Advanced Methods in Health Services Research
PH 147 / QBS 139
Spring Term 2019

Course co-Directors:

Tracy Onega, PhD, MA, MS; Tracy.L.Onega@dartmouth.edu
Division of Biomedical Informatics
Associate Professor, Dept. Biomedical Data Science, Dept. of Epidemiology, TDI, and, Norris Cotton Cancer Center
Office: Room 858 Rubin Bldg. Norris Cotton Cancer Center
Office hours (in classroom) 8:45 – 9 am, 12-12:30, or other times by appt.

Erika Moen, PhD: Erika.L.Moen@dartmouth.edu
Assistant Professor, Dept. Biomedical Data Science
Office: Room 327 Williamson Translational Research Building
Office hours (in classroom) 8:45 – 9 am, 12-12:30, or other times by appt.

Guest Faculty Lecturers:

Thérèse Stukel, PhD, Professor TDI, Professor of Health Policy, Management and Evaluation, University of Toronto, Senior Scientist Institute for Evaluative Clinical Sciences (ICES), Toronto, Ontario

Teaching assistants:

Andrew (Drew) Bohm, MS, PhD Student, TDI Andrew.R.Bohm.gr@dartmouth.edu
Office hours (in classroom) 8:45-9:00 am, 12-12:30, and by appointment

Date, time, and place:

Spring Term March 25 – May 31, 2019
Mondays 9:00 – noon; 1:00– 4:30;
except Friday, May 31 (9:00-12:00)

Prerequisites:

(TDI) PH 140, 141 (pass or high pass required)

Overall course Aims: This course will develop student analytic competencies to the level necessary to conceptualize, plan, carry out, and effectively communicate small research projects in patient care, epidemiology, or health services. Lectures, demonstrations, and labs will be used to integrate and extend methods introduced in other TDI courses. The course will also cover new methods in epidemiology, health services and data science. The students will use national publicly available data and synthetic research datasets resembling Medicare claims and electronic health record data in classroom lab exercises and course assignments. Course topics focus on key aspects observational research including cohort derivation, multilevel analyses, small area analysis, and network analysis. Practical skill areas will include programming in STATA and/or R, developing an analytic workflow, data visualization (designing tables and figures), and data structure and management. Emphasis is on becoming independent in research processes. The instructors will mentor students as they develop their own analytic projects. The main goal of the course is to firmly ground you in the scientific ***process*** of observational research.

LEARNING OBJECTIVES [LINKED TO TDI CORE COMPETENCIES]	COMPETENCIES TO DEMONSTRATE
1. To develop an understanding of advanced methods commonly used in health services research and to be an informed end-user of some of these methods. [5,7]	For each method: Understand when the method is useful, how it is used, its data requirements, and limitations of its application. To know when regression, propensity scores, instrumental variables and multilevel modeling are applicable
2. To develop skills in defining, planning, and conducting a small research project. [5,7]	Define a research question within the context of current theories and past analyses. Identify datasets used in health services research; create Stata datasets from source files, and analyze the contents for data quality. Operationalize the research question into a feasible research plan. Conduct the appropriate analyses to partially or completely answer the research question Apply newly learned Stata commands to conduct analyses. Understand the components of data analysis
3. To enhance skills needed in interpreting and communicating a	Apply the methods of technical writing used in health services research. Develop efficient and understandable tables and figures.

research project. [5,7,9]	Interpret analytic results, communicate the results to peers, and identify limitations.

TDI Core Competencies

1. Understand the contributors to health and disease
2. Understand the relationships between health, health care, and public policy
3. Understand the organizations, dynamics, delivery, impact, and financing of health care
4. Understand decision-making and variation in health care
5. Apply and understand statistics, research methods, and quality improvement approaches and measurement of individual, community, and population health
6. Critically evaluate health and health care information from diverse sources with a healthy skepticism by questioning assumptions and considering alternative explanations and conclusions
7. Apply diverse problem-solving skills including the ability to define a question, access appropriate data, apply analytic tools, and combine creativity and systems thinking to generate informed assessments and feasible recommendations
8. Effectively engage stakeholders, manage change, monitor processes, and evaluate outcomes to improve health
9. Communicate professionally, effectively, and persuasively in written, spoken, and visual formats
10. Exercise the skill of listening. Seek, accept, and provide constructive feedback
11. Contribute individually, collaborate effectively, and promote a productive work environment
12. Convey respect for others. Incorporate ethics in decision making, management, and behavior
13. Lead improvement by organizing teams and motivating others with an inspiring, realistic vision

