Lunch 'n' Learn Webinar

Finding Meaning in a Universe of Data:

Exploring Opportunities for Learning Machines to Advance Health Care Planning and Delivery

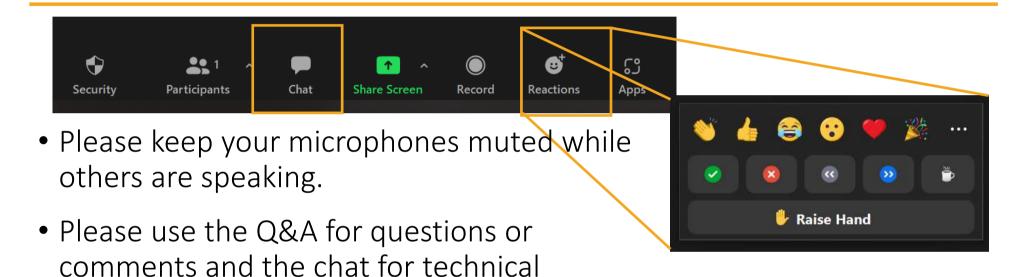
Dr. Jaky Kueper, PhD, Western University October 7, 2022



B

Housekeeping





• Open the "reaction" button to access "raise hand" and other tools.



assistance.



Acknowledgement of Traditional Indigenous Territories

We recognize that the work of the Alliance for Healthier Communities, our members, and the POPLAR Network takes place across what is now called Ontario, on traditional territories of Indigenous people. They have lived here since time immemorial and have deep connections to these lands. We further acknowledge that Ontario is covered by 46 treaties, agreements and land purchases, as well as unceded territories. We are grateful for the opportunity to live, meet and work on this territory.

Ontario continues to be home to vibrant, diverse Indigenous communities who have distinct and specific histories, needs, and assets as well as constitutionally protected and treaty rights. We honour this diversity and respect the knowledge, leadership and governance frameworks within Indigenous communities. In recognition of this, we commit to building allyship relationships with First Nation, Inuit and Métis peoples in order to enhance our knowledge and appreciation of the many histories and voices within Ontario. We also commit to sharing and upholding our responsibilities to all who now live on these lands, the land itself, and the resources that make our lives possible.





Welcome & Introduction



Dr. Jaky Kueper, PhD Postdoctoral Fellow, Western University







Finding Meaning in a Universe of Data: Exploring Opportunities for Learning Machines to Advance Health Care Planning and Delivery

Alliance Lunch 'n' Learn Webinar

Jaky Kueper, PhD

Postdoctoral Associate, Department of Computer Science, Western University

October 7, 2022

Getting to Know You



Poll Question

What organizational setting do you primarily work in?

Poll Question

What is your role?

Poll Question

What are you most excited about learning today?

Motivation



- Primary health care is the foundation of the healthcare system
 - Broad scope of practice
 - Team-based care across the lifespan
- ► **Challenge**: How to provide the best care possible, 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
- ▶ Potential support: deriving insights and value through careful analysis of electronic health records
 - Learn from what was done in the past to inform the future

Opportunity: Data to Support Care Delivery

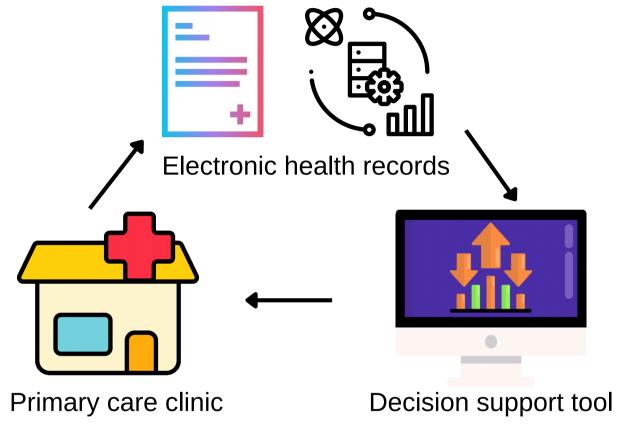


- ► Increases in "everyday" data and computing resources
 - Electronic health records from healthcare delivery
- ► Potential to analyze these data to increase knowledge and improve care, e.g., personalized decision support tools
 - ► Artificial intelligence (AI) and machine learning are powerful tools to help us accomplish this
- Many opportunities to ask questions, learn, and work together towards innovations



Decision Support Tool Development





Example Learning Health System (LHS) paradigm

Decision support tools may or may not rely on Al to function.

