Using Epidemiology and Artificial Intelligence to Describe a Complex Primary Care Population in a Learning Health System

50th NAPCRG Annual Meeting

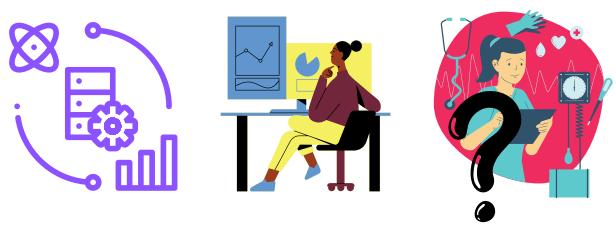
Jaky Kueper, PhD

Co-Authors: Jennifer Rayner, Merrick Zwarenstein, Dan Lizotte

Contact
- Email jkueper@uwo.ca
- Twitter @jk_kueper

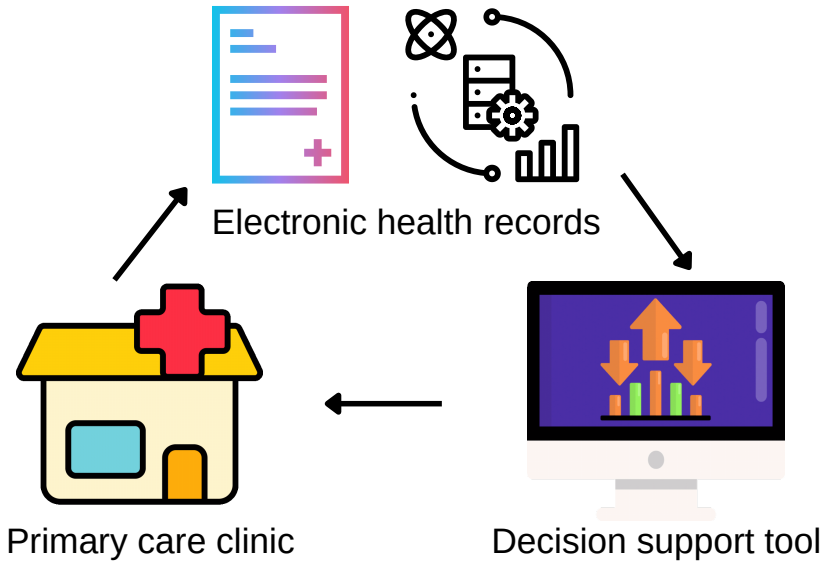
Motivation

- **Challenge**: Providing the best health care, especially for clients experiencing complex clinical and/or social situations
  - Not well represented in research literature or clinical guidelines
  - Burden of treatment may be high
  - Need for whole-person, client-centred care
- **Opportunity**: Increases in "everyday" data and computing resources
  - Electronic health records generated through care delivery
- **Potential support**: Careful analysis of these data to derive value e.g., personalized decision support tools
  - Primary health care is a relatively understudied area



Background: Decision Support Tools

- Need to start with real challenges and end up with impactful solutions - many possibilities!
- AI for health and LHS guidelines emphasize
  - Organization culture and capacity
  - Stakeholder engagement and leadership
  - Data availability, provenance, quality

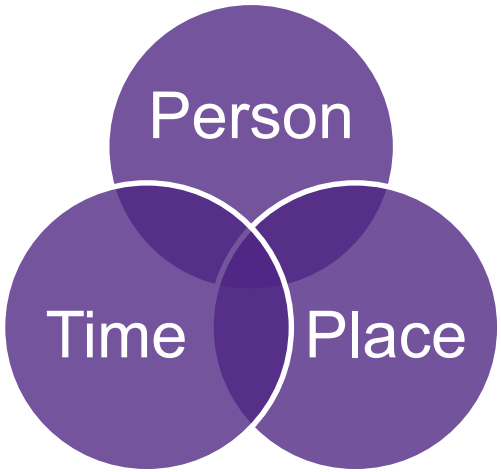


Example Learning Health System (LHS) paradigm

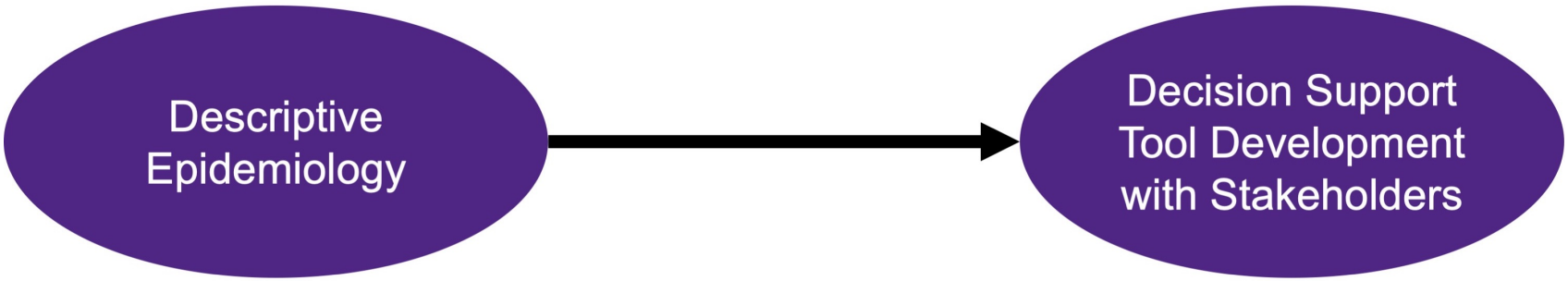
Descriptive epidemiology as a valuable **add in** to early stages of decision support tool development?