Course requirements (% of final grade):

- **Participation (10%)** Attendance and active participation in class. This requires reading and/or watching the assigned materials before class and discussion during the components of the morning sessions, and active participation in group planning portions of the labs.
- **Eight labs (40%)** To be handed in weekly per course schedule
- **Analytic memo & conceptual model (10%)** Due WEEK 4
- **Research roadmap (5%)** Due WEEK 9
- **Student course project (35%)** This is the "final project" for the course and will be comprised of several graded elements discussed below.

Grades: Students will receive a single grade for the course, using Dartmouth's standard grading for graduate level courses:

HP - High Pass, indicating work of high quality, which is distinctly superior to that normally expected of a graduate student. Denotes high standing and satisfactory progress

P - Pass, indicating work of good quality, worthy of graduate credit.

LP - Low Pass, indicating work which is acceptable for graduate credit, but in which the student exhibited one or more serious deficiencies.

NC - No Credit, indicating work, which is not acceptable for graduate credit.

Participation: Inherent in the "real world" of research is a constant exchange and interplay of ideas, interpretations, questions, etc. No one can function successfully without engaging in this exchange. We want the course to provide somewhat of a microcosm of the research world, and would like this course to be an opportunity to "practice" sharing ideas, refining arguments, exploring questions, etc. Because we view this as an important aspect of the course and of your education, we will be grading this element of your 'work' in the course.

Lab assignments: These are due by midnight the *Thursday* following the lab and will be returned to you the following Monday. Labs are graded on a 5-point scale based on the following aspects: 1.analytic approach and workflow summary, 2. do file documentation, and 3. lab interpretation and write up. Late labs (up to 24 hours) are strongly discouraged, but please let Tracy, Erika, or Drew know if this becomes necessary prior to the due time. More than 2 late labs will be marked down by a minimum of 2 points. This is to help keep the temporal aspect of learning and feedback on track. Please let Tracy, Erika, and Drew know if you are having persistent difficulty in completing the labs on time. Lab write-ups are individual assignments. Helping out your fellow student is fine, but it is expected that the materials you turn in are your own work.

Unless otherwise indicated, lab assignments require that you submit through Canvas five parts bundled together in **one** pdf file: 1. A report of your analyses in a JAMA abstract format (300 words), 2. Responses to the lab questions/write-up if applicable for that week, 3. Your Stata do file or R code, 4. Your Stata log file, if using a Stata do file, and 5. Your workflow

summary figure. Please remember to submit both the do **and** the log file. For requirements and suggestions regarding do files, see the "Do Files & Workflow Documentation" document on Canvas. Please submit all completed lab assignments as a pdf.

The format used by JAMA for abstracts can be found at:
http://jama.ama-assn.org/misc/ifora.dtl#AbstractsforReportsofOriginalData

Analytic memo and conceptual model This will be developed as a prelude to your Student Course Project, and will follow a format that we will go over in the course. This assignment will allow you to define, refine, and "map out" your course project --- crucial elements in a sound and efficient workflow! It also provides a glimpse into how scientific teams really operate.

Research roadmap Students will develop a 'research roadmap'. There really is a process to research – and although there are many roads leading to the same end - understanding and using a guiding framework for the entire research process can be the difference between success and failure. This is particularly true in our very time-constrained lives. Researchers, physicians, and others often only have a fraction of the time available for research (overall, or on a given project) that they would like to have. Research projects that get bogged down in inefficiency, lack of clarity, false-starts, or multiple re-visitations, often fall by the wayside and never come to fruition. I want you to leave this course with a solid process map for a high quality and efficient research process. This in-class assignment is intended to promote serious thought and consideration to such a research roadmap. As in most science these days – it's a team effort. So collectively, teams will develop these in class and will present to the other groups. All of the research roadmaps will be shared as course products to benefit from the collective effort. Students may work individually or together on this assignment.

