

THE UNIVERSITY OF OKLAHOMA



Biostatistics, Epidemiology, and Research Design
Key Component Activity

OSCTR Biostatistics, Epidemiology and Research Design Core

WORKSHOP

INTRODUCTION TO DATA INTEGRATION FOR COMBINING PROBABILITY AND NONPROBABILITY SAMPLES

Sixia Chen, Ph.D.

Associate Professor of Biostatistics
University of Oklahoma Health Sciences Center

**Friday, September 29, 2023
12:00PM – 2:00PM**

Location: Hudson College of Public Health Auditorium (CHB 220)
and Virtual through Zoom.

Box lunches will be provided for the first 25 attendees.

Registration required in advance for this meeting.

**After registration, you will receive a confirmation email containing the zoom link
and information about the workshop materials.**

SCHEDULE: The workshop will be held at September 29th from 12pm to 2pm, 2023. The workshop will be a mixture of lecture and in-class hands-on practice.

SOFTWARE: Prior to the workshop, please install SAS and R on your laptop. Please bring your laptop to the workshop.

PREREQUISITE: Pre-requisite requirements include basic knowledge of SAS and R programming and statistical methods such as linear models, descriptive statistics, two sample t tests, ANOVA, chi-square tests, and logistic regression models. Study the material for my previous workshop 'Introduction to Complex Survey Data Analysis' at <https://osctr.ouhsc.edu/short-course>

DESCRIPTION: Nonprobability samples have been used frequently in practice including education, medical study, and public opinion research. Due to selection bias, naïve estimates without adjustments by using nonprobability samples may lead to misleading results. In this course, we will discuss some commonly used data integration methods for reducing the selection bias of nonprobability samples.

WORKSHOP CONTENT: 1. Introduction to probability samples, nonprobability samples, and their applications in practice; 2. Calibration weighting approach; 3. Propensity score weighting approaches; 4. Mass imputation approaches. For each of the previous topics, we will provide hands on exercises by using some real data applications.

Course materials can be downloaded or printed for personal use prior to attendance from the following website <https://osctr.ouhsc.edu/short-course>

END of WORKSHOP EVALUATION SURVEY: Please complete the survey at the following link:
<https://bbmc.ouhsc.edu/redcap/surveys/?s=KFRD3NDKKWLD9JME>. You will also receive the link by email after the workshop.

REGISTRATION: Registration is required by September 27th at 6pm. Registration can be completed by clicking the hyperlink in the first page or at the following website <https://osctr.ouhsc.edu/short-course>.

SPONSOR ACKNOWLEDGEMENT: Sponsored by the Clinical Epidemiology Unit of the Biostatistics and Epidemiology Research Design (BERD) Core of the Oklahoma Shared Clinical and Translational Resources, National Institutes of Health, National Institute of General Medical Sciences Grant U54GM104938

PARKING: The College of Public Health Building is located on the corner of 13th

Street and Phillips Avenue. Parking is available on the north side of the building. If you are driving north on Phillips Ave, you will see a sign for Lot A on the east side of the street. Pull into this lot. You will need to press the emergency button on the speaker box and indicate that you are attending a workshop in the College of Public Health Building for the parking gate to be raised.

FACULTY BIOGRAPHICAL SUMMARY: Dr. Chen has more than 15 years practical and theoretical research experience with survey sampling and missing data analysis problems. Before joining OUHSC, Dr. Chen served as the senior sampling statistician at Westat research organization (a world leading research organization in survey sampling and methods research) from 2012 to 2015. Dr. Chen has served as PI, Co-PI and Co-I for multiple local and national grants including NIH. He has extensive inter-/multi-disciplinary research experience in data integration, missing data analysis, survey sampling, biostatistics, tobacco research, and Native American health disparity. Recently, Dr. Chen has started to work on machine learning and big data research with three funded NIH grants (two as the PI) related to machine learning. He has more than 37 independent methodological publications in top statistical journals (31 are first/senior author articles) and more than 48 collaborative publications in other fields. Dr. Chen is currently an elected member in International Statistical Institute. He is now serving as associate editor for the following four well-known peer reviewed journals: Journal of Korean Statistical Society, Scandinavian Journal of Statistics, Journal of Survey Statistics and Methodology (Top 2 journals in survey statistics and methodology in the world), Survey Methodology (Top journal in survey research), and director of Novel Methodologies Unit of Biostatistics, Epidemiology, and Research Design Core of the Oklahoma Shared Clinical and Translational Resources (OSCTR). In addition, Dr. Chen is now serving as the president of American Statistical Association Oklahoma Chapter. Dr. Chen is a Faculty Affiliate of the TSET Health Promotion Research Center. He performs as faculty biostatistician for Biomedical and Behavioral Methodology Core, Biostatistics Core for Stephenson Cancer Center and Sooner Survey Center at OUHSC. He has also served as a consultant for the Tribal Epidemiology Center of the Southern Plains Tribal Health Board of Oklahoma since 2016.