Descriptive Epidemiology

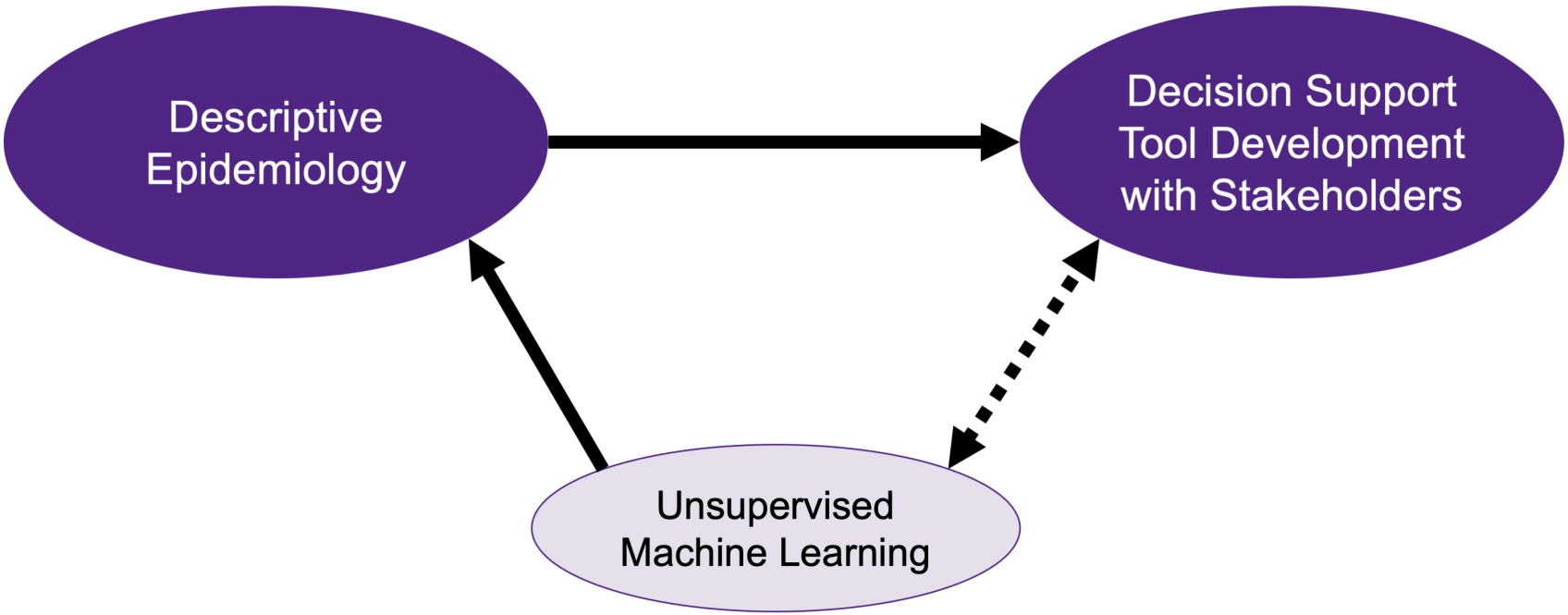
- Understand health-related states and needs of a population
  - Who? What? When? Where? How?
- Valuable tool in public health practice and system planning
- Population-level descriptions may complement internal stakeholder expertise
  - Problem refinement and selection
  - Understanding of data availability and quality
  - Identify methodological considerations
  - Evaluation of long-term progress



Proposed Role for Descriptive Epidemiology



Proposed Role for Descriptive Epidemiology

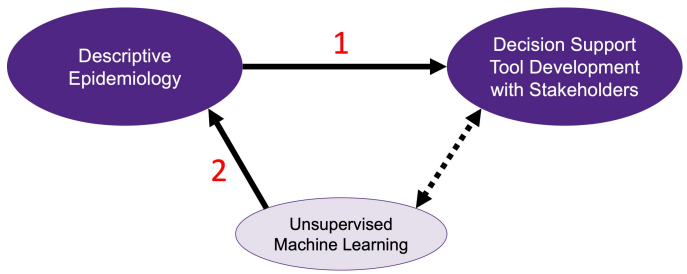


- To properly understand complex populations, both simple statistical techniques and techniques that can capture more complex patterns may be useful.

Study Goals

1. To **describe sociodemographic, clinical, and healthcare use characteristics** of a complex primary care population for the purpose of supporting future initiatives, including the development of decision support tools.

2. To demonstrate how **both simple statistical and unsupervised learning techniques**, applied with an epidemiological lens, can be used to describe complex populations.



Alliance for Healthier Communities

- Provide team-based primary health care through 72 Community Health Centres (CHCs) across Ontario
  - 18 Urban-at-Risk CHCs
  - 9 Rural geography CHCs
- Focus on people facing challenges or barriers to care that increase their risk for poor health
- Central, structured EHR database (we use 2009-2019)
- In 2020, adopted a learning health system framework[1]

Alliance for Healthier Communities
Alliance pour des communautés en santé

1. Nash DM, Rayner J, Bhatti S, Zagar L, Zwarenstein M. The Alliance for Healthier Communities' journey to a learning health system in primary care. Learn Health Syst. 2021:e10321. doi:10.1002/lrh2.10321

Methods: Cohort & Characteristics of Interest

- **Base cohort**: Adult primary care clients with 1+ encounter in 2009-2019
- **Sociodemographic characteristics**
  - Structured table fields
  - ENCODE-FM codes
- **Clinical characteristics**
  - 20 chronic conditions[1] + 4 Alliance-identified conditions
  - ENCODE-FM and ICD-10 codes
- **Health care use characteristics**
  - Providers involved in care
  - Complexity of care: number of distinct issues addressed per visit
  - Frequency of care: number of days of care access per year or quarter-year