Course Project: The student course project is comprised of three graded portions: 1. Podium/oral presentation. This will be a 10 minute presentation and will be graded based on instructor and peer review scores (10% [of total course grade]); 2.Manuscript (15%); 3. Analyses (10%). The purpose of these separate components is to provide the opportunity to hone the range of skills necessary to carry out and communicate research. A separate document with details about the analytic project will be discussed in Week 2.

Course evaluation: We encourage feedback anytime and ideally every week. Please feel free to speak directly to Tracy or Erika. You may email us, or post to Canvas if you have general comments to which the whole class might want to respond.

E-mail queries: Students occasionally e-mail questions to the course instructors during the week. Often the question and its answer will be of general interest to the class. Complex queries are not appropriate for e-mail; contact one of the instructors or teaching assistant for a meeting.

Classroom details: Class will start right at 9:00 am. Tracy, Erika, and Drew will be available every morning before the course at 8:45 for questions, extra help, or discussion. We will also be happy to come earlier or meet at other times by arrangement. There will be a ten-minute break each hour approximately on the hour (e.g. 10:00 AM to 10:10 AM). We will break at noon for lunch most weeks. For Week 10 (Friday, May 31), we will start at 9:00 am as usual, but will go straight through until 1:00 at the latest (lunch provided). For the 8 labs, they will occur in the afternoon session which will begin at 1 PM and will finish around 4:30 PM. Food and drink are welcome in the classroom.

Computational requirements: A laptop with Stata and/or R is needed for class labs. After the group planning portion of some labs, students should generally work on labs individually to develop problem solving skills, but reciprocally helping your fellow students is appropriate and encouraged.

R is a free and open-source software environment for statistical computing and graphics. Prior to the beginning of the course, all students should download R (https://www.r-project.org) and RStudio (https://www.rstudio.com/products/rstudio/download/). If this is your first time using R, please use the swirl() package to familiarize yourself with the R programming language.

Teaching datasets: Several datasets are used during the course. **The files from dartbox are not public and should not be moved from your laptop or shared with anyone else. At the end of the course, it is expected that you will delete the files from your computer. Please *do* to help us maintain compliance with our Data Use Agreement (DUA) with the Centers for Medicare and Medicaid Services (CMS).** Because of the nature of the course datasets, the course data plan was submitted to and approved by Dartmouth's Institutional Review Board, the Committee on the Protection of Human Subjects. As part of this approval, you will be required to sign a course data use agreement (DUA) at the beginning of the course.

Datasets: (Documentation can be also be accessed through Canvas)

Publicly Available Data:

1. **Medicare Claims Synthetic Public Use Files (SynPUFs)** These data files were created by the Centers for Medicare & Medicaid Services (CMS) for the express purposes of: a) training researches on the use and complexity of claims data; b) allowing entrepreneurs to develop software and applications for claims data; c) support data mining innovations. These data maintain the actual CMS Limited Data Sets (available for tens of thousands of dollars!), but with fewer variables (mostly related to processing claims for payment) -- while being de-identified to safeguard beneficiaries' privacy. You will need to visit the website for more information and to download (which will part of Lab 1 instructions). The SynPUFs cover beneficiaries and claims from 2005-2011, although we'll be working with 2008-2010 files. The data are split into 20 samples due to the sheer size of the data. For each sample, there are 6 files. In this course, we will use only 2 of the samples and 5 of the file types in each of the samples (so you'll download 10 files total). That way the data size should be manageable for everyone's computers. The specifics of the data and files will become VERY clear as part of Lab 1. (record length varies, variable count varies, comma delimited files) http://www.cms.gov/Research-Statistics-Data-and-Systems/Research-Statistics-Data-and-Systems.html
2. **Synthea Synthetic Electronic Health Record Data** These data are derived from an open source, synthetic patient generator that models patient medical histories. The patient's diseases, conditions, and medical care are defined by one or more generic modules informed by clinicians and real-world statistics collected by the CDC, NIH, and other research sources. Read more about Synthea here https://syntheticheath.github.io/synthea/
3. **Physician Compare datasets** CMS provides datasets including general information for individual eligible providers (EP). The downloadable file "Physician Compare National Downloadable File" is organized at the individual EP level. Eps affiliated with multiple group practices are listed on multiple lines for each group PAC ID. Additional datasets include the "Physician Compare 2016 Group Public Reporting – Patient Experience" file which contains the Consumer Assessment of Healthcare Providers and Systems (CAHPS) for Physician Quality Reporting System (PQRS) measure performance rates by groups. https://data.medicare.gov/data/physician-compare
4. **CMS Physician Shared Patient Patterns** CMS provides datasets with a list of pairs of physicians who deliver health services to the same patient within 30 days, 60 days, 90 days, and 180 days. https://www.cms.gov/Regulations-and-Guidance/Legislation/FOIA/Referral-Data-FAQs.html
5. **Dartmouth Atlas Downloadable Data:** http://archive.dartmouthatlas.org/tools/downloads.aspx?tab=39#research_files Specific files will be detailed in Week 8.