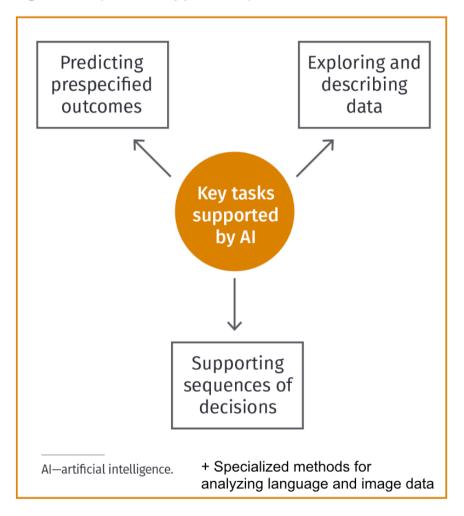
Artificial Intelligence (AI)



One part of a larger digital transformation

- Interdisciplinary field focused on computers performing humanlike "intelligent" tasks
- Machine learning subfield: Math based on data and human decisions

Figure 1. Key tasks supported by Al



Kueper JK. Primer for artificial intelligence in primary care. *Canadian Family Physician*. 2021;67:889-893.



Al has the potential to *support* existing relationships and functions in primary health care



All has the potential to *support* existing relationships and functions in primary health care

- ► Risk prediction tool running in the background of an EHR system
 - Generate alerts when someone is at high risk of a negative outcome



All has the potential to *support* existing relationships and functions in primary health care

- ► Risk prediction tool running in the background of an EHR system
 - ► Generate alerts when someone is at high risk of a negative outcome
- ► Tools to perform clerical tasks, e.g., note-taking
 - Mitigate burnout and free up time for client interactions



All has the potential to *support* existing relationships and functions in primary health care

- Risk prediction tool running in the background of an EHR system
 - Generate alerts when someone is at high risk of a negative outcome
- ► Tools to perform clerical tasks, e.g., note-taking
 - Mitigate burnout and free up time for client interactions
- Personalized care decision support tool, e.g., synthesize relevant research literature, clinical guidelines, and EHR data
 - Provides recommendations for care providers and clients to discuss



All has the potential to *support* existing relationships and functions in primary health care

- Risk prediction tool running in the background of an EHR system
 - ► Generate alerts when someone is at high risk of a negative outcome
- ► Tools to perform clerical tasks, e.g., note-taking
 - Mitigate burnout and free up time for client interactions
- Personalized care decision support tool, e.g., synthesize relevant research literature, clinical guidelines, and EHR data
 - Provides recommendations for care providers and clients to discuss

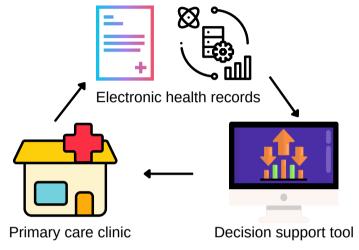
Poll Question

What is your level of familiarity with AI?

Decision Support Tool Development Context



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



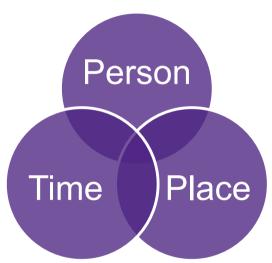
Example Learning Health System (LHS) paradigm

Can we use **descriptive epidemiology** to help guide early stages of decision support tool development?

Descriptive Epidemiology

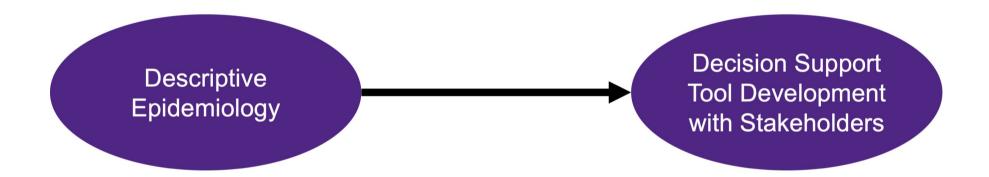


- Understand health-related states and needs of a population
 - ► Who? What? When? Where? How?
- Valuable tool in public health practice and system planning
- Population-level descriptions may complement stakeholder expertise
 - Problem refinement and selection
 - Understanding of data availability and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



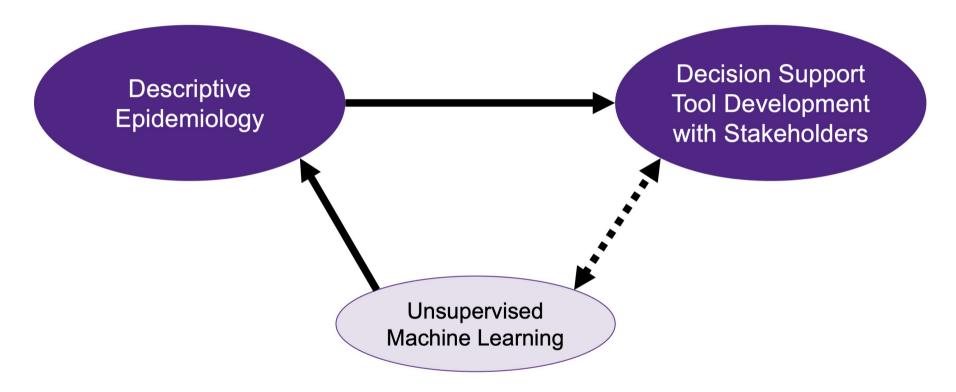
Proposed Role for Descriptive Epidemiology





