

COLLEGE OF POPULATION HEALTH

## *PopTalk Webinar Series*

# What is Health Data Science?

*An Introduction for Health Care Professionals and Clinicians*

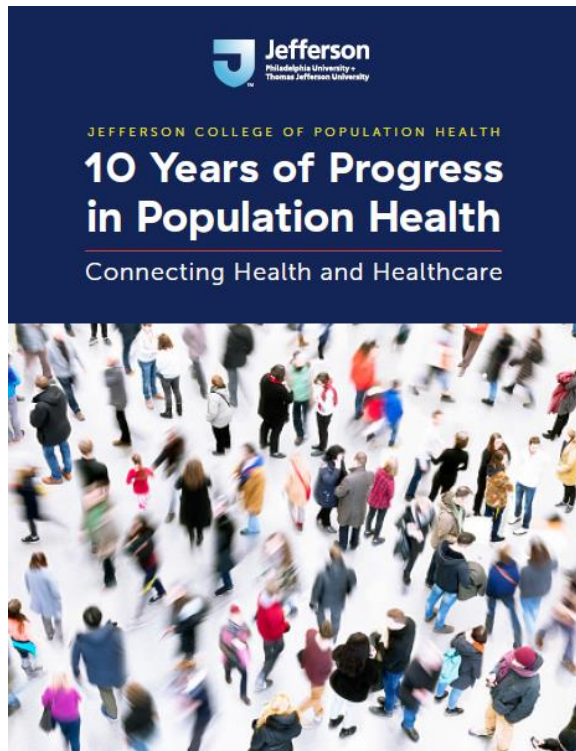
January 13, 2020

Harm Scherpbier, MD, MS  
Karen Walsh, DHSc, MS, MBA



**Jefferson**  
Thomas Jefferson University

# Jefferson College of Population Health



To prepare leaders with global vision to *develop, implement* and *evaluate* health policies and systems that improve the health of populations and thereby enhance the quality of life

# Today's Presenters

---



**Karen Walsh, DHSc, MS, MBA**  
*Program Director, Health Data  
Science*  
*Jefferson College of Population  
Health*

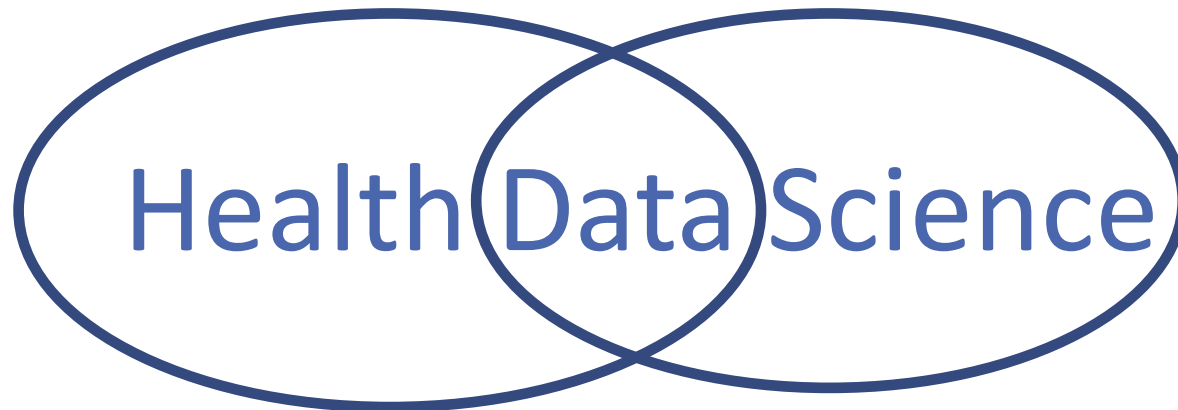


**Harm Scherpbier, MD, MS**  
*Chief Medical Information  
Officer HealthShare Exchange*

# What is Health Data Science?

---

A multi-disciplinary field that involves using statistical inference, algorithmic development, and technology to make insights about data in healthcare settings.



