0) | (**2.05**, **0.52**, 0) | (2.13, 0.51, 0) |