**Impute Data with Multiple Imputation by Chained Equations (MICE)**

We then impute the remaining missing data using Multiple (or also called Multivariate) Imputation by Chained Equations (MICE), where Random Forest approach, a machine learning technique, is used to impute each variable. We consider this an appropriate method to impute missing data in our case as it needs not assume data are missing at random which is otherwise required for parametric MICE. It also accommodates nonlinearity and interaction between variables and yields estimates with no material bias with better confidence intervals.[[1]](#footnote-1)

1. Anoop D. Shah, Jonathan W. Bartlett, James Carpenter, Owen Nicholas, and Harry Hemingway. “Comparison of Random Forest and Parametric Imputation Models for Imputing Missing Data Using MICE: A CALIBER Study”, Am J Epidemiol. 2014 Mar 15; 179(6): 764–774. [↑](#footnote-ref-1)