Using statistical methods and reproducible tools to gain new insights from biomedical and public health data

Ariel Mundo Ortiz

Centre de Recherches Mathematiques, Université de Montréal

MfPH Next Generation Seminar Series

3/15/23

Introduction

- Data is the core of research. However, data is not information, as it needs to be processed before we can get information from it.

Introduction

- Data is the core of research. However, data is not information, as it needs to be processed before we can get information from it.
- This is specially true in the case of health research: public health, or biomedical data can be complex, and decisions along the analysis can result in different interpretations.

Introduction

- Data is the core of research. However, data is not information, as it needs to be processed before we can get information from it.
- This is specially true in the case of health research: public health, or biomedical data can be complex, and decisions along the analysis can result in different interpretations.
- In this talk I will focus on two examples that showcase how we can get more insight from data

The Case of Public Health Data

COVID-19

- COVID-19 vaccination has been an important component of public health strategies aimed at managing the pandemic.

[1] Nafilyan et al. 2021.
[2] Gerretsen et al. 2021.

COVID-19

- COVID-19 vaccination has been an important component of public health strategies aimed at managing the pandemic.
- However, COVID-19 vaccination has not been equal across different population segments.

[1] Nafilyan et al. 2021.
[2] Gerretsen et al. 2021.

COVID-19

- COVID-19 vaccination has been an important component of public health strategies aimed at managing the pandemic.
- However, COVID-19 vaccination has not been equal across different population segments.

[1] Nafilyan et al. 2021.
[2] Gerretsen et al. 2021.

COVID-19

- COVID-19 vaccination has been an important component of public health strategies aimed at managing the pandemic.
- However, COVID-19 vaccination has not been equal across different population segments.
- Individuals with lower income, and those belonging to a racial/ethnic minority have had lower vaccination uptake[1,2].

[1] Nafilyan et al. 2021.
[2] Gerretsen et al. 2021.

COVID-19: The Case of Ontario

- The Fields Institute collected some very nice data regarding COVID-19 vaccination in Ontario, the *Survey of COVID-19 related Behaviours and Attitudes*.

COVID-19: The Case of Ontario

- The Fields Institute collected some very nice data regarding COVID-19 vaccination in Ontario, the *Survey of COVID-19 related Behaviours and Attitudes*.
  - The survey ran between late 2021 and early 2022 and collected socio-demographic information along with self-reported vaccination status ("Have you received the first dose of the Covid vaccine?")

COVID-19: The Case of Ontario

Table 1: Selected socio-economic factors from the survey

Variable	Levels
Age group	16-34,35-54,55 and over
Income bracket (CAD)	under 25,000, 25,000-59,999, 60,000 and above
Race/ethnicity	Arab/Middle Eastern, Black, East Asian/Pacific Islander, Indigenous, Latin American, Mixed, South Asian, White Caucasian, Other

COVID-19: The Case of Ontario

- Other studies have analyzed the dependency on vaccination status using socio-economic data.

COVID-19: The Case of Ontario

- Other studies have analyzed the dependency on vaccination status using socio-economic data.
- We could do the same, but what other information could we get from this data?

COVID-19: The Case of Ontario

- Other studies have analyzed the dependency on vaccination status using socio-economic data.
- We could do the same, but what other information could we get from this data?
- From a Public Health Perspective, there have been some relatively recent developments in Ontario.

COVID-19: The Case of Ontario

- However, Ontario adopted in late 2019 the Health Regions for healthcare and phased out the Local Health Integration Network (LHIN) approach.

COVID-19: The Case of Ontario

- However, Ontario adopted in late 2019 the Health Regions for healthcare and phased out the Local Health Integration Network (LHIN) approach.
- The change is relatively new, and therefore, geographical data can be used to analyze data within the different Health Regions.