Proposed Role for Descriptive Epidemiology



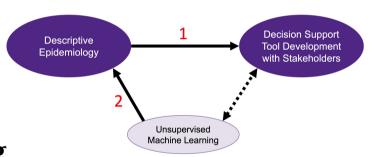


► To properly understand complex populations, both simple statistical techniques and techniques that can capture more complex patterns may be useful.

Study Goals



- 1. To describe an ongoing primary care population for the purpose of supporting future initiatives, including the development of decision support tools.
- 2. To demonstrate how simple statistical and unsupervised learning techniques, applied with an epidemiological lens, can be used to describe complex populations.



Cohort Overview

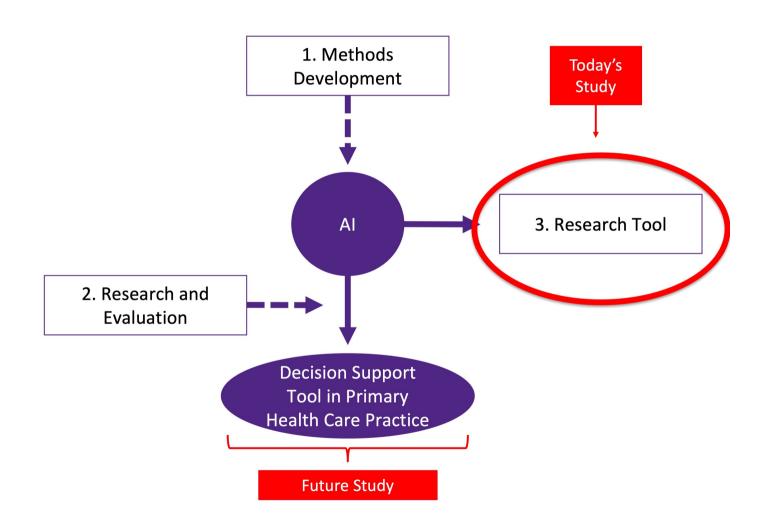


- Clinical, Health Care Use, and Sociodemographic characteristics
- ► Electronic health record data
- ► Base cohort eligibility (n=221,047)
 - Ongoing primary care clients
 - ► 1+ encounter at a Community Health Centre (CHC) in 2009-2019
 - ► At least 18 years old

Aside: Al as a Tool to Capture Complexity



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



Warning: Lots of Information Coming!



General Strategy

- Take it one slide at a time.
- Polls embedded throughout.
- Put ideas and questions in the chat!

Reflection Questions

- 1. Does each population-level description match your expectations?
- 2. What other questions about the data do you have?
- 3. Can you imagine any areas to develop decision support tools?



Clinical Characteristics

11-Year Period Prevalence



Poll Question

Are these prevalence estimates surprising to you?

- 1. Some are higher than expected
- 2. Some are lower than expected
- 3. Both 1 & 2
- 4. About what I expected
- 5. I don't have prior expectations

Variable	All CHCs	Urban At Risk CHCs
Number of clients	165125	27256
Multimorbidity 2+	133704 (80.97)	24129 (88.53)
Multimorbidity 3+	103172 (62.48)	19237 (70.58)
Hypertension Depression or Anxiety Chronic Musculoskleletal Arthritis Osteoporosis	68177 (41.29) 23828 (14.43) 104304 (63.17) 37201 (22.53) 11462 (6.94)	12304 (45.14) 5533 (20.3) 18842 (69.13) 6906 (25.34) 1950 (7.15)
Asthma, COPD, or Chronic Bronchitis	43837 (26.55)	9190 (33.72)
Cardiovascular Disease	23311 (14.12)	4673 (17.14)
Heart Failure	7994 (4.84)	1564 (5.74)
Stroke or TIA	2967 (1.8)	585 (2.15)
Stomach Problem	36175 (21.91)	7620 (27.96)
Colon Problem Chronic Hepatitis Diabetes Thyroid Disorder Any Cancer	24949 (15.11) 13288 (8.05) 35704 (21.62) 24793 (15.01) 14024 (8.49)	4974 (18.25) 2954 (10.84) 6912 (25.36) 4217 (15.47) 2636 (9.67)
Kidney Disease or Failure	8290 (5.02)	1555 (5.71)
Chronic Urinary Problem	59677 (36.14)	11131 (40.84)
Dementia or AD	4776 (2.89)	898 (3.29)
Hyperlipidemia	67175 (40.68)	11659 (42.78)
Obesity	38408 (23.26)	6455 (23.68)
Hepatitis C	2436 (1.48)	1173 (4.3)
Smoking or Tobacco Use	37355 (22.62)	9597 (35.21)
Substance Use	20853 (12.63)	7508 (27.55)
Social Isolation or Loneliness	17947 (10.87)	5149 (18.89)