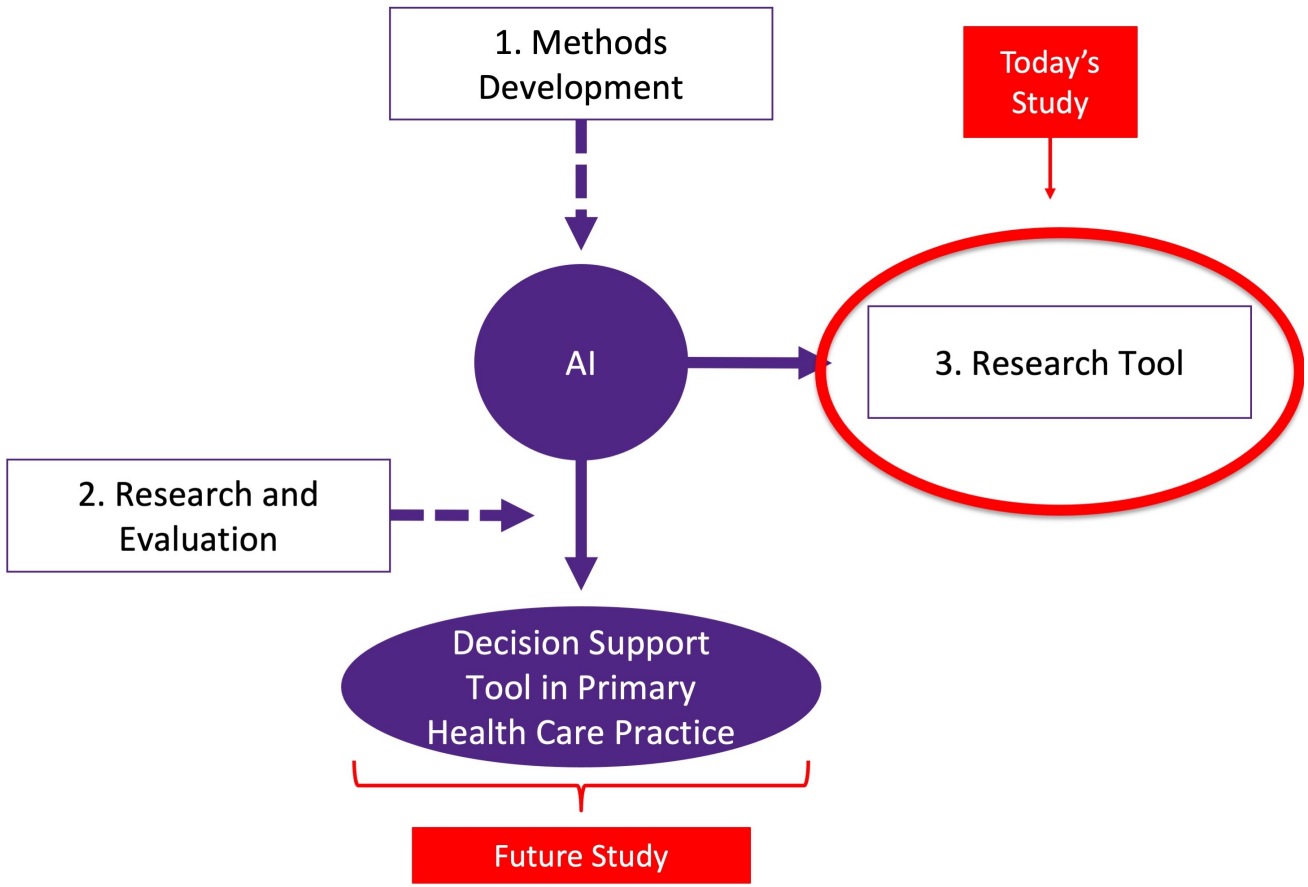
1. Fortin M, Almirall J, Nicholson K. Development of a research tool to document self-reported chronic conditions in primary care. J Comorb. 2017;7(1):117-123. doi:10.15256/joc.2017.7.122

Analysis Techniques

- **Sociodemographic characteristics**
  - Table-based summaries, overall & stratified
- **Clinical characteristics**
  - Prevalence by calendar and observation time
  - Cumulative incidence
  - *Ising model*: co-occurence patterns
- **Healthcare use characteristics**
  - Table-based summaries, overall & stratified
  - *Non-negative matrix factorization*: common care provider teams
  - *K-medoids time series clustering*: care access frequency trends

Aside: AI as a Research Tool

- AI is a tool that can help to explore and describe more complex patterns in data than is possible with simple statistics



Results: Sociodemographic characteristics

- Base cohort: 221,047 clients
- Summarized 13 characteristics, e.g., age, rural residence geography, household income, food insecurity, residence stability
- Social determinants expected to increase risk of poor health were more prevalent in Urban-at-Risk CHC and multimorbidity strata
- Completeness varied by characteristic, client, provider, and clinic factors.
  - Level of analysis matters, e.g., self-report health measures collected by a subset of CHCs

Clinical characteristics

11-year period prevalence: Burden of conditions from system planning perspective

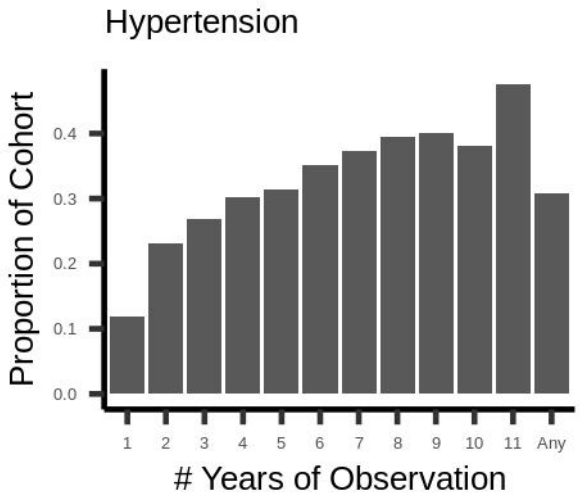
Estimates ranged from 2% (Hepatitis C) to 63% (Chronic musculoskeletal problem)

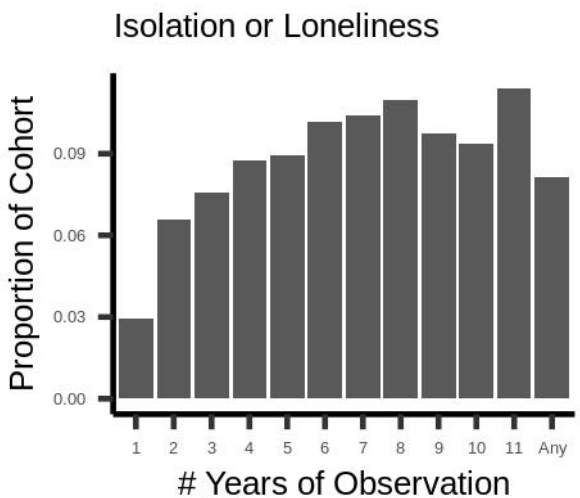
Clinical characteristics

Observation-based period prevalence: Burden from client perspective

11-year period prevalence: Burden of conditions from system planning perspective

Estimates range from 2% (Hepatitis C) to 63% (Chronic musculoskeletal problem)