# Health Data Science

---

To uncover actionable insights, skilled healthcare data scientists are needed to:

- ✓ Combine large disparate data sources
- ✓ Build statistical and predictive models
- ✓ Create effective data visualizations
- ✓ Communicate findings to technical and non-technical audiences

# Volume & Velocity of big data in healthcare

- Since 2016, there has been an 878% growth in healthcare data -- more than manufacturing, financial services or media industries
- There is an exponential growth of healthcare data each year
- Healthcare organizations must develop techniques to protect, store, manage, and process these data



Resource: [https://www.dellemc.com/sl-si/collaterals/unauth/briefs-handouts/solutions/h17823\\_solution\\_brief\\_driving\\_real\\_clinical\\_business\\_outcomes\\_with\\_a\\_modern\\_it.pdf](https://www.dellemc.com/sl-si/collaterals/unauth/briefs-handouts/solutions/h17823_solution_brief_driving_real_clinical_business_outcomes_with_a_modern_it.pdf)

Resource: <https://healthitanalytics.com/news/understanding-the-many-vs-of-healthcare-big-data-analytics>

# Variety of big data in healthcare

## Internet of Things (IoT)

**Wearables and trackers**  
(heart rate, weight, activity, stress levels)

**Apps** (exercise, sleep quality, etc)

Medical devices and sensors (glucose monitors, oximeters, blood pressure monitors)

## Insurance Claims Data

Can be used to assess **medication compliance**

Captures **healthcare services, procedures, tests**

## Electronic Medical Records (EMRs)

**Improves condition identification**  
(e.g., use of lab result or vital signs)

**More timely**  
(do not have to wait for receipt and process of claim to identify certain healthcare interactions)

**Additional clinical information**  
(vital signs, lab results, smoking and alcohol use, survey results [depression screenings, asthma control tests, etc])

## Other Relevant Data

**Genomics data**

**Doctor's notes**

**Environmental data**  
(Census data, air quality data, neighborhood-level factors)

**Social Media**

Resource: <https://www.optum.com/content/dam/optum/resources/whitePapers/Benefits-of-using-both-claims-and-EMR-data-in-HC-analysis-WhitePaper-ACS.pdf>

Resource: <https://emerj.com/ai-sector-overviews/where-healthcares-big-data-actually-comes-from/>

# Value of big data in healthcare

- The application of big data in healthcare can:
  - Reduce healthcare cost
  - Improve patient outcomes
  - Aid in novel drug discovery
  - Prevent progression of disease
  - Accelerate personalized medicine in cancer treatment
  - Reduce fraud and enhance security
  - Enhance accuracy of image diagnostics





# What is Statistical Learning?

- Statistical learning refers to a “set of tools for modeling and understanding complex datasets” (James et al. *Intro to Statistical Learning*)
  - Arose from the field of statistics and merged with new developments in computer science, in particular machine learning
- With the explosion of “big data” statistical learning techniques have been in high demand to answer questions such as:
  1. Identify the risk factors for breast cancer
  2. Predict whether a patient will have a stroke based on clinical, demographic, and diet data
  3. Identify fraudulent claims
  4. Classify tissue samples into cancer classes

# Population Health Definition

---

Population Health is:

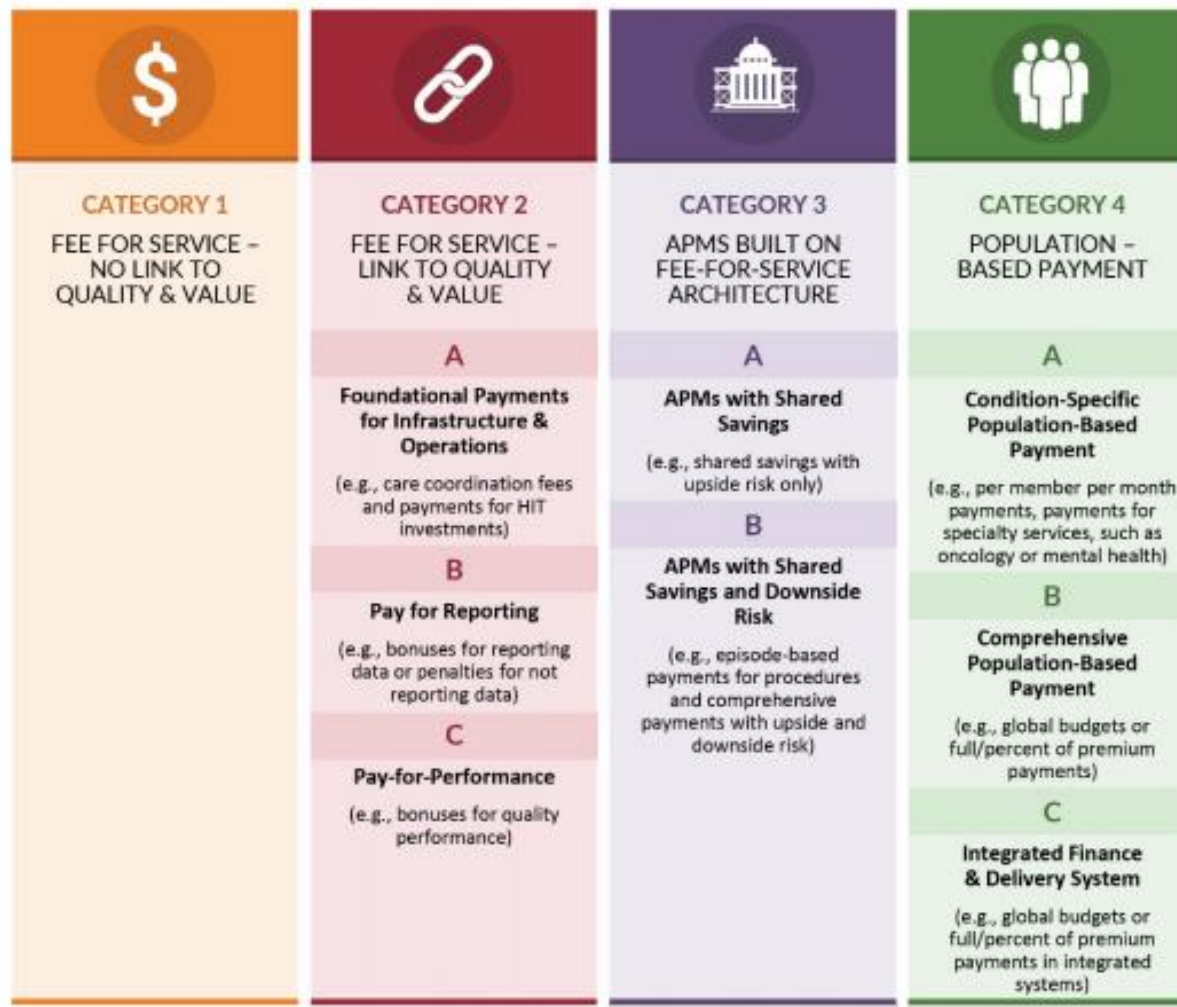
A Healthcare Delivery Model,

Coupled with a Reimbursement or Payment Model,

