





# IMPRS

INTERNATIONAL MAX PLANCK  
RESEARCH SCHOOL  
population. health. data science.

## PHDS – Population Health Course Outline

**Coordinators** Jonas Schöley  [schoeley@demogr.mpg.de](mailto:schoeley@demogr.mpg.de)  
Mine Kühn  [kuehn@demogr.mpg.de](mailto:kuehn@demogr.mpg.de)

**Date** November 21 to Dec 2, 2022

**Time** 10am to 4pm CET with lunch break from 12 to 2pm

**Location** In person @ MPIDR room 005 with [Zoom](#) option and Nextcloud repository  
Zoom Meeting-ID: 997 1515 3278  
Passcode: 641642

### Course description

This intensive two-week course introduces key substantive and methodological topics in population health. Participants will be exposed to various research areas and methods within population health, with experts on global health, mortality trends, socio-economic differentials, intersectionality, family dynamics, migration, aging, causal inference, and multistate methodology coming together to give a 360-degree view of the current research landscape.

### Organization

The course will take place in person at MPIDR with a Zoom stream available. Exceptions are made for two seminars which will take place virtually on Zoom. These exceptions are highlighted in the course outline below.

Each day covers one topic taught by a domain expert. Depending on the topic and the instructor, the lecture format varies and may include a pre-recorded part, a coding session, or group work. Lecture materials will be made available at least two days before the corresponding session in a Nextcloud folder. In general, students should expect to spend about 6-8 hours of effort per day on the course (lectures, discussions, readings, exercises).

### Course prerequisites

Students are expected to have basic knowledge of R and Stata. Participants need a laptop or desktop computer with the latest versions of R and RStudio installed. Instructions on how to download and install R can be found in “A (very) short introduction to R” by Torfs and Brauer (2014): <https://cran.r-project.org/doc/contrib/Torfs+Brauer-Short-R-Intro.pdf>.

If you don't have sufficient knowledge about R, you can use the following websites to familiarize yourself with the program: [swirlstats.com](http://swirlstats.com), [coursera.org/course/rprog](https://coursera.org/course/rprog), [stats.idre.ucla.edu/r](http://stats.idre.ucla.edu/r).

The two seminars on multistate models will require some basic familiarity with linear algebra. You should familiarize yourself with vectors, matrices, and basic operations on them by reading the following literature:

Chapters 1-3 (except section 3.9) of "Matrix Operations for Engineers and Scientists" by A. Jeffrey, [10.1007/978-90-481-9274-8](https://doi.org/10.1007/978-90-481-9274-8). If you do not have time to read the full three chapters, focus on chapter 1 and sections 2.1-2.3, 3.1 (skip section 3.1.1), 3.2, and 3.4.

Sections 1.1.1 and 1.1.2 of the Appendix to the paper "Flexible Transition Timing in Discrete-Time Multistate Life Tables Using Markov Chains with Rewards" by D. C. Schneider et al., [10.4054/MPIDR-WP-2021-002](https://doi.org/10.4054/MPIDR-WP-2021-002).

### **Examination**

Active participation in all classes is required to earn the course certificate. No grades will be given beyond a pass/fail.

### **General readings**

To be announced. Instructors will typically assign one or two core readings, and several optional readings to develop a deeper understanding of the topic. The readings will be made available with the course materials.

## Week 1 Topics in Population Health

**Mortality trends** Nov 21, 10am to 4pm CET

Enrique Acosta  [acosta@demogr.mpg.de](mailto:acosta@demogr.mpg.de)

Materials will be available in Nextcloud latest by Nov 17

Tallies of the dead are one of the oldest and most widespread measures of population health. Thus, we start the week by positioning historical global mortality changes within the epidemiological transition theory, distinguishing between extrinsic and intrinsic causes of death, and explaining the distinction between age, period, and cohort effects on mortality. We implement the learned concepts by analyzing mortality data in R.

**Seminar is fully virtual via [zoom](#)**

**Family and Health** Nov 22, 10am to 4pm CET

Mine Kühn  [kuehn@demogr.mpg.de](mailto:kuehn@demogr.mpg.de)

Materials will be available in Nextcloud latest by Nov 18

Family members can provide social support and resources that are beneficial for health. At the same time, certain family contexts can be characterized by stress and conflicts that may be harmful. The course covers theoretical concepts and empirical evidence on how health and well-being are associated with family dynamics. In the second part of the course, we will discuss “research in action” by looking at MPIDR research projects that investigate how family life events and critical periods in life impact health and well-being. The lecture includes interactive elements that require active student participation.

**Intersectionality, Cumulative Disadvantage & Health** Nov 23, 10am to 4pm CET

Jo Mhairi Hale  [Jo.Hale@st-andrews.ac.uk](mailto:Jo.Hale@st-andrews.ac.uk)

Materials will be available in Nextcloud latest by Nov 21

In this session, we will employ theories of intersectionality and cumulative (dis)advantage to explain health disparities over the life course. We will look at some examples in papers by Jo with various other Max-Planckers and learn about various methods to investigate how racial/ethnic, nativity, socio-economic, and gender identities intersect to affect health/wealth outcomes.


**Global Health** Nov 24, 10am to 4pm CET

Joshua Wilde  [wilde@demogr.mpg.de](mailto:wilde@demogr.mpg.de)

Materials will be available in Nextcloud latest by Nov 22

This section will cover the bi-directional interaction between economic development and population health and dynamics in low and middle-income country settings. Particular attention will be given to the effects of various demographic transitions which have occurred since the 1920s in LMICs – particularly of mortality, fertility, and infectious disease incidence – and their feedback to economic outcomes, both at the micro and macro levels.