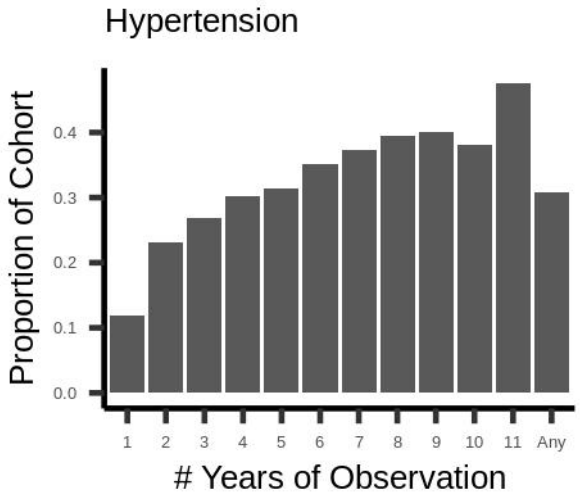


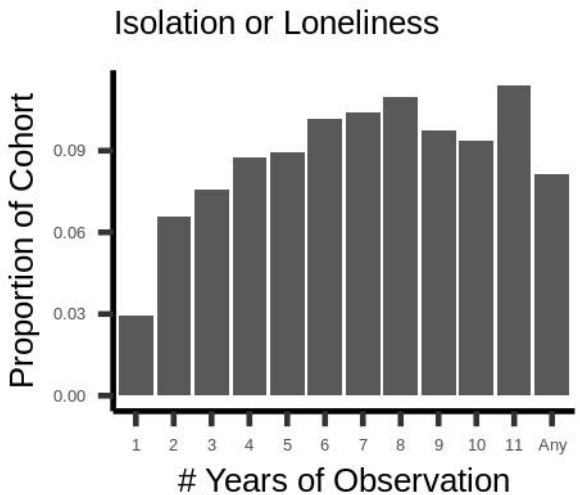
Clinical characteristics

11-year period prevalence: Burden of conditions from system planning perspective

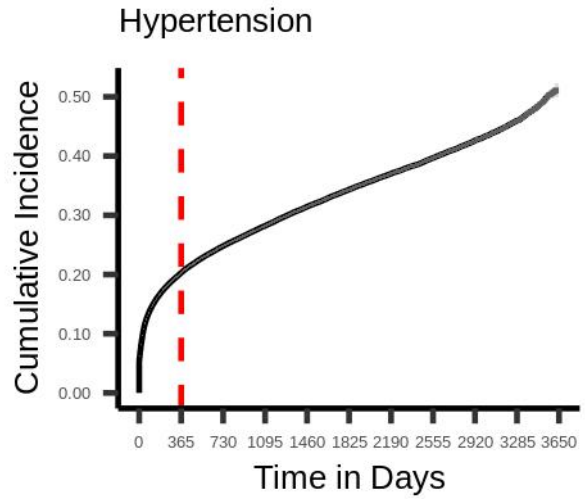
Estimates range from 2% (Hepatitis C) to 63% (Chronic musculoskeletal problem)

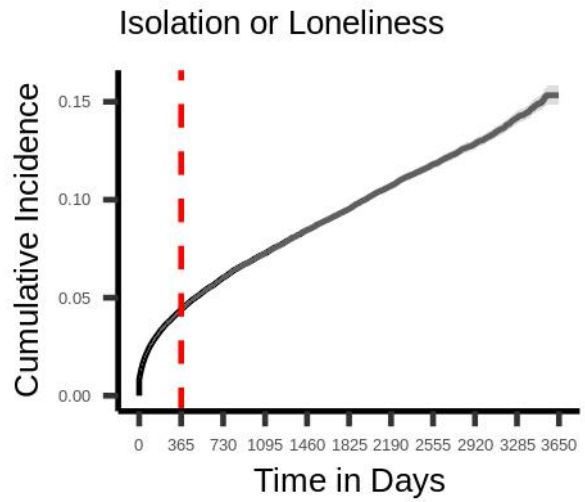
Observation-based period prevalence: Burden from client perspective





Cumulative incidence: Rate of condition indications by observation days





Condition Co-occurence Patterns

11-year period prevalence of 2+ multimorbidity

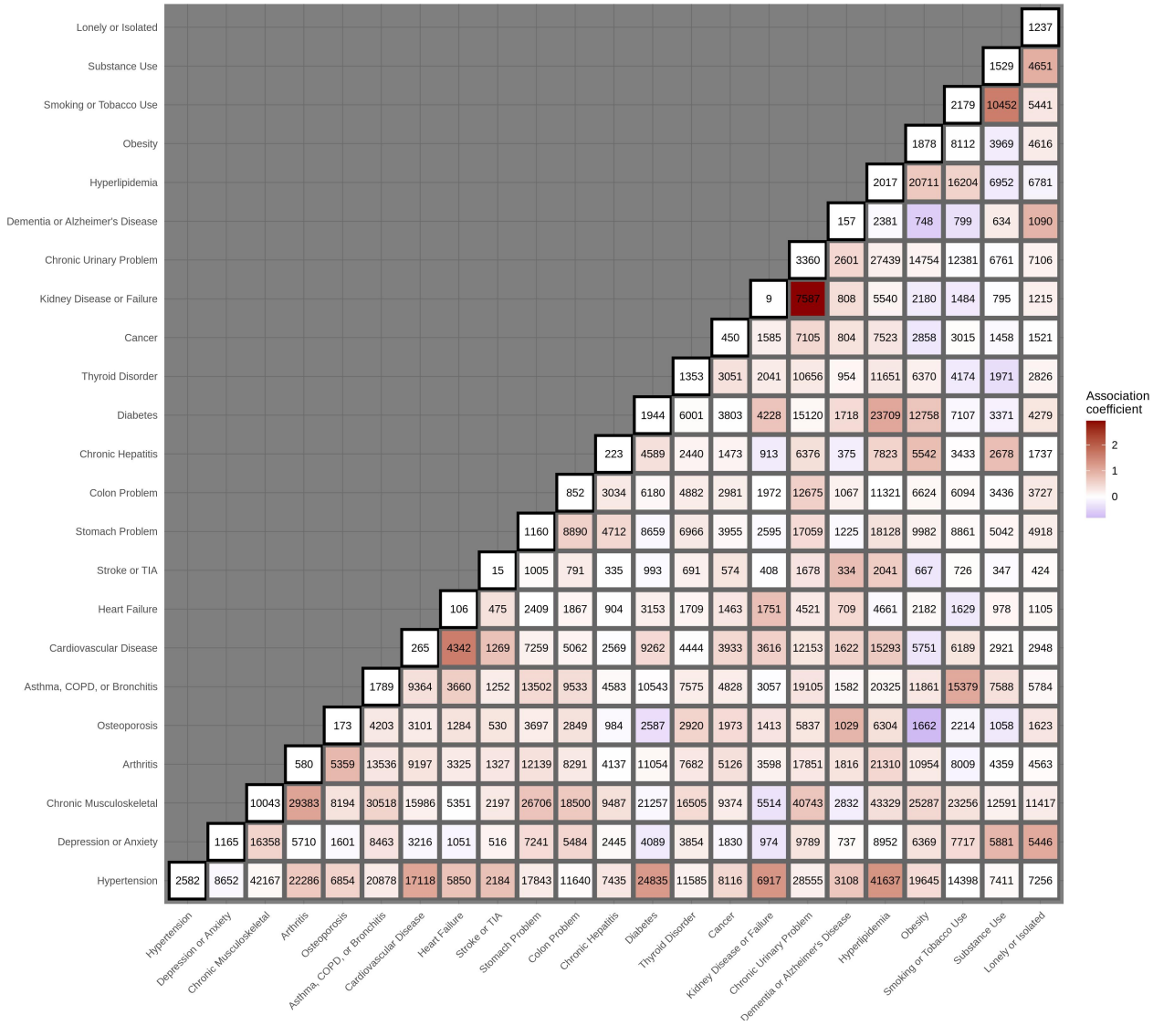
133,704 (81%) clients with 25,346 unique compositions ranging from 1 (<0.1%) to 2840 (2%)

