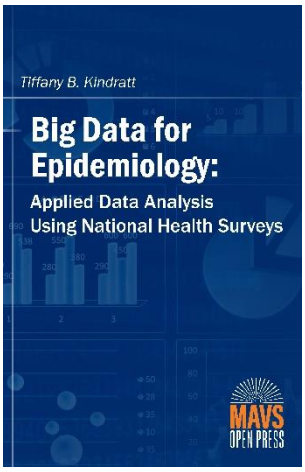
Chapter 1: Introduction
Slides Transcript



Slide 1

Hi everyone, my name is Tiffany Kindratt. This presentation covers “Chapter 1, Introduction,” in the textbook Big Data for Epidemiology: Applied Data Analysis using National Health Surveys.

Slide 2

Here is an outline of the chapter. In this chapter, I will review the purpose of the textbook in Section 1.1., the health services research focus of the textbook in Section 1.2, outline the chapters in Section 1.3, provide a summary in Section 1.4, discuss COVID-19 pandemic changes in Section 1.5 and list references used to write the chapter in Section 1.6.

Slide 3

Section 1.1. Purpose. The purpose of this textbook is to train future public health professionals, particularly Master of Public Health (or MPH) students, how to conduct applied data analysis using secondary data collected from nationally representative health surveys. The textbook is meant to help eliminate gaps in knowledge, skills and analytical abilities that may prohibit MPH student graduates from being successful in entry level public health practice and research focused positions. A recent study of local health departments identified that several entry-level public health professionals lacked knowledge, skills and abilities for data collection, data management, data cleaning, quantitative data analysis and statistics, as well as specifically conducting data analysis using SAS statistical software, which is a software commonly used within the public health practice setting.

Slide 4

Section 1.2. Health Services Research Focus. The textbook was developed based on previous research studies I have been involved in, as well as the research focus of my current health survey research lab at the University of Texas at Arlington. The lab was established in fall of 2019 in the Public Health Program, Department of Kinesiology, and College of Nursing and Health Innovation at the University of Texas at Arlington. The goal of the lab is to conduct epidemiologic research studies focused on evaluating predisposing and enabling factors that influence individuals’ morbidity, mortality and use of health services across the life course with big data methodologies.

Kindratt, Tiffany B (2022). Chapter 1. Introduction. In: *Big Data for Epidemiology: Applied Data Analysis Using National Health Surveys.* UTA PressBooks. Mavs OpenPress. In Press. Available at: https://hsrlab.uta.edu/big-data-for-epidemiology/. Transcript updated January 17, 2022.

Slide 5

Studies conducted in this lab have used multiple large data sources, including but not limited to the Medical Expenditure Panel Survey or MEPS, the National Health Interview Survey or NHIS, the Health Information National Trends Survey or HINTS. American Community Survey or ACS, as well as studies using state and local health department data sources. The research that I design and lead typically involves the analysis of secondary data sources, but I also collaborate and partner with other individuals doing primary data collection and collaborations in medical education settings and in the community. The lab website is provided on this slide if you are interested in learning more about current research projects.

Slide 6

Here is an overview of the theoretical model for the Health Survey Research Lab based on Anderson's Behavioral Model of Health Services. Anderson's model posits that health services use is influenced by predisposing, enabling and need factors. Predisposing factors include factors like demographics, biology, or social structure factors that cannot be changed. Enabling factors include community or personal resources that are available to people that may enable or disable them to receive health services. Finally, examples of perceived and evaluated needs include whether or not someone has been diagnosed with the disease in the past that that may influence their need to receive health services. Within this model, my research focuses on predisposing and enabling factors as the main predictors or independent variables. The predisposing factors are race and ethnicity. In the studies that I lead, I collaborate with a group of individuals that focus on research, looking at disaggregating Arab Americans, broadly including individuals who were born in or trace their heritage to the Middle East or North Africa, from the non-Hispanic white racial and ethnic group. Arab Americans, or MENA individuals – which is the term used to describe Middle Eastern or North Africans -- are classified as white by the US federal government. This limits their ability to receive dedicated funding or resources that has been allocated for reducing health disparities. Currently, they are not considered an underrepresented minority group in the US. It also masks potential health disparities that may exist that are higher than non-Hispanic whites and may even be higher than non-Hispanic black or African American individuals or Hispanic or Latino individuals. So, that's the main predisposing factor that I look at in my research. I'm also involved in some other studies looking at Hispanics and non-Hispanic, black or African American individuals, and particularly by nativity status. Nativity status describes whether or not someone is born in the US versus someone who is foreign-born. The enabling factors evaluated in my lab include patient experiences. Within Anderson's model, there are no enabling factors listed that are focused on health care professionals. My research recognizes the impacts that health care professionals can have on individual's health outcomes. Health care professionals can really contribute to someone's likelihood of receiving recommended screenings and providing advice for health behavior change. The patient experiences evaluated within this model include those interactions with the provider, which can be evaluated as