Notes: Denominator was the approximated average population size across all years (2009-2019).

Clinical Characteristics: Different Perspectives Western



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

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

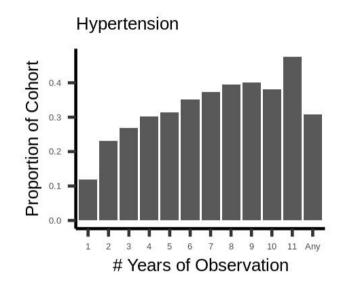
Clinical Characteristics: Different Perspectives

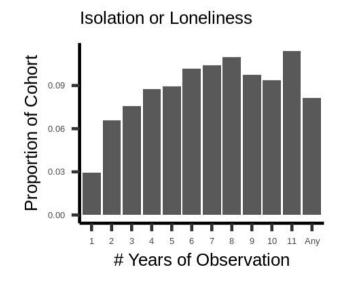


Observation-based period prevalence: Burden from client perspective

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

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





Clinical Characteristics: Different Perspectives

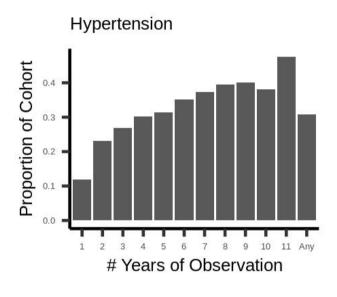


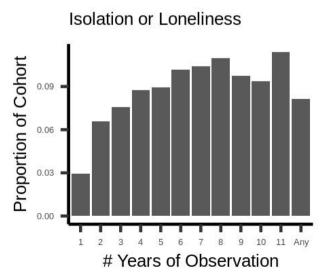
Observation-based period prevalence: Burden from client perspective

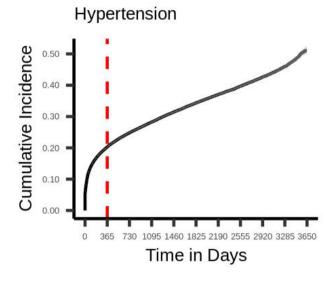
Cumulative incidence: Rate of condition indications by observation days

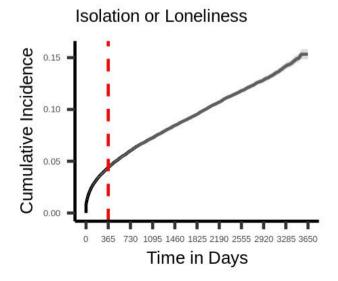
11-year period prevalence:
Burden of conditions from system planning perspective

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









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%)

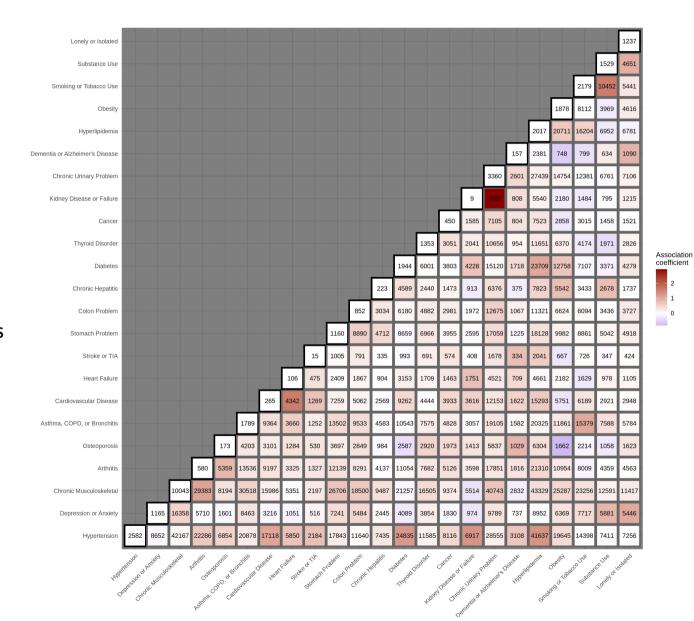
Condition Co-occurence Patterns



Ising model: Tendency for conditions to co-occur

11-year period prevalence of 2+ multimorbidity

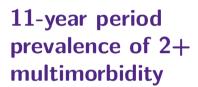
133,704 (81%) clients with 25,346 unique compositions ranging from 1 (<0.1%) to 2840 (2%)



