**Data S3.** Combined Inference after Multiple Imputation.

For both imputation approaches, results were pooled from the 50 imputed datasets as follows. In this study, $β$ is the regression coefficient of a linear regression model (or linear mixed model). Suppose we have imputed $m$ complete datasets (in this case, $m=50$). Let $\hat{β}^{(k)}$ and $\hat{σ}^{2(k)}$ denote the point estimate and variance from the $k$th dataset ($k=1,…, m$). Then the multiple imputation point estimate of $β$ is the average of the $m$ complete-data estimates:
$$\overbar{β}=\frac{1}{m}\sum\_{k=1}^{m}\hat{β}^{(k)}$$

The SE of the multiple imputed point estimate $\overbar{β}$ consists of two components: within-imputation variance and between-imputation variance. The within-imputation component is specified as:
$$W=\frac{1}{m}\sum\_{k=1}^{m}\hat{σ}^{2(k)}$$

The between-imputation component is specified as:

$$B=\frac{1}{m-1}\sum\_{k=1}^{m}\left(\hat{β}^{(k)}-\overbar{β}\right)^{2}$$

The total variance of the pooled estimate of $β$ is then given by:

$$V=W+\left(1+\frac{1}{m}\right)B$$

Inference (hypothesis testing and confidence interval estimation) is based on $\sqrt{V}\left(β-\overbar{β}\right)∼t\_{λ}$, where the degrees of freedom $λ=\left(m-1\right)\left\{1+\frac{W}{\left(1+1/m\right)B}\right\}^{2}$.1

REFERENCE

1. Little RJ and Rubin DB. *Statistical Analysis with Missing Data (2nd edition).* John Wiley & Sons; 2002.