Condition Co-occurence Patterns

Ising model: Tendency for conditions to co-occur

11-year period prevalence of 2+ multimorbidity

133,704 (81%) clients with 25,346 unique compositions ranging from 1 (<0.1%) to 2840 (2%)

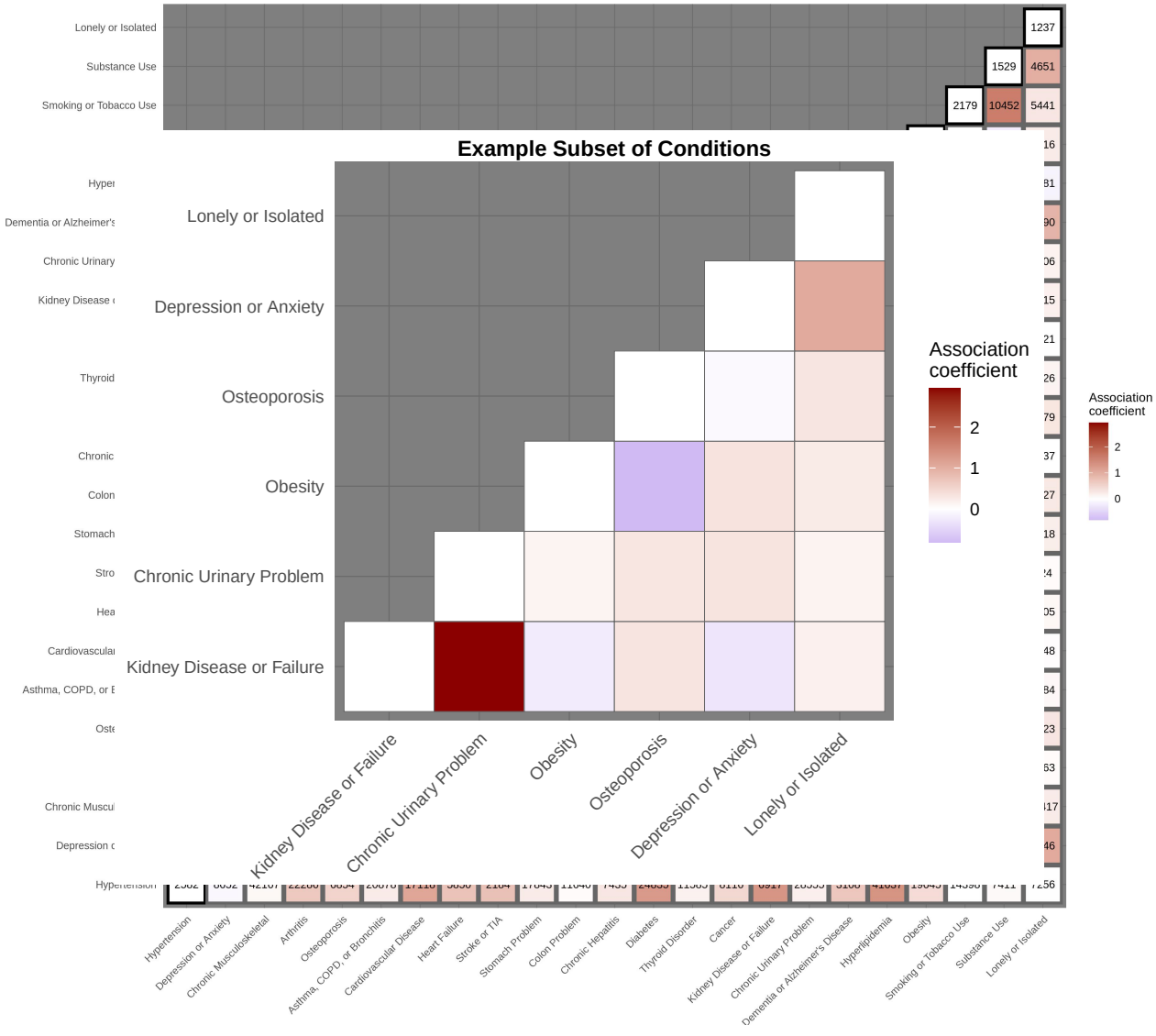


Condition Co-occurence Patterns

Ising model: Tendency for conditions to co-occur

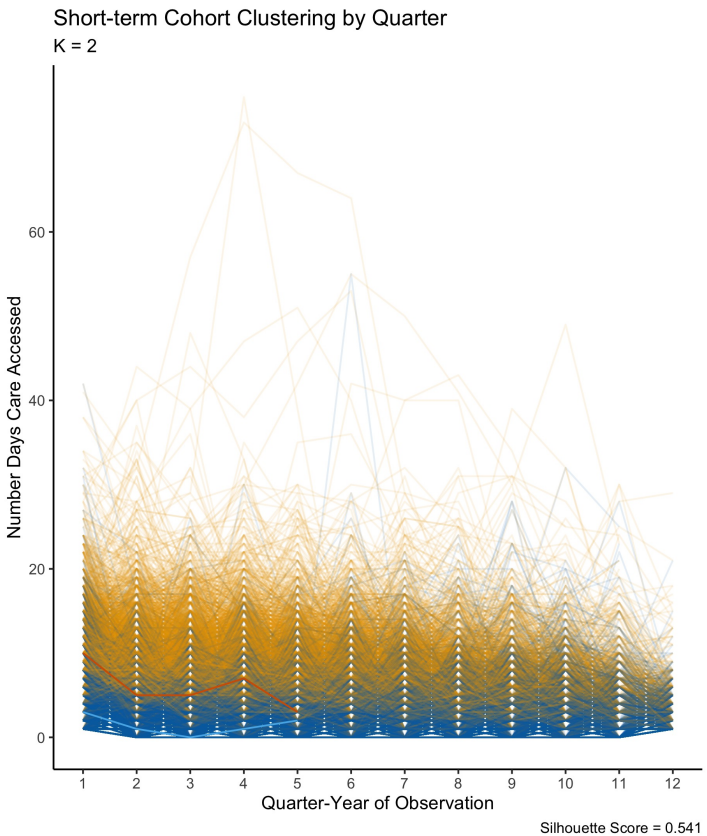
11-year period prevalence of 2+ multimorbidity

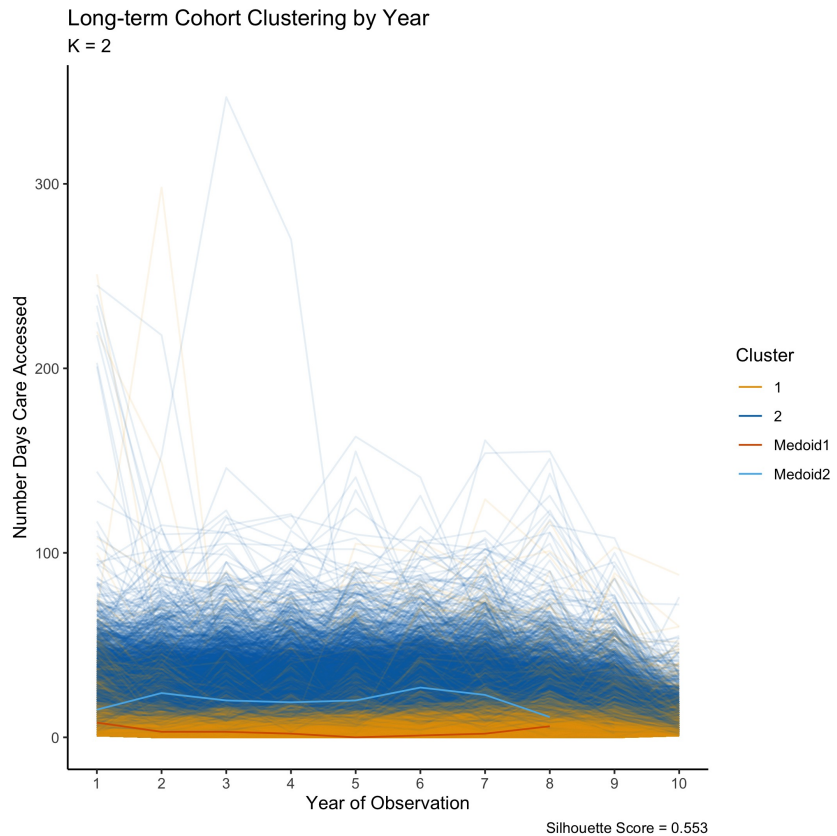
133,704 (81%) clients with 25,346 unique compositions ranging from 1 (<0.1%) to 2840 (2%)



Care access frequency

- K medoids clustering with dynamic time warping distance metric
  - Short-term (2-3yr) & long-term (8-10yr) client sub-cohorts
  - By year and by quarter-year frequencies
  - 2-5 clusters allowed