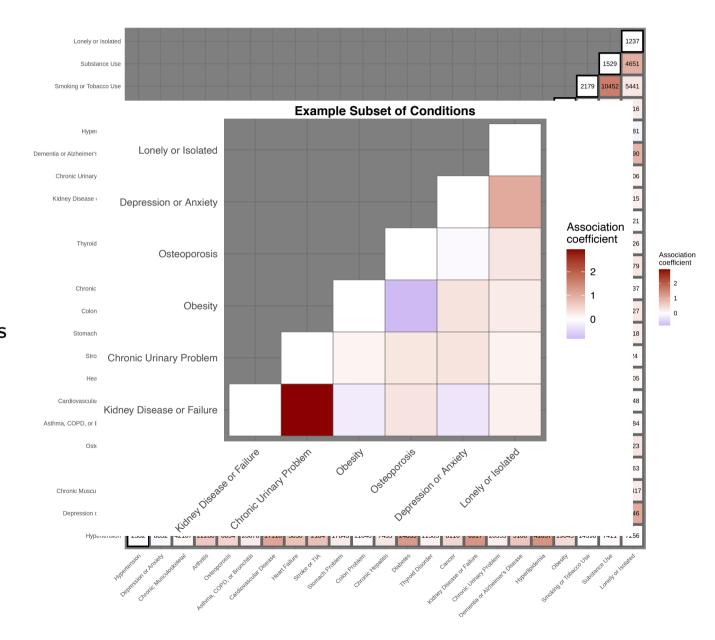
Condition Co-occurence Patterns



Ising model: Tendency for conditions to co-occur



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



Common Comorbidities: Different Perspectives Western



Most frequent co-occurring conditions (Simple counts)

- 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

Largest positive pairwise associations (Ising model)

- 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

Poll Question

Which type of measure do you think is more important to guide focus areas for decision support tools?

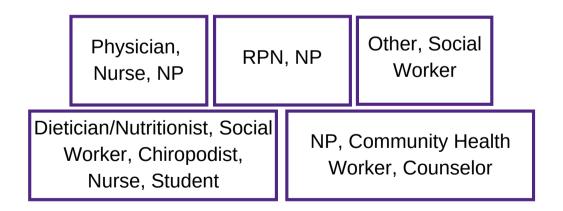


Health Care Use Characteristics

Provider Teams



- ▶ 19,394 unique combinations of 68 provider types
- ► We used topic modelling to identify common "ever-seen" and "relative amount seen" teams
- Care amount 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 Teams: Ever-Seen



2 Topics	Physician, Nurse, NP, Other, Dietician	RPN, NP					
3 Topics	Physician, Nurse, NP, Dietician	RPN, NP	Other, Dietician, Social Worker, Chiropodist				
5 Topics	Physician, Nurse, NP	RPN, NP	Other, Social Worker	Dietician, Social Worker, Chiropodist, Nurse, Student/Trainee	NP, Community Health Worker, Counselor		
10 Topics	Nurse, Physician	RPN, RN	Other	Dietitian	NP		
	Physician	Counselor, Community Health Worker	Student/Trainee	Chiropodist, Physiotherapist	Social Worker		
15 Topics	Nurse, Physician	RPN, NP	Other	Dietitian	NP	Social Worker	Medical Technician
	Physician	Counselor	Student/Trainee	Chiropodist	Community Health Worker	Physiotherapist, Service Access Coordinator	Outreach Worker Health Promoter/Educator,
•							Respirologist

Poll Question

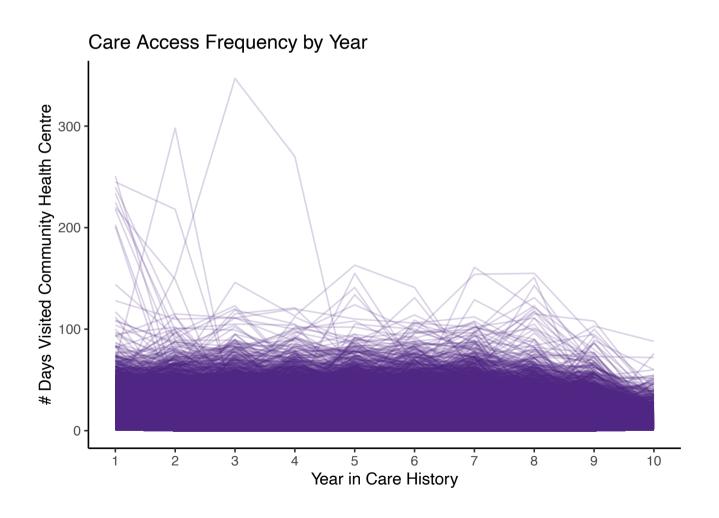
Do any of the common provider type combinations surprise you?