Designed to deliver the most health, at the lowest cost, to a defined population,

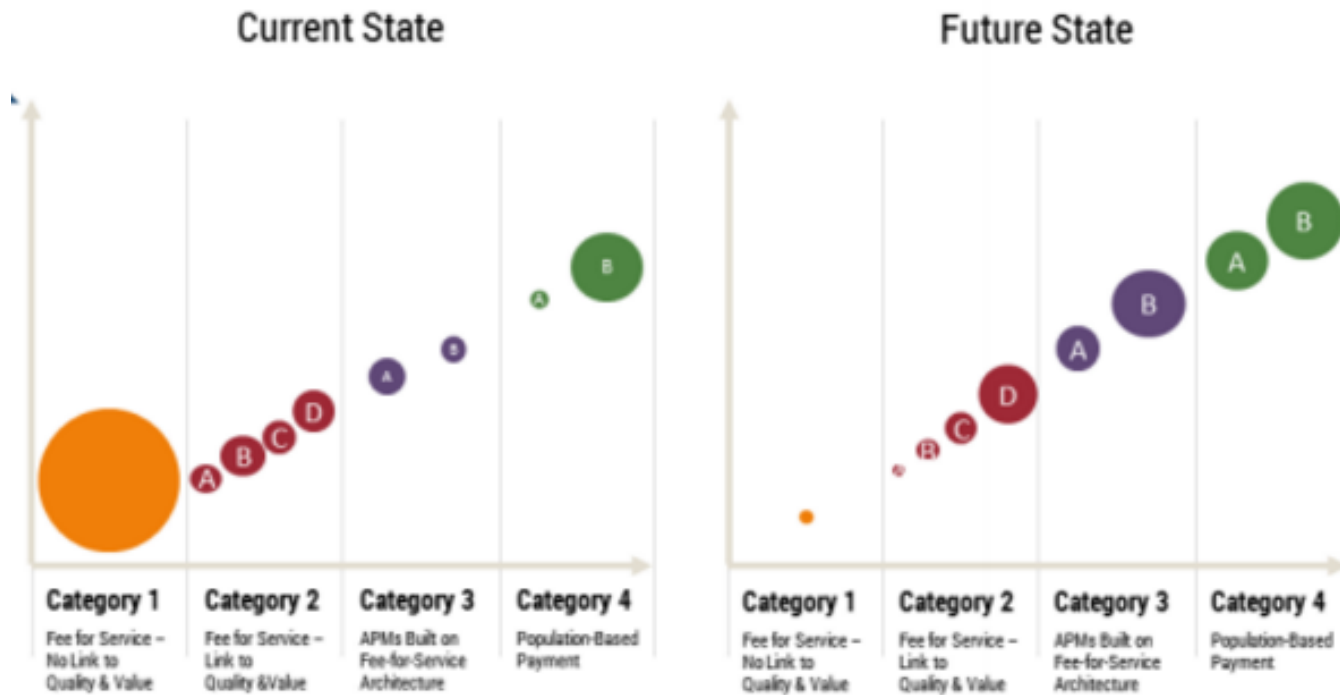
Supported by a Data Infrastructure.