**Migration and Health** Nov 25, 10am to 4pm CET

Silvia Loi  [loi@demogr.mpg.de](mailto:loi@demogr.mpg.de)

Materials will be available in Nextcloud latest by Nov 23

We will focus on native-immigrant health disparities, starting with an introduction to the basic concepts, terminology, and literature on migration and health. We will describe how

every single phase of the migration experience is associated with different health risks, from the decision to migrate in the origin country to the arrival in the receiving country (and eventually to the return migration in later life). We will then cover key issues, core theories, and hypotheses commonly used to explain native-immigrant health disparities. We will cover “the healthy immigrant effect” and “the salmon bias,” as well as the process of deterioration of immigrant health over time, the “immigrant-native health convergence process,” and the mechanisms behind it. We will continue illustrating health disparities across immigrant generations. Lastly, we will study some empirical examples on the relationship between early life exposures of immigrants and later health consequences and the immigrant-native health gap.

## Week 2 Methods in Population Health

### **Multistate Models I** Nov 28, 10am to 4pm CET

Christian Dudel  [dudel@demogr.mpg.de](mailto:dudel@demogr.mpg.de)

Daniel Schneider [schneider@demogr.mpg.de](mailto:schneider@demogr.mpg.de)

Mikko Myrskylä  [myrskylä@demogr.mpg.de](mailto:myrskylä@demogr.mpg.de)

Materials will be available in Nextcloud latest by Nov 24

Many social processes can be represented by individuals being in, and transitioning between, states. Statistical modelling of such processes can be done using multistate models. Multistate models are becoming an increasingly popular tool for social scientists, as they provide many quantities to describe longitudinal data in general and life course trajectories in particular. For instance, multistate models are commonly used to estimate healthy life expectancy, i.e., the average lifetime spent in good health.

Multistate models come in two different variants: discrete-time and continuous-time. While continuous-time multistate models have been extensively described in the literature, there is little guidance on their discrete-time counterparts.

In this course we provide a mostly non-technical introduction to discrete-time multistate models, their theoretical foundations and assumptions, and their empirical estimation with R and Stata. Students will be guided through all steps required to generate some of the key outputs of discrete-time multistate models, such as state expectancies. This equips them with the skills and tools needed to use and implement multistate models in their own research.

### **Required reading**

Chapters 1-3 (except section 3.9) of “Matrix Operations for Engineers and Scientists” by A. Jeffrey, [10.1007/978-90-481-9274-8](https://doi.org/10.1007/978-90-481-9274-8). If you do not have time to read the full three chapters, focus on chapter 1 and sections 2.1-2.3, 3.1 (skip section 3.1.1), 3.2, and 3.4.

Sections 1.1.1 and 1.1.2 of the Appendix to the paper “Flexible Transition Timing in Discrete-Time Multistate Life Tables Using Markov Chains with Rewards” by D. C. Schneider et al., [10.4054/MPIDR-WP-2021-002](https://doi.org/10.4054/MPIDR-WP-2021-002).

### **Multistate Models II** Nov 29, 10am to 4pm CET

Christian Dudel  [dudel@demogr.mpg.de](mailto:dudel@demogr.mpg.de)

Daniel Schneider [schneider@demogr.mpg.de](mailto:schneider@demogr.mpg.de)

Mikko Myrskylä  [myrskylä@demogr.mpg.de](mailto:myrskylä@demogr.mpg.de)

Materials will be available in Nextcloud latest by Nov 25

**Predictive Modeling** Nov 30, 10am to 4pm CET

Jonas Schöley  [schoeley@demogr.mpg.de](mailto:schoeley@demogr.mpg.de)

Materials will be available in Nextcloud latest by Nov 28

In population health prediction takes many forms: we may be interested in the future development of fertility, the future care home demand, the infant mortality in regions of the world without reliable data collection. We may ask “what-if” questions that require prediction for a scenario that never happened: How many deaths without COVID-19? How many deaths without the summer heat wave? These questions are concerned with the prediction of an outcome. The focus is on the left-hand side of the regression equation, not on the interpretation of coefficients. In this seminar I will demonstrate the predictive modeling workflow: formulation of a prediction model, out-of-sample model validation, and simulation-based uncertainty quantification and propagation. We will discuss a range of applications for predictive modeling within the context of population health.

**Instead of the morning lecture, watch the recording  
Join in person at 2pm for exercises**

**Causality I** Dec 1, 2pm to 4pm CET

Angelo Lorenti  [lorenti@demogr.mpg.de](mailto:lorenti@demogr.mpg.de)

Materials will be available in Nextcloud latest by Nov 29

In this part of the course, you will gain an understanding of the basic ideas, principles, and methods for causal inference on observational data. In a pre-recorded lecture video, I will first introduce the potential outcomes model as a useful tool to formulate causal research questions and identify potential problems for inference. I will then discuss three basic methodological approaches (fixed effects, difference-in-differences designs, instrumental variables) to resolve some of these issues. In the live tutorial session, we will work together through a real-world dataset to see how these methods are implemented in practice, how their output can be interpreted and how to avoid common pitfalls.

**Causality II** Dec 2, 10am to 4pm CET

Angelo Lorenti  [lorenti@demogr.mpg.de](mailto:lorenti@demogr.mpg.de)

Materials will be available in Nextcloud latest by Nov 30

Although many research questions in the social sciences are inherently causal, it is common to avoid using the word ‘causal’ in favor of the word association when writing research articles. Building upon the understanding gained in the Causality I course, we further explore the potential outcome framework to understand how a design can help us toward causal thinking and to lay out the assumptions needed to identify causal effects. The theoretical introduction is complemented with a description and some practical applications of propensity-score-based analysis methods, as inverse probability treatment weighting, marginal mean weighting, and matching.