TDI-Derived Data:

1. **Medicare Denominator files** gi_den_yr1.txt; gi_den_yr2.txt; gi_den_yr3.txt – The complete records from over 1.5 million Medicare beneficiaries from 14 Hospital Referral Regions brought into an anonymous data set that maintains their geographic relationships. The fictional area is termed "Glover Isle". These files are the "Denominator Files" that includes individual characteristics of beneficiaries, one file for each year. (Year 1 Record count = 1,458,621; 19 variables; fixed length)

2. **Medicare MEDPAR files** gi_medpar_yr1.txt; gi_medpar _yr2.txt; gi_medpar _yr3.txt – The complete discharge records from the above Medicare beneficiaries for three years. (Year 1 record count = 450,478; 42 variables; fixed length)

3. **Glover Isle Contextual files** gi_pop_zcta.txt; gi_pop_pcsa.txt – Population and physician information about the ZIPs and Primary Care Service Areas that constitute Glover Isle. Useful for defining contextual variables in multi-level analyses. (ZIP record count 1,369; 44 variables; fixed length)

4. **Primary Care Service Area (V1) database** – The PCSA data are two files that represent a geographic hierarchy -- a file of characteristic of PCSAs, and a file of characteristics of the ZIP Codes from which the PCSAs were aggregated. Variables of both files include numbers of clinicians, rural health centers, community health clinics, population counts, median household income, and more. (PCSA level file: Record count = 6,102, 290 variables, comma delimited; ZIP Code level file record count = approx 29,868; 259 variables, comma delimited file)

Textbooks:

Given the ubiquitous nature of information these days – the suggested textbook is optional (an oldie but a goodie)

OPTIONAL TEXTBOOK:

The Workflow of Data Analysis Using Stata
Author: J. Scott Long
Publisher: Stata Press
2009
ISBN-10: 1-59718-047-5

ISBN-13: 978-1-59718-047-4
Price: $54

Assigned readings - available on Canvas

Additional Materials on Canvas:
Worth looking at. These are reference materials that may be useful during the course. We'll add to these as the need arises.

Academic Integrity/Honor Code
Upholding the highest standards of academic integrity is a professional obligation. Students are expected to know, abide by, and help preserve the TDI Honor Principle and Academic Code of Conduct, both inside and outside of the classroom. All TDI students receive the Honor Principle and Code of Conduct upon entering the program and must sign an acknowledgment that they agree to abide by them. Please see the instructor, a teaching assistant, or member of the Office of Educational Programs if you have any questions.
Related Resource: TDI provides access to TurnItIn (anti-plagiarism software) that allows students to check their work (and learn from any mistakes) prior to submission. While it is recommended that students use the software for all work, it is especially encouraged for group work in which each student whose name is on the final product is responsible for the integrity of the entire document. A link to TurnItIn is provided on the Canvas home page.

Dartmouth Policies
At Dartmouth, we value integrity, responsibility, and respect for the rights and interests of others, all central to our Principles of Community. We are dedicated to establishing and maintaining a safe and inclusive campus where all have equal access to the educational and employment opportunities Dartmouth offers. We strive to promote an environment of sexual respect, safety, and well-being. In its policies and standards, Dartmouth demonstrates unequivocally that sexual assault, gender-based harassment, domestic violence, dating violence, and stalking are not tolerated in our community.

The Sexual Respect Website (sexual-respect.dartmouth.edu) at Dartmouth provides a wealth of information on your right with regard to sexual respect and resources that are available to all in our community. Please note that, as a faculty member, I am obligated to share disclosures regarding conduct under Title IX with Dartmouth's Title IX Coordinator.