1) Some combinations are unexpected 2) Some expected combinations are missing 3) 1 & 2 4) Looks about right 5) Not sure!

Care Access Frequency



- Number of days that clients received care by year and by quarter-year
- We did not identify distinct groups of people or "frequent visitor" cut-offs



Care Complexity



► CHC-perspective

- ▶ 80% of client-visits had a single issue addressed (low complexity)
- ightharpoonup < 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



Sociodemographic Characteristics

Sociodemographic Characteristics (Last Poll!)



Variable	Values	All (n=221047)	UAR (n=35998)	MM (n=103172)
Age in 2015	25-44	101151 (45.76)	15516 (43.10)	24888 (24.12)
	45-64	82501 (37.32)	14976 (41.60)	49560 (48.03)
	65+	37395 (16.92)	5506 (15.30)	28724 (27.84)
Rural Geography	Rural	49275 (22.29)	6131 (17.03)	26818 (25.99)
	Urban	167728 (75.88)	28538 (79.28)	75011 (72.7)
	Missing	4044 (1.83)	1329 (3.69)	1343 (1.3)
Sex	Female	127070 (57.49)	18699 (51.94)	59946 (58.1)
	Male	93294 (42.21)	17151 (47.64)	43124 (41.8)
	Other	331 (0.15)	43 (0.12)	19 (0.02)
	Missing	352 (0.16)	105 (0.29)	83 (0.08)
Primary Language	English	167163 (75.62)	31658 (87.94)	79599 (77.15)
	French	22547 (10.2)	944 (2.62)	11091 (10.75)
	Other	26847 (12.15)	2948 (8.19)	10710 (10.38)
	Missing	4490 (2.03)	448 (1.24)	1772 (1.72)
Highest Level of Education	Post-secondary or equivalent	84888 (38.4)	12056 (33.49)	35763 (34.66)
	Secondary or equivalent	61831 (27.97)	11783 (32.73)	32617 (31.61)
	Less than high school	18941 (8.57)	3266 (9.07)	10618 (10.29)
	Other	8507 (3.85)	719 (2)	4078 (3.95)
	Do not know	4860 (2.2)	1318 (3.66)	2350 (2.28)
	Prefer not to answer	2950 (1.33)	422 (1.17)	1585 (1.54)
	Missing	39070 (17.67)	6434 (17.87)	16161 (15.66)
Household Income	\$0 to \$14,999	40519 (18.33)	8729 (24.25)	17757 (17.21)
	\$15,000 to \$24,999	21102 (9.55)	3555 (9.88)	11081 (10.74)
	\$25,000 to \$39,999	20877 (9.44)	2988 (8.3)	10736 (10.41)
	\$40,000 to \$59,999	17245 (7.8)	2421 (6.73)	8671 (8.4)
	\$60,000 or more	28494 (12.89)	3862 (10.73)	12868 (12.47)
	Do not know	15408 (6.97)	2658 (7.38)	6264 (6.07)
	Prefer not to answer	27621 (12.5)	4130 (11.47)	14890 (14.43)
	Missing	49781 (22.52)	7655 (21.27)	20905 (20.26)
Stable Residence	True	199349 (90.18)	28227 (78.41)	90479 (87.7)
Food Insecurity	True	10985 (4.97)	2947 (8.19)	7323 (7.1)

Note:

Conclusions



- ► Community Health Centres serve a complex population in terms of sociodemographic, clinical, and healthcare use characteristics
 - Multimorbidity is common but specific profiles are diverse
 - Ongoing primary care is typically led by a physician or nurse practitioner with heterogeneous combinations of other provider types
 - No major separation in care use frequency patterns
- ► Al techniques were useful to explore comorbidity patterns and care provider teams

Reflection Questions

- 1. Does each population-level description match your expectations?
- 2. What other questions about the data do you have?
- 3. Can you imagine any areas to develop decision support tools?