COVID-19: The Case of Ontario

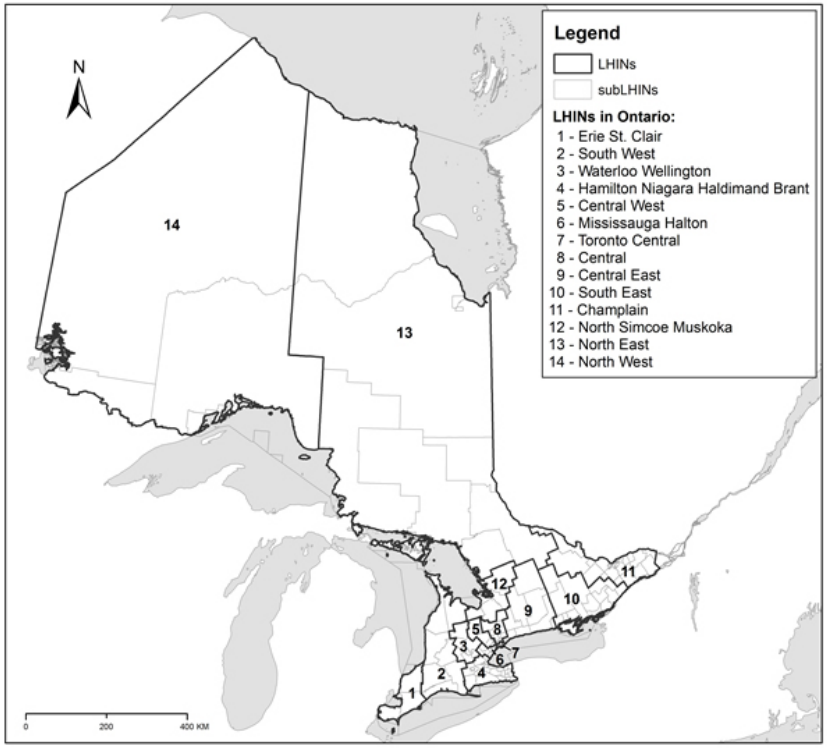

Figure 1: Ontario LHINs (Crighton et al. 2015)

COVID-19: The Case of Ontario

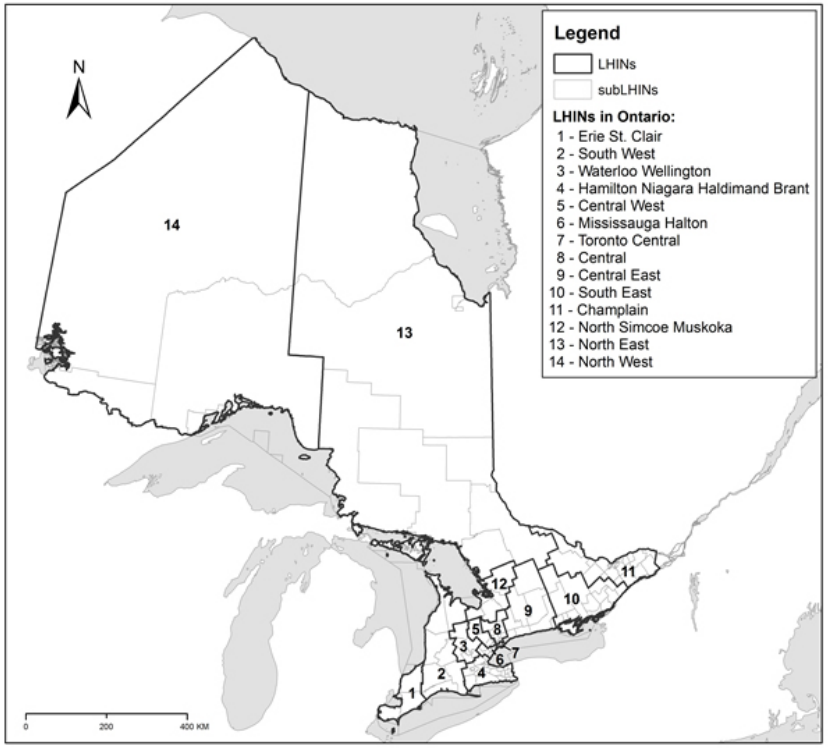

Figure 1: Ontario LHINs (Crighton et al. 2015)