Should you have any questions, please feel free to contact Dartmouth's Title IX Coordinator (Kristi.Clemens@Dartmouth.edu) or the Deputy Title IX Coordinator for Geisel (Leslie.Henderson@Dartmouth.edu) or for Guarini (Gary.Hutchins@Dartmouth.edu).

Disability Services/Accommodations
Students with learning disabilities or other special needs should contact a Course Director at the beginning of the term to arrange for any accommodations that may be necessary. The Course Director will consult with the Associate Director of Education to verify the appropriate level of accommodations.

WEEK 1 March 25

Learning Objectives:

1. **What is meant by workflow of data analysis**
2. **Conceptual frameworks – who, what, where, when, why, how**
3. **Basic tenets of units of observation and analysis**
4. **Basic tenets of data models & analytic documentation**
5. **Data management and data quality**

Readings:
Required

1. Read the information found at the following link and download and read the pdfs listed below:
   https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/SynPUFs/index.html

   a. CMS Data Disclaimer – User Agreement – Public Use Data (pdf)
   b. DE 1.0 Data Users Document (pdf)
   c. DE 1.0 Codebook (pdf)
   d. DE 1.0 Frequently Asked Questions (pdf)

2. On the same webpage as above – navigate to the link found in the slate-blue-gray box at the left --- click on:
   CMS 2008-2010 Data –
   https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/SynPUFs/DE_Syn_PUF.html
3. Morshed S, Tornetta P III, Bhandari M. Analysis of observational studies: a guide to understanding statistical methods. *The Journal of Bone & Joint Surgery*. 2009;91:50-60.

   JUST **SKIM** THIS BELOW……..
4. Bastani R, Glenn BA, Taylor VM, Chen MS, et al. Integrating theory into community interventions to reduce liver cancer disparities: The Health Behavior Framework. *Preventive Medicine*. 2010;50;63-67. OK just to skim this, focusing on the use of the conceptual model……

Optional

From: Long, J.Scott: *The Workflow of Data Analysis Using Stata*

Chapters 1 and 2; and…….
pp. 47-58
pp. 125-139
Chapter 5.6 – 5.9.2 pp. 143 – 165 (this section all review from PH141)
Chapter 6.1.1 to 6.1.2 (partial) pp. 198-201

From Stata Data Management commands:

insheet (review from PH 141)
infix (fixed format)
outsheet
merge (review from PH 141)
collapse
reshape

WEEK 2 April 1

Learning Objectives:

1. Learn structure and utility of analytic planning
2. Understand units of analysis and observation
3. Gain familiarity with using data at multiple levels
4. Understand basic data structure models and key aspects

Required

Slides (pdf) from Beth Virnig, PhD as part of ResDAC/CMS introductory workshop on using Medicare claims for research

Optional

From: Long, J.Scott: *The Workflow of Data Analysis Using Stata*

p. 92-105
6.2. – 6.2.3 (partial) pp. 210 – 229 (This will be a review from PH141)
pp. 279-285

From Stata Data Management

egen
boxplot
dotplot (for turnip plots)

WEEK 3 April 8

Learning Objectives:

1. Understand concepts and practicalities of small area analysis
2. Gain practical skills in complex analytic workflow & analytic memos

3. Simulate research team environment to develop research project
4. Understand how to measure & interpret regional variation

Required

Skinner J. A response to the critics of the Dartmouth Atlas

Goodman DC, Mick S, Bott D, Chang CH, Carretta H, Marth N. Primary Care Service Areas: A new tool for the evaluation of primary care services. *Health Services Research,* 2003;38:287-309. (**Focus on the methods only pp. 290-296)**

Optional
Glover JA. The incidence of tonsillectomy in school children. *Proceedings of the Royal Society of Medicine*. 1938. 31:95-113.

Diez-Roux AV, Multilevel analysis in public health research. *Annual Rev Public Health*. 2000;21:171-92.

Diez-Roux AV. The study of group-level factors in epidemiology: rethinking variables, study designs, and analytical approaches. *Epidemiologic Reviews*. 2004. 26:104-111.