# HCP-LAN Alternative Payment Model Categories



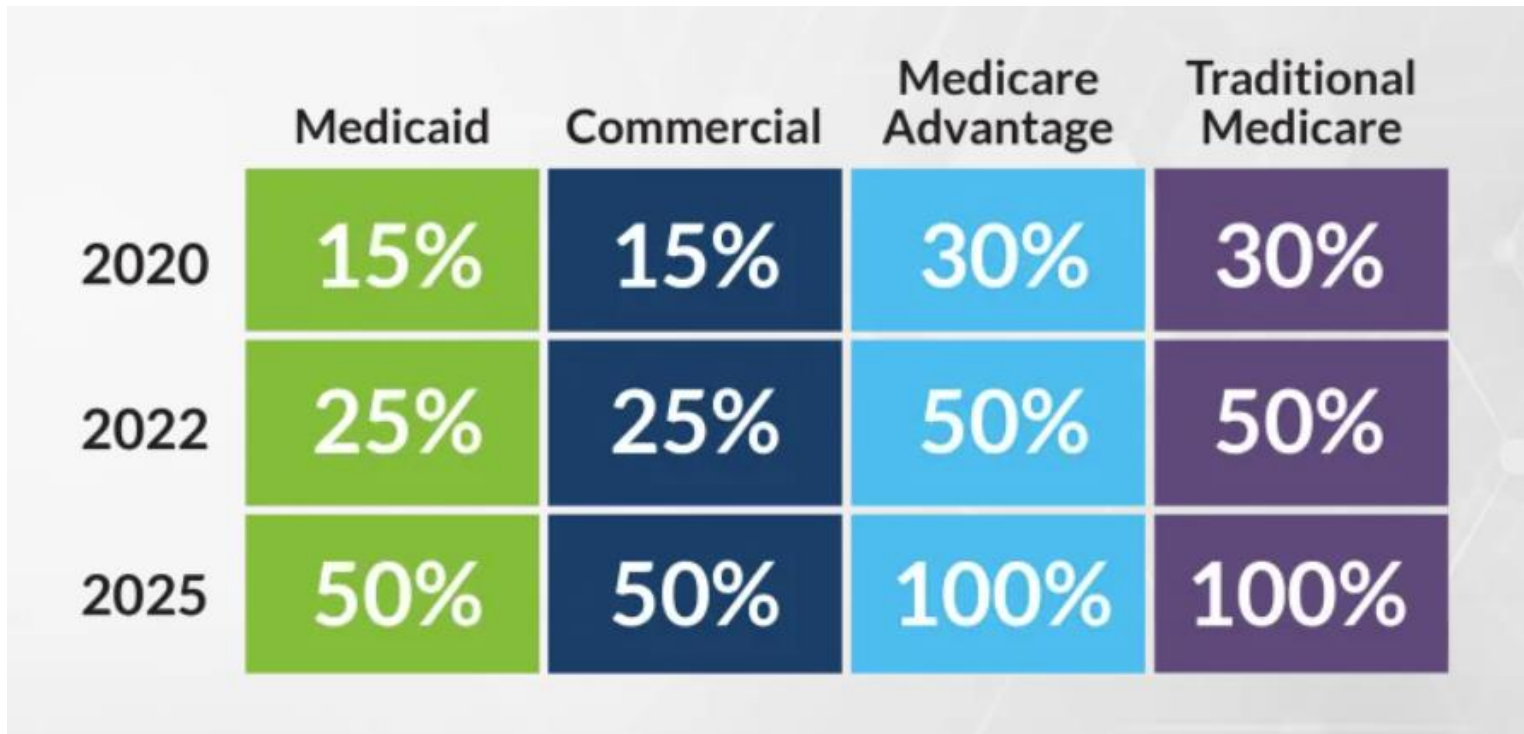
# Current State and Future State

Level of Financial Risk for Provider



# APM Trajectory

---



# Data Science as Teamwork – following Davenport’s Analytical Process

## Analytics-Based Decision Making—in Six Key Steps

When using big data to make big decisions, non-quants should focus on the first and the last steps of the process. The numbers people typically handle the details in the middle, but wise non-quants ask lots of questions along the way.

### Recognize the problem or question

Frame the decision or business problem, and identify possible alternatives to the framing.

### Review previous findings

Identify people who have tried to solve this problem or similar ones—and the approaches they used.

### Model the solution and select the variables

Formulate a detailed hypothesis about how particular variables affect the outcome.

### Collect the data

Gather primary and secondary data on the hypothesized variables.

### Analyze the data


Run a statistical model, assess its appropriateness for the data, and repeat the process until a good fit is found.

### Present and act on the results

Use the data to tell a story to decision makers and stakeholders so that they will take action.

Thomas Davenport, Keeping up with your Quants, HBR 2013

# HIMSS Analytics Maturity Stack

STAGE	 Adoption Model for Analytics Maturity Cumulative Capabilities
7	Personalized medicine & prescriptive analytics
6	Clinical risk intervention & predictive analytics
5	Enhancing quality of care, population health, and understanding the economics of care
4	Measuring and managing evidence based care, care variability, and waste reduction
3	Efficient, consistent internal and external report production and agility
2	Core data warehouse workout: centralized database with an analytics competency center
1	Foundation building: data aggregation and initial data governance
0	Fragmented point solutions