Acknowledgements



Research Team Collaborators:

- ▶ Dr. Dan Lizotte
- ▶ Dr. Jennifer Rayner
- ▶ Dr. Merrick Zwarenstein







EXTRA SLIDES

Sociodemographic Characteristics: A



Variable	Values	All	UAR	MM
Number of clients		221047	35998	103172
Age in 2015	25-34	55505 (25.11)	7976 (22.16)	9346 (9.06)
	35-44	45646 (20.65)	7540 (20.95)	15542 (15.06)
	45-54	44653 (20.20)	8186 (22.74)	23982 (23.24)
	55-64	37848 (17.12)	6790 (18.86)	25578 (24.79)
	65-74	23162 (10.48)	3644 (10.12)	17780 (17.23)
	75+	14233 (6.44)	1862 (5.17)	10944 (10.61)
Rural Geography Residence	Rural	49275 (22.29)	6131 (17.03)	26818 (25.99)
	Urban	167728 (75.88)	28538 (79.28)	75011 (72.70)
	Missing	4044 (1.83)	1329 (3.69)	1343 (1.30)
Sex	Female	127070 (57.49)	18699 (51.94)	59946 (58.1)
	Male	93294 (42.21)	17151 (47.64)	43124 (41.80)
	Other	331 (0.15)	43 (0.12)	19 (0.02)
	Missing	352 (0.16)	105 (0.29)	83 (0.08)
Gender	Female	41352 (18.71)	5509 (15.30)	21831 (21.16)
	Gender Diverse	340 (0.15)	112 (0.31)	144 (0.14)
	Male	29366 (13.28)	4585 (12.74)	14733 (14.28)
	Prefer not to answer	1001 (0.45)	51 (0.14)	376 (0.36)
	Missing	148988 (67.40)	25741 (71.51)	66088 (64.06)
Sexual Orientation	Bisexual Gay Heterosexual Lesbian Queer Two-Spirit Other Do not know Prefer not to answer Missing	1578 (0.71) 708 (0.32) 57065 (25.82) 485 (0.22) 323 (0.15) 128 (0.06) 246 (0.11) 924 (0.42) 7561 (3.42) 152029 (68.78)	285 (0.79) 192 (0.53) 8447 (23.47) 70 (0.19) 34 (0.09) 80 (0.22) 34 (0.09) 201 (0.56) 877 (2.44) 25778 (71.61)	690 (0.67) 306 (0.30) 29105 (28.21) 244 (0.24) 91 (0.09) 61 (0.06) 143 (0.14) 485 (0.47) 4078 (3.95) 67969 (65.88)

Legend: MM = Multimorbidity; UAR = Urban At Risk CHC

Sociodemographic Characteristics: B



Variable	Values	All	UAR	MM
Number of clients		221047	35998	103172
	Post-secondary or equivalent	84888 (38.40)	12056 (33.49)	35763 (34.66)
	Secondary or equivalent	61831 (27.97)	11783 (32.73)	32617 (31.61)
	Less than high school	18941 (8.57)	3266 (9.07)	10618 (10.29)
Highest Level of Education	Other	8507 (3.85)	719 (2.00)	4078 (3.95)
	Do not know	4860 (2.20)	1318 (3.66)	2350 (2.28)
	Prefer not to answer	2950 (1.33)	422 (1.17)	1585 (1.54)
	Missing	39070 (17.67)	6434 (17.87)	16161 (15.66)
	English	167163 (75.62)	31658 (87.94)	79599 (77.15)
Drimon, Longue	French	22547 (10.20)	944 (2.62)	11091 (10.75)
Primary Language	Other	26847 (12.15)	2948 (8.19)	10710 (10.38)
	Missing	4490 (2.03)	448 (1.24)	1772 (1.72)
	Black	8861 (4.01)	725 (2.01)	3757 (3.64)
	East/SouthEast Asian	3739 (1.69)	484 (1.34)	1545 (1.50)
	Indigenous	2944 (1.33)	1577 (4.38)	1641 (1.59)
	Latino	4350 (1.97)	206 (0.57)	1708 (1.66)
	Middle Eastern	2046 (0.93)	344 (0.96)	838 (0.81)
Race and Ethnicity	Other	567 (0.26)	148 (0.41)	306 (0.3)
	South Asian	3597 (1.63)	323 (0.90)	1852 (1.80)
	White	38464 (17.40)	4531 (12.59)	21504 (20.84)
	Do not know	838 (0.38)	151 (0.42)	487 (0.47)
	Prefer not to answer	2649 (1.20)	261 (0.73)	1513 (1.47)
	Missing	152992 (69.21)	27248 (75.69)	68021 (65.93)

