

## Course Title: Multiple Imputation Analysis for Missing Data in Practice

March 1-3, 2023

**Learning Objectives:** By the end of the course, participants are expected to learn some fundamental concepts of missing data/nonresponse problems and multiple imputation. They are expected to know the great potential of multiple imputation in handling missing data in different study and research contexts. They are expected to know some basic multiple imputation programs in major software packages (e.g., SAS, R, and Python) as well as key strategies in constructing appropriate imputation models for data analysis.

**Content and Instructional Methods:** The course will have slides for the content. However, multiple real data examples are to be presented throughout the course. The programming code for examples are to be demonstrated to the audience. In addition, key references in missing data and multiple imputation analysis will be provided. All registered participants will be contacted by the instructors prior to and after the event with additional course information (e.g., Virtual Classroom link and programming code).

**Instructor Background:** Drs. Yulei He and Guangyu Zhang are mathematical statisticians at the National Center for Health Statistics, the U.S. Centers for Disease Control and Prevention. Dr. Chiu-Hsieh Hsu is a Professor of Biostatistics at the University of Arizona. All authors have researched, taught, and consulted in multiple imputation and missing data analysis in the past 20 years. The course material is related to and also expanded from their recent book “Multiple Imputation of Missing Data in Practice: Basic Theory and Analysis Strategies”  
<https://www.routledge.com/Multiple-Imputation-of-Missing-Data-in-Practice-Basic-Theory-and-Analysis/He-Zhang-Hsu/p/book/9781498722063#>

**Registration** is limited to 25 persons with the cost of registration being \$25 for Full-time Students, \$50 for current WSS members, and \$100 for non-members.

Interested persons should visit <https://www.eventbrite.com/e/multiple-imputation-analysis-for-missing-data-in-practice-tickets-520786445277> for more information about the content and how to register. If you have any questions, please contact Yan Li ([yli6@umd.edu](mailto:yli6@umd.edu)).

### Outline:

*Day 1: 11am-1pm EST on March 1 (Wed.), 2023*

Part 1: Basic concepts of missing data problems - Topics include missing data definition, missing data pattern, missing data mechanisms (MCAR/MAR/MNAR), and some principled as well as ad-hoc missing data analysis methods (e.g., complete-case analysis).

Part 2: Basic ideas of multiple imputation - Topics include the motivation of multiple imputation, its procedure, and the combining rules, and multiple imputation models for univariate missing data.

*Day 2: 11am-1pm EST on March 2 (Thur.), 2023*

Part 1: Multiple imputation for multivariate missing data – Topics include imputation using the joint modeling approach and fully conditional specification approach; Multiple imputation in some specific data and design settings (e.g., for longitudinal data, survival data, survey data, and measurement error models).

Part 2: Sample code and implementation using SAS/R/Python

*Day 3: 11am-1pm EST on March 3 (Fri.), 2023*

Part 1: Two real data applications: (1) Multiple imputation for survival analysis; (2) Multiple imputation sensitivity analysis for missing data not at random.

Part 2: A case study: Multiple imputation for handling CDC COVID-19 case reports with missing race/ethnicity.

Part 3: Floor discussion and wrap-up.