Source: <https://www.himssanalytics.org/amam>

# Core Health Data Science Training

---

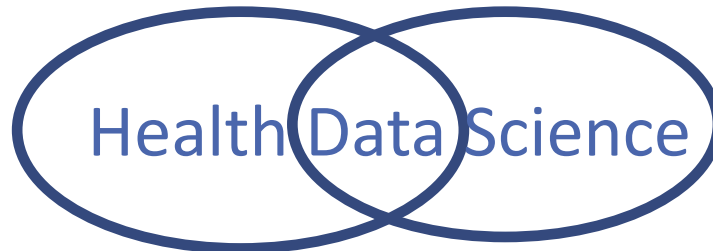
Healthcare  
Informatics

Data Science

Population  
Health

Data  
Visualization

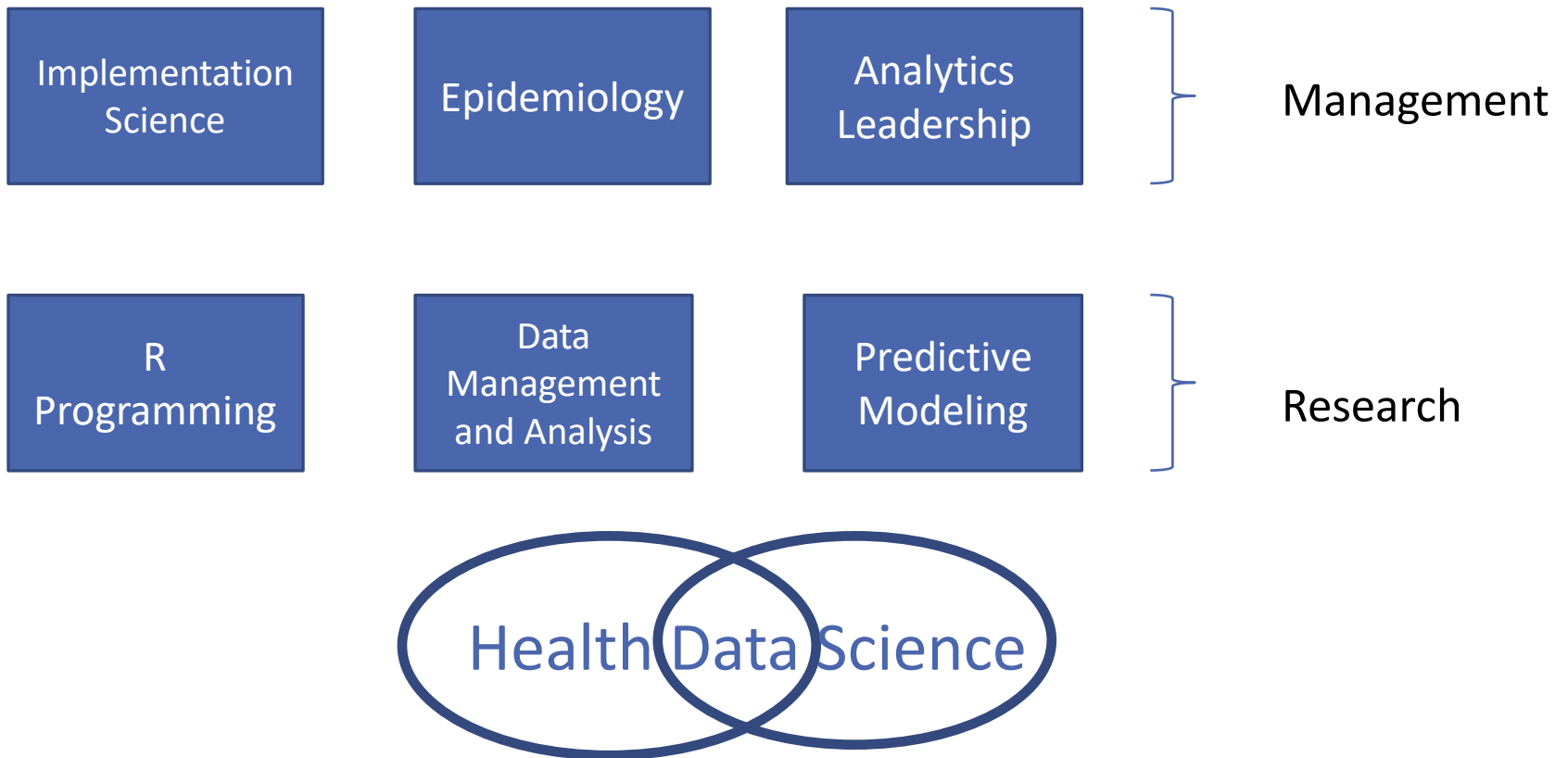
Statistics





# Health Data Science Training

---



# New Master's Degree in Health Data Science

## Management Track

- Competencies in HDS data, statistics, predictive analytics, and the ability to interpret the results and gain insights on data.
- Target audience will include individuals who wish to start a career in the healthcare industry and individuals in healthcare who want to expand or advance their career.

## Research Track

- Competencies in data wrangling, statistical, and predictive analytics to work on HDS projects
- The targeted audience will include individuals who want to perform HDS research in industry.



# Q & A

# New Master's Degree in Health Data Science

Complete a graduate certificate in 1 year  
or Master's degree in 2 years



Health Data Science is an ever-evolving multi-disciplinary field that involves using statistical inference, algorithmic development, and technology to make insights about data.

- Two Tracks: Management and Research
- 100% online
- Accelerated 7-week courses
- Expert practitioner faculty

Learn more at: [Jefferson.edu/HDS](https://jefferson.edu/HDS)

Questions: [JCPH.Admissions@jefferson.edu](mailto:JCPH.Admissions@jefferson.edu)

**JCPH Virtual Open House**  
**February 10 | 5:30-7:30 pm**  
[Register Here](#)

# PopTalk

## Webinar Series

One-hour webinars featuring experts in population health.



## Upcoming PopTalks

**Trust, Timing and COVID: Attitudes & Vaccine Policy**

January 21, 2021 | 12:00-1:00 pm ET

[Register Now](#)

**Making Philadelphia Parks Safe & Healthier**

February 3, 2021 | 12:00-1:00 pm ET

[Register Now](#)

**Economic Evaluation of Vaccines: Challenges & Opportunities**

February 17, 2021 | 12:00-1:00 pm ET

[Register Now](#)

[Visit our Website](#)  
for more information