Legend: MM = Multimorbidity; UAR = Urban At Risk CHC

Sociodemographic Characteristics: C

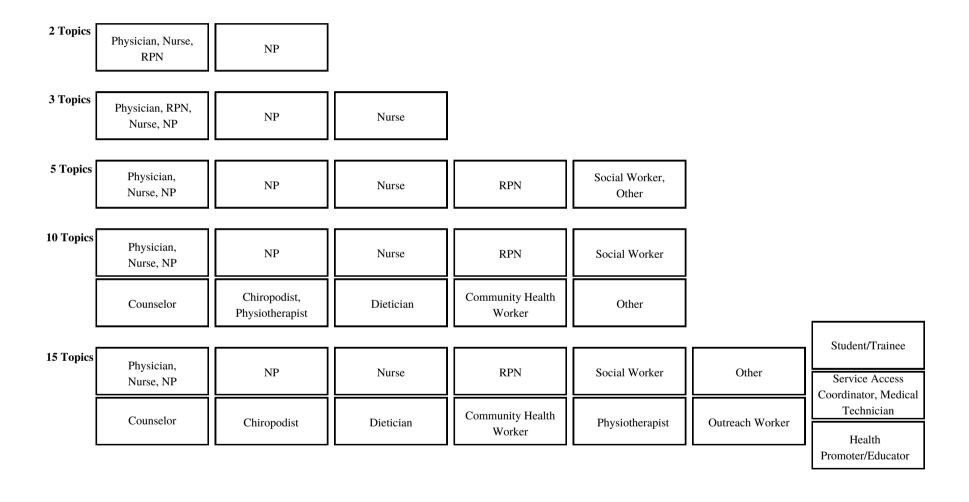


Variable	Values	All	UAR	MM
Number of clients		221047	35998	103172
	0-5 years	13654 (6.18)	1191 (3.31)	3047 (2.95)
Years since Arrival in Canada	6+ years	51815 (23.44)	4940 (13.72)	22722 (22.02)
	None recorded	155578 (70.38)	29867 (82.97)	77403 (75.02)
	\$0 to \$14,999	40519 (18.33)	8729 (24.25)	17757 (17.21)
	\$15,000 to \$24,999	21102 (9.55)	3555 (9.88)	11081 (10.74)
	\$25,000 to \$39,999	20877 (9.44)	2988 (8.3)	10736 (10.41)
Household Income	\$40,000 to \$59,999	17245 (7.80)	2421 (6.73)	8671 (8.40)
Household income	\$60,000 or more	28494 (12.89)	3862 (10.73)	12868 (12.47)
	Do not know	15408 (6.97)	2658 (7.38)	6264 (6.07)
	Prefer not to answer	27621 (12.50)	4130 (11.47)	14890 (14.43)
	Missing	49781 (22.52)	7655 (21.27)	20905 (20.26)
	Couple with children	53398 (24.16)	6759 (18.78)	20713 (20.08)
	Couple without child	39664 (17.94)	5945 (16.51)	22950 (22.24)
	Extended Family	7632 (3.45)	1123 (3.12)	3581 (3.47)
	Grandparents with Grandchild(ren)	1746 (0.79)	247 (0.69)	1183 (1.15)
	Siblings	1622 (0.73)	250 (0.69)	669 (0.65)
Household Composition	Single Parent	14445 (6.53)	2527 (7.02)	6348 (6.15)
	Sole Member	32782 (14.83)	7445 (20.68)	18597 (18.03)
	Unrelated housemates	8622 (3.90)	1567 (4.35)	2849 (2.76)
	Other	8913 (4.03)	1476 (4.10)	4202 (4.07)
	Do not know	2475 (1.12)	643 (1.79)	1279 (1.24)
	Prefer not to answer	3727 (1.69)	491 (1.36)	1927 (1.87)
	Missing	46021 (20.82)	35998 1191 (3.31) 4940 (13.72) 29867 (82.97) 8729 (24.25) 3555 (9.88) 2988 (8.3) 2421 (6.73) 3862 (10.73) 2658 (7.38) 4130 (11.47) 7655 (21.27) 6759 (18.78) 5945 (16.51) 1123 (3.12) 247 (0.69) 250 (0.69) 2527 (7.02) 7445 (20.68) 1567 (4.35) 1476 (4.10) 643 (1.79)	18874 (18.29)
Stable Residence	True	199349 (90.18)	28227 (78.41)	90479 (87.70)
Food Insecurity	True	10985 (4.97)	2947 (8.19)	7323 (7.10)

Legend: MM = Multimorbidity; UAR = Urban At Risk CHC

Provider Type Teams: Relative Amount-Seen Western





Questions/Discussion

Please raise your hand and we'll call on you to ask your question aloud.

If you prefer, you may type questions into the Q&A panel and we will moderate them.



Want more? Let's take a deeper dive together!



Brainstorming Session: How can we use health care data to create decision-support tools?

Friday, October 21, 2022

12:00-1:00 pm

Info: Sara.Bhatti@AllianceON.org





Thank you!

For follow-up questions:

jkueper@uwo.ca LHS@AllianceON.org