WEEK 4 April 15

Learning Objectives:
1. Develop healthcare service areas: understand utilities, features, limitations
2. Simulate a team science approach to research
3. Understand key elements of an analytic memo and be able to apply

Required:

Critics Question Study Cited in Health Debate: Data Used to Justify Health Savings Effort is Sometimes Shaky, by Reed Abelson and Gardiner Harris, *New York Times*, June 2, 2010.

Dartmouth Atlas: A Response to Its Researchers. *New York Times*, June 18, 2010

Optional:
From: Long, J Scott: *The Workflow of Data Analysis Using Stata*
Chapter 7 (can skim 7.2.1, 7.4.3 – 7.4.5, 7.6.2 – 7.7.1)

WEEK 5 April 22

Learning Objectives:
1. **Explain general principles of multilevel modeling: when to use it, the different types and interpretations**
2. **Discuss nomenclature of multilevel modeling methods.**
3. **Apply multilevel modeling techniques in practice**
4. **Gain familiarity with the R package lme4**

Required:

Stukel TA, Fisher ES, Wennberg JE, et al. Analysis of Observational Studies in the Presence of Treatment Selection Bias. *JAMA*. 2007;(3): 278-285.

Optional:
Diez Roux AF. A Glossary for Multilevel Analysis. *Journal of Epidemiology & Community Health*. 2002; 56(8):588-94.

WEEK 6 April 29

Learning Objectives:

1. **Gain facility with planning your own research project.**
2. **Simulate research team environment to develop research project.**
3. **Understand predictive modeling in theory and application.**

WEEK 7 May 6

Learning Objectives:

1. **Understand the conceptual frameworks of physician patient-sharing networks**
2. **Gain familiarity with the derivation and interpretation of various network measures**
3. **Apply network analysis to publicly available CMS data**

Required:

DuGoff EH, Fernandes-Taylor, S, Weissman GE, et al. A scoping review of patient-sharing network studies using administrative data. *Trans Behav Med.* 2018; 8(4):598-625.

Landon BE, Keating NL, Onnela JP, et al. Patient-sharing networks of physicians and health care utilization and spending among Medicare beneficiaries. JAMA Intern Med. 2018; 178(1):66-73.

WEEK 8 May 13

Learning Objectives:

1. **Gain familiarity with geospatial approaches to health care delivery, and GIS in particular.**
2. **Learn concepts and methods of spatial analysis applied to health care**
3. **Understand access to care and methodological approaches to measuring**

Required:

Wang F, Onega T. Accessibility of cancer care: disparities, outcomes, and mitigation. *Annals of GIS*. 2015;21(2):119-125.

Onega T, Alford-Teaster, Andrews S, Ganoe C, Perez M, King D, Shi X. Why health services research needs geoinformatics: rationale and case example. *Journal of Health & Medical Informatics.* 2014. http://dx.doi.org/10.4172/2157-7420.1000176 2. UI: 25893140.

WEEK 9 May 20

Learning Objectives:

1. **Understand structure, utility, and basics of electronic health record data**
2. **Apply course principles to use of electronic health record data**
3. **Synthesize course concepts and demonstrate mastery of key points**
4.

Required:

Weiskopf NG, Weng C. Methods and Dimensions of Electronic Health Record Data Quality Assessment: Enabling Reuse for Clinical Research. *JAMIA*. 2013;20:144-151.

Optional:

Madigan D, Ryan PB, Schuemie M, Stang PE, et al. Evaluating the Impact of Database Heterogeneity on Observational Study Results. *Am J Epid* 2013;178(4):645-651.

WEEK 10 May 31 (Friday)

Learning Objectives:

1. **Hone skills of distilling research projects into effective presentations**
2. **Practice the role of peer review process.**

Required:

Seals DR and Tanaka H. Manuscript Peer Review: A Helpful Checklist for Students and Novice Referees*. Adv Physiol Educ.* 2000;22:52-58.