- Expected "frequent visitor" cluster did not emerge





Conclusions

- The Alliance serves a complex population in terms of sociodemographic, clinical, and healthcare use characteristics
  - Multimorbidity is common but specific profiles are diverse
  - No major separation in care use frequency patterns
  - Care is typically led by physician or nurse practitioners with heterogeneous combinations of other provider types (not shown)
- Unsupervised learning techniques were useful in our population-level descriptions
- Substantive and methodological insights set the stage for future work and learning initiatives

Acknowledgements

- **Collaborators**:
  - Dr. Jennifer Rayner
  - Dr. Merrick Zwarenstein
  - Dr. Dan Lizotte

Questions

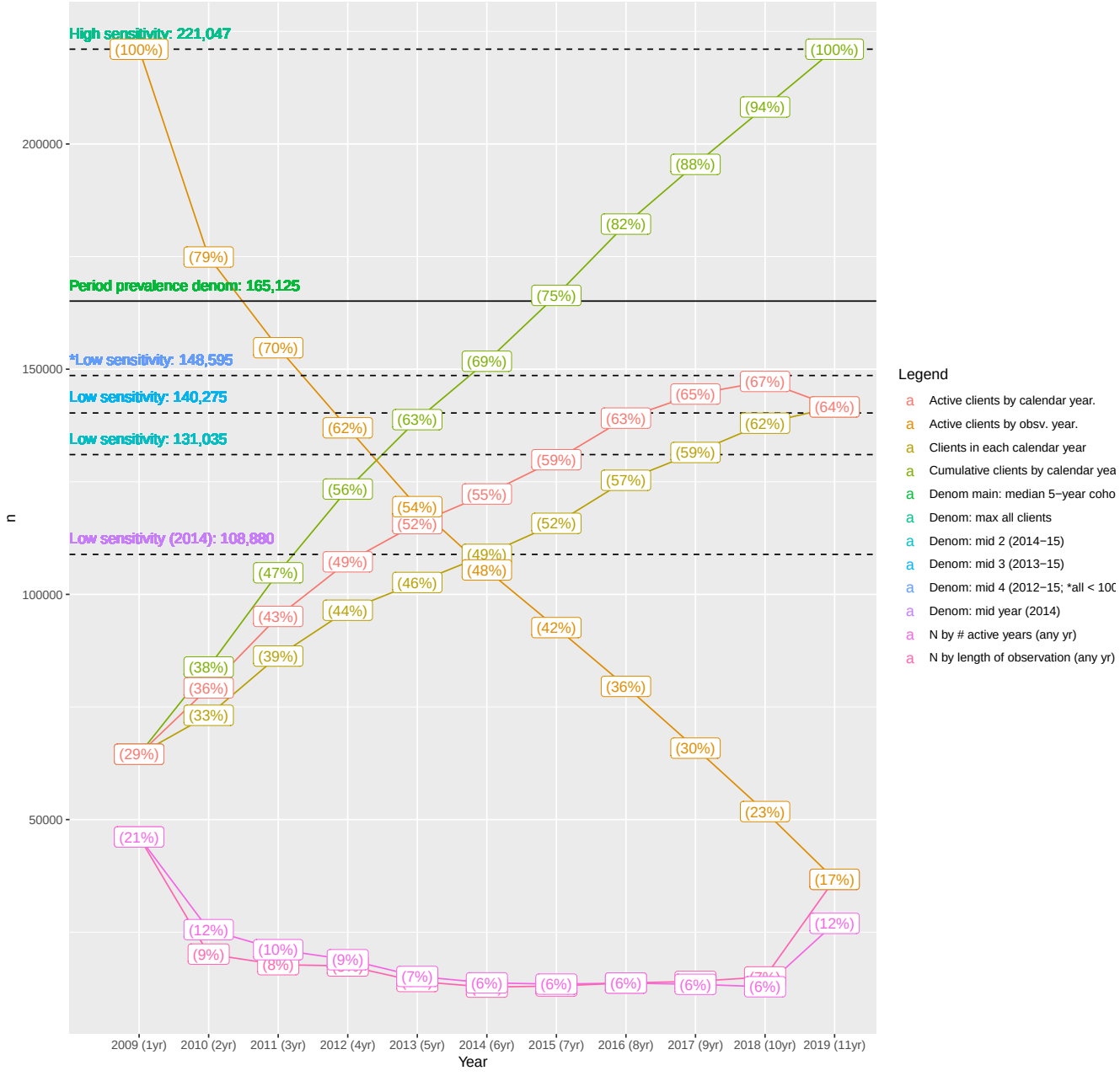
Contact:

- Jaky Kueper, PhD
- **Email**: jkueper@uwo.ca
- **Twitter**: @jk_kueper

Kueper JK, Rayner J, Zwarenstein M, Lizotte D. Describing a complex primary health care population to support future decision support initiatives. *International Journal of Population Data Science*. 2022;7(1). doi:10.23889/ijpds.v7i1.1756

EXTRA SLIDES

Cohort Sizes



Point labels are relative to all clients: % of 221,047.

Horizontal lines are candidate denominators for period prevalence.

Common Conditions

Largest positive associations

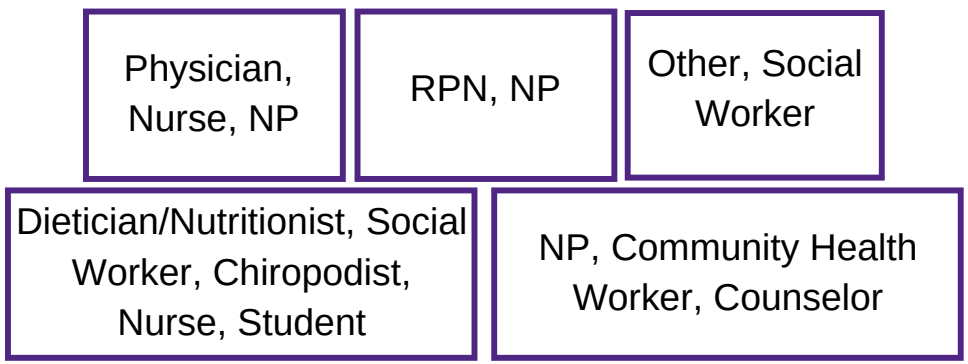
1. Kidney Disease or Failure—Chronic Urinary Problem
2. Smoking or Tobacco Use—Substance Use
3. Cardiovascular Disease—Heart Failure
4. Hypertension—Hyperlipidemia
5. Hypertension—Kidney Disease

Most frequent co-occurring conditions

1. Hyperlipidemia—Chronic Musculoskeletal
2. Hypertension—Chronic Musculoskeletal
3. Hyperlipidemia—Hypertension
4. Chronic Urinary Problem—Chronic Musculoskeletal
5. Asthma or COPD or Chronic Bronchitis—Chronic Musculoskeletal