Kindratt, Tiffany B (2022). Chapter 1. Introduction. In: *Big Data for Epidemiology: Applied Data Analysis Using National Health Surveys.* UTA PressBooks. Mavs OpenPress. In Press. Available at: https://hsrlab.uta.edu/big-data-for-epidemiology/. Transcript updated January 17, 2022.

independent or predictor variables or also as an outcome. Examples include patient-provider communication quality, electronic patient-provider communication through patient portals or using emails, shared decision-making, and patient-provider concordance. So, for patient-provider concordance, looking at whether or not people are more comfortable with a health care provider who is the same gender or race and ethnicity. The dependent variables or outcomes related to morbidity, mortality and health services use in some current studies, include a focus on cognitive health outcomes, including cognitive impairment and decline that often precedes Alzheimer disease and related dementias or ADRD, a wide range of maternal and child health topics including low birth weight among infants, health behaviors during pregnancy, and COVID-19 cases and deaths among Arab American individuals. More details on research studies conducted in the Health Survey Research Lab are provided on the website.

Slide 7

Section 1.3 Outline of Textbook Chapters. This textbook has four sections. The first section is the introduction to national health surveys. The second section focuses on basic applied data analysis. The third section describes common national health surveys and includes case studies that you will practice your basic analytic skills with each of the different surveys. The fourth section includes details on how to disseminate results of research studies and includes a chapter to conclude the textbook.

Slide 8

Section 1.3.1. Textbook Section 1: Introduction to National Health Surveys. Section 1 in the textbook will provide an introduction to national health surveys. Chapter 1, which I am discussing right now, is the introduction. Chapter 2 provides some basics of all of the different national health surveys that are covered in the textbook, including the National Health Interview Survey or NHIS, the Medical Expenditure Panel Survey or MEPS, the Health Information National Trends Survey or HINTS, Behavioral Risk Factor Surveillance System or BRFSS, and the National Health and Nutrition Examination Survey or NHANES. Chapter 3 will provide a literature review related to the case studies that are within the Section 3 of the textbook. So, you should get some sort of a background feel of why I used those examples based on my former, current, and future research interests and trajectory.

Slide 9

Section 1.3.2. Textbook Section 2: Basic Applied Data Analysis. In Section 2 of the textbook, there are two chapters that are meant to be a review of what you've learned about basic data analysis before and give an overview of the complex survey design features that you will be learning for analyzing national data sources. Chapter 4 includes basic information on how to conduct frequencies, percentages, means, and standard deviations using SAS statistical software. Chapter 5 provides more details of special survey procedures, with clustering, stratum, and weighting variables required for conducting analysis with complex samples.

Kindratt, Tiffany B (2022). Chapter 1. Introduction. In: *Big Data for Epidemiology: Applied Data Analysis Using National Health Surveys.* UTA PressBooks. Mavs OpenPress. In Press. Available at: https://hsrlab.uta.edu/big-data-for-epidemiology/. Transcript updated January 17, 2022.

Slide 10

Section 1.3.3. Textbook Section 3: Common National Health Surveys. This section is divided into five chapters. Each chapter is dedicated to one national health survey. Each chapter includes a general overview of the survey and its purpose. It includes details on what types of data files are available. What are some of the advantages and disadvantages of using that survey? It includes some practical tips for conducting the analysis for the research case study that will be used for students to really apply what they are learning in on how to analyze national data.

Slide 11

Here is an outline of the chapters. Chapter 6 covers the National Health Interview Survey. Chapter 7 covers the Medical Expenditure Panel Survey. Chapter 8 covers the Health Information National Trends Survey. Chapter 9 covers the Behavioral Risk Factor Surveillance System. Chapter 10 covers the National Health and Nutrition Examination Survey.

Slide 12

In Chapter 6, covering the National Health Interview Survey, the data that will be analyzed for the research case study will include 2018 NHIS person and sample adult files. The data will be analyzed to explore whether foreign-born Arab American adults are more or less likely to receive an annual flu vaccine in comparison to other racial and ethnic groups including non-Hispanic whites, non-Hispanic blacks or African Americans, and Hispanics or Latinos.