PH 147 SPRING 2018

	8:45-9:00 AM	9:00 to 10:00 AM and 10:10 to 11:00 AM		11:10 to noon	Noon to 1:00	1:00 - 4:30 PM	
	Help & Individual Discussion (optional)	**Core Content and Group Synthesis**		**Methods in Practice**	**Lunch**	**Data Analysis Lab**	**Learning Objectives for Week**
Week 1 - March 26		Course Overview; Observational Research Kudos and Caveats	Project Session I: Conceptual modeling and translating into analyses	Datasets & documentation & Data Management (download data)	Lunch	Data Day: Loading Data, Assessing Data Quality, and Workflow Management	1. What is meant by workflow of data analysis 2. Conceptual frameworks – who, what, where, when, why, how 3. Basic tenets of units of observation and analysis 4. Basic tenets of data models & analytic documentation 5. Data management and data quality
Week 2 - April 2		Data Structure	Class Project Session II: Developing an Analytic Plan	Medicare Claims & other Secondary Data Planning Analytic Workflow	Lunch	Indirect Adjustment across geographic areas using Medicare data	1. Gain familiarity with administrative data – strengths and weaknesses 2. Understand units of analysis and observation 3. Gain familiarity with using data at multilple levels 4. Gain familiarity with risk adjustment in principle and practice 5. Apply a method of adjustment to assess geographic variation
Week 3 - April 9		Project Session III: Analytic Plan (part b): student project planning	Small Area Analysis: Part I	Group Work to Plan Lab Analytic Workflow	Lunch	Measuring regional variation in utilization	1. Understand concepts and practicalities of small area analysis 2. Gain practical skills in complex analytic workflow & analytic memos 3. Simulate research team environment to develop research project 4. Understand how to measure & interpret regional variation
Week 4 - April 16		Analytic Memos and Analyst-investigator pairs: Student project peer feedback	Small Area Analysis: Part II	Group Work to Plan Lab Analytic Workflow	Lunch	Defining health care markets* (aka creating service areas a la the Atlas)	1. Develop healthcare service areas: understand utilities, features, limitations 2. Simulate a team science approach to research 3. Understand key elements of an analytic memo and be able to apply
Week 5 - April 23		Quick tour: Bayes, Empirical Bayes, & Instrumental Variables	Propensity Scores	Inverse propensity weighting class exercise/demo	Lunch	Propensity Scores	1. Gain an introductory applied understanding of Bayesian methods 2. Gain an introductory applied understanding of instrumental variables 3. Understand the basics of propensity score methods 4. Apply regression and propensity score methods in practice
Week 6 - April 30		Risk Adjustment from Stakeholders Perspective: Student-led: Journal Club	Mixed Methods	Pablo Martinez Camblor, PhD	Lunch	Structured student project lab	1. Gain facility with planning your own research project. 2. Understand risk adjustment from multiple stakeholders' perspectives 3. Learns selection and application of best methodological approaches
Week 7 - May 7		Mutlilevel Analyses -- Therese Stukel, PhD Institute for Clinical Evaluative Sciences, Toronto			Lunch	Readmission Rates Lab	1. Understand general principles of multilevel modeling – when to use, different types, how to interpret 2. Understand nomenclature of multilevel modeling methods 3. Gain experience in applying multilevel modeling techniques and apply longitudinal methods to analysis 4. Independently demonstrate the art and application of analytic workflow
Week 8 - May 14		Health Informatics Intro: New Methods in the Mainstream	Social Network Analysis: Erika Moen, PhD	Introduction to R and Advanced Analytics	Lunch	R and Stata Lab: Platform-Agnostic Advanced Data Analytics	1. Gain familiarity with new informatics-based approaches to observational research, such as natural language processing (NLP) and data integration 2. Learn concepts and methods of network analysis applied to health care 3. Understand current data analytic tools, and applications beyond Stata
Week 9 - May 21		Electronic Health Records	EHR-hands-on exercise (mini-lab)	Group Activity: JEOPARDY!!!	Lunch	Indiviudal Course Project Help (Optional)	1. Understand structure, utility, and basics of electronic health record data 2. Apply course principles to use of electronic health record data 3. Synthesize course concepts and demonstrate mastery of key points
Week 10 - June 1 FRIDAY		CLASS BEGINS AT 9:00 a.m.	Oral Presentations of Individual Course Projects & Peer Review		CLASS ENDS BY 12:00; Lunch Provided		1. Hone skill of distilling research projects into effective presentations 2. Practice the role of peer reviewer

	General content type
	Conceptual/Didactic
	Skills/Practicum
	Statistical methods
	Integration of all