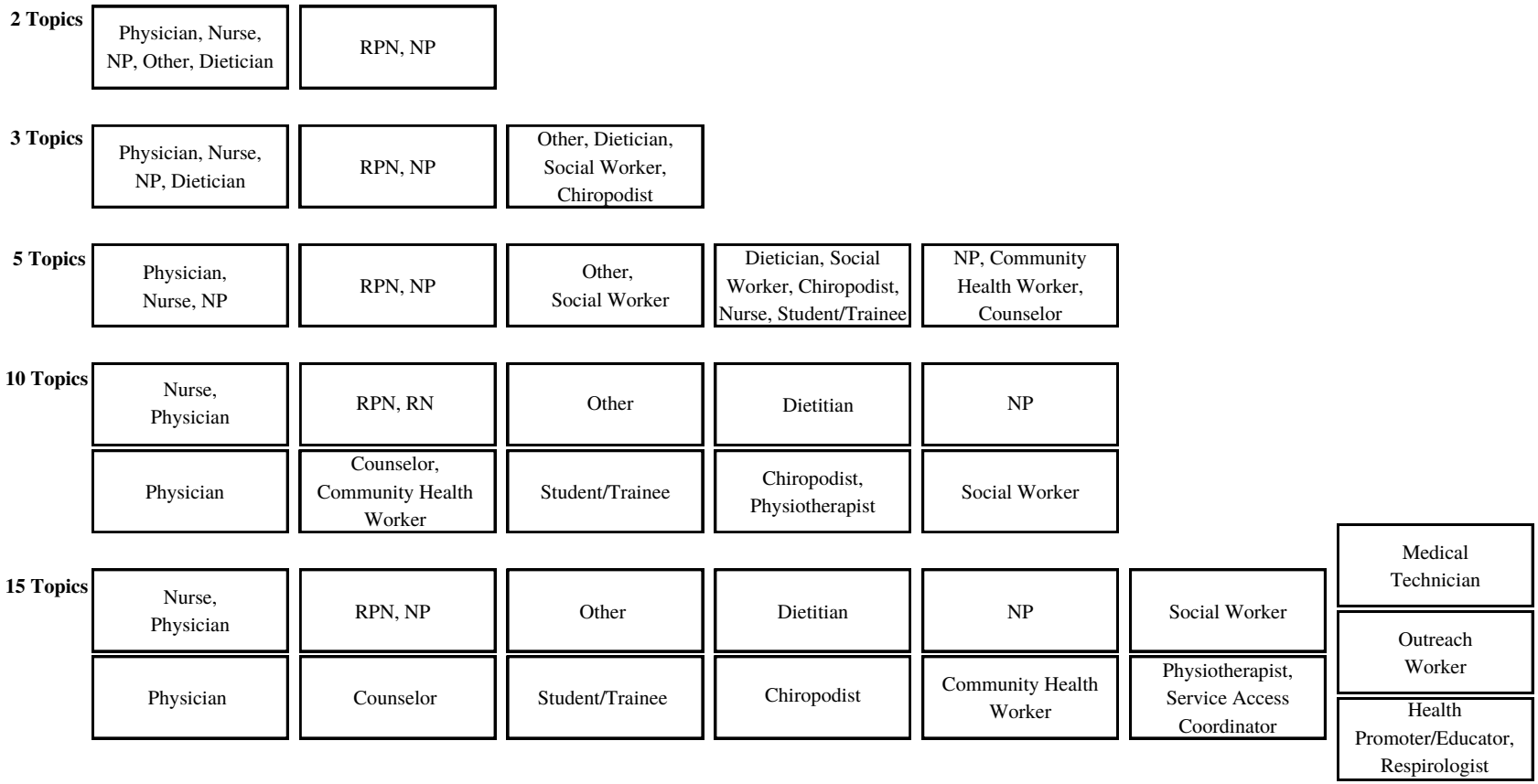
Providers involved in care

- 19,394 unique combinations of 68 provider types
- Non-negative matrix factorization to identify common “ever-seen” and “relative amount seen” teams
- Care is typically led by physician or nurse practitioners with heterogeneous combinations of other provider types
  - End-users for a given decision support tool may differ by client

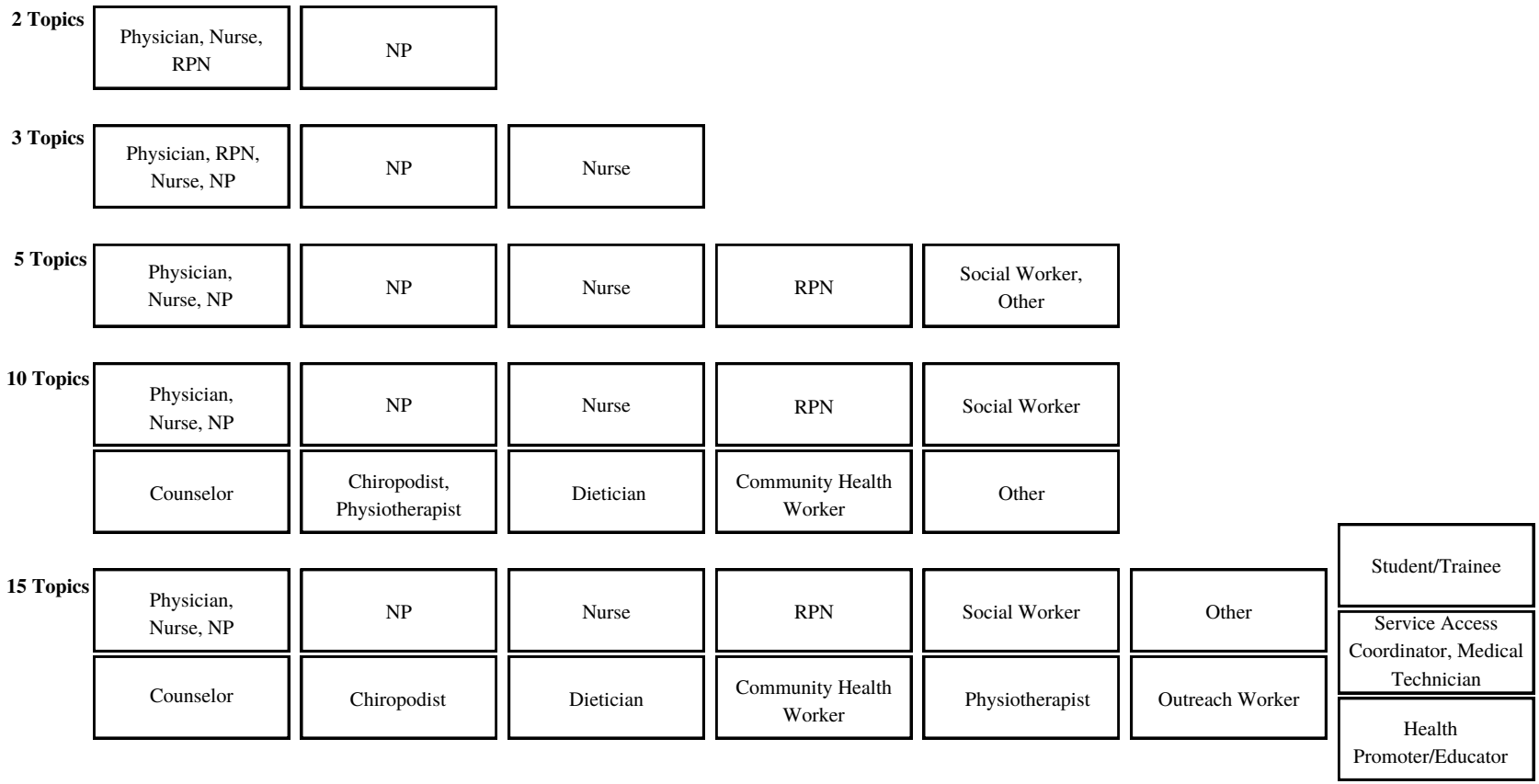


Example: 5 topic "ever-seen" analysis.

Provider Type NMF: Ever-Seen



Provider Type NMF: Relative Amount-Seen



Care Complexity

- **CHC-perspective**
  - 80% of client-visits had a single issue addressed (low complexity)
  - $<$ 1% client-visits had over 5 issues addressed (high complexity)
- **Client-perspective**
  - 17% only ever experienced one issue per visit
  - 11% had at least one visit with over 5 issues addressed
    - Did not tend to also have high care frequency
- Reminder that these analyses only capture what is recorded in EHRs, but these are the data that EHR-based decision support tools would have access to