Slide 13

In Chapter 7, data from 2015 and 2016 Medical Expenditure Panel Survey household level in-person interviews and self-administered questionnaires will be analyzed. The research case study objective in this chapter is to determine whether adults who perceived their health care provider exhibited quality patient-provider communication during their visits over the last 12 months are more or less likely to receive an annual flu vaccine compared to those who did not receive quality patient-provider communication.

Slide 14

In Chapter 8, HINTS 5 Cycle 2 data will be analyzed to determine associations between electronic patient-provider communication and breast cancer screening uptake among women.

Slide 15

In Chapter 9, state level data from the 2019 BRFSS will be analyzed. The objective of the research case study is to explore how differences in caregiving experiences among urban and rural adults in Texas are moderated by race and ethnicity. Meaning that, in this research case study, we will use data from the optional caregiver module to explore

Kindratt, Tiffany B (2022). Chapter 1. Introduction. In: *Big Data for Epidemiology: Applied Data Analysis Using National Health Surveys.* UTA PressBooks. Mavs OpenPress. In Press. Available at: https://hsrlab.uta.edu/big-data-for-epidemiology/. Transcript updated January 17, 2022.

the differences in caregiving experiences among urban and rural residents and then stratify or separate our findings by race and ethnicity.

Slide 16

In Chapter 10, self-reported and examination data from 2011-2012 and 2013-2014 data files will be analyzed to explore whether sociodemographic factors and subjective cognitive decline are associated with cognitive functioning performance among older adults.

Slide 17

Section 1.3.4. Textbook Section 4: Dissemination and Conclusions. The fourth, and last section of the textbook has a chapter focused on disseminating the results of research studies and a chapter that just summarizes everything that was covered in the textbook. Chapter 11 includes content on how to present the results of the research case studies in an abstract, as well as a full-length manuscript format. Chapter 12 summarizes what has been covered in the textbook and includes some recommendations for future versions.

Slide 18

Section 1.4. Summary. So, in summary, this textbook provides instructions on how to conduct some basic applied data analysis using secondary data collected from national health surveys. The textbook was developed based on a previous course that I developed as part of my PhD training. This was an independent study in epidemiology course at the UT Health, School of Public Health at Dallas Regional Campus. Within that course, I created teaching tools for three different surveys in order for me to learn more applied data analysis skills that I was not receiving within the recommended coursework at that time. I created a teaching tool for three surveys, including MEPS, BRFSS and NHANES. I've provided these teaching tools in the textbook in Appendix 1. The teaching tools include examples are a little bit older than those in this textbook but can provide students will some other ways to look at how to analyze these data. The outcomes are different, including celiac disease and immunizations during pregnancy. One of the teaching tools also provides syntax using STATA instead of SAS.

Slide 19

Here is an example of one of the teaching tools that is provided in the appendix. It provides a brief overview of the survey. It has a description of the data files, the advantages and disadvantages, practical tips, research questions used for the research case study, and some simple examples of coding.

Slide 20

Section 1.5. COVID-19 Pandemic Changes. The initial version of this textbook was written from June through December 2020 during the early waves of the COVID-19 pandemic. The methods described for the national surveys in this textbook represent

Kindratt, Tiffany B (2022). Chapter 1. Introduction. In: *Big Data for Epidemiology: Applied Data Analysis Using National Health Surveys.* UTA PressBooks. Mavs OpenPress. In Press. Available at: https://hsrlab.uta.edu/big-data-for-epidemiology/. Transcript updated January 17, 2022.

“pre-pandemic” methodologies. Many surveillance systems and surveys had to be modified due to safety concerns, stay-at-home orders, and data collection needs from 2020 onward (Lau et al. 2021). Chapters 6 through 10 include a brief section that discusses these changes for the respective survey.

Slide 21

Section 1.6. References. Here are some of the references that were used in writing this chapter of the textbook.

Slide 22

That concludes my presentation on “Chapter 1: Introduction,” in the textbook Big Data for Epidemiology: Applied Data Analysis Using National Health Surveys. Thank you so much for listening to this presentation. I hope you found it helpful. Thank you.

Kindratt, Tiffany B (2022). Chapter 1. Introduction. In: *Big Data for Epidemiology: Applied Data Analysis Using National Health Surveys.* UTA PressBooks. Mavs OpenPress. In Press. Available at: https://hsrlab.uta.edu/big-data-for-epidemiology/. Transcript updated January 17, 2022.