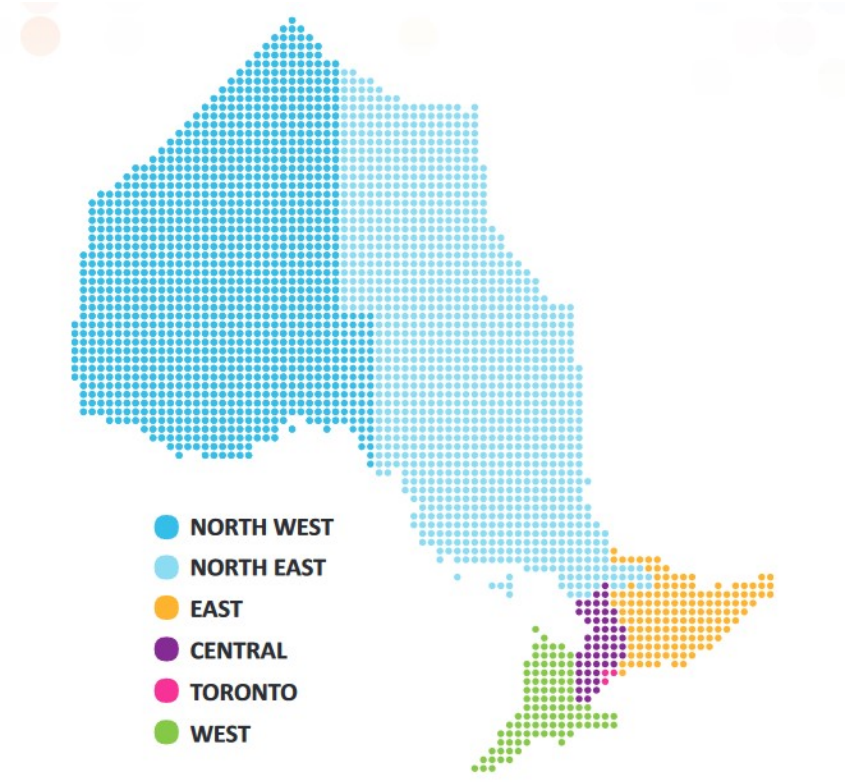

Figure 2: Ontario Health Regions (Ontario Business Health Plan 2022-2023)

COVID-19

- Therefore, we decided to integrate the different Health Regions in our analysis to determine the odds of vaccination.

$$\log\left(\frac{p(\text{vac})}{1-p(\text{vac})}\right) = \beta_0 + \beta_1(\text{Age group}) + \beta_2\ \text{Race} + \beta_3\ \text{Health Region} + \beta_4\ \text{Income} + \beta_5(\text{Health Region} \times \text{Race}) + \beta_6\ (\text{Income} \times \text{Race}) \quad (1)$$

Results

Table 2: Multivariable Regression Results

Characteristic	OR	95% CI	p-value
Age Group			
16_34	—	—	
35_54	0.90	0.67, 1.21	0.5
55_and_over	0.99	0.74, 1.32	>0.9
Income			
60000_and_above	—	—	
25000_59999	0.59	0.39, 0.89	0.011
under_25000	0.37	0.25, 0.56	<0.001
Race			
white_caucasian	—	—	
arab_middle_eastern	0.31	0.14, 0.69	0.004
black	0.32	0.17, 0.60	<0.001
east_asian_pacific_islander	1.15	0.50, 2.66	0.7
indigenous	0.44	0.19, 1.02	0.056
latin_american	0.28	0.11, 0.67	0.004
mixed	0.64	0.25, 1.65	0.4
other	0.22	0.12, 0.41	<0.001
south_asian	0.91	0.49, 1.69	0.8
Health Region			
Toronto	—	—	
Central	1.47	0.92, 2.35	0.11
East	1.42	0.90, 2.23	0.13
West	1.55	1.05, 2.30	0.029
Income * Race			
25000_59999 * arab_middle_eastern	1.79	0.67, 4.83	0.2
under_25000 * arab_middle_eastern	3.05	1.26, 7.39	0.013
25000_59999 * black	1.34	0.59, 3.05	0.5
under_25000 * black	3.19	1.45, 6.99	0.004
25000_59999 * east_asian_pacific_islander	0.42	0.17, 1.05	0.062
under_25000 * east_asian_pacific_islander	1.16	0.47, 2.86	0.8
25000_59999 * indigenous	1.36	0.48, 3.89	0.6
under_25000 * indigenous	1.45	0.55, 3.80	0.5
25000_59999 * latin_american	1.24	0.45, 3.43	0.7

Results

Table 3: Multivariable Regression Results

Characteristic	OR	95% CI	p-value
under_25000 * latin_american	2.80	1.04, 7.51	0.041
25000_59999 * mixed	0.85	0.32, 2.26	0.7
under_25000 * mixed	1.10	0.37, 3.27	0.9
25000_59999 * other	6.93	2.65, 18.1	<0.001
under_25000 * other	4.59	2.33, 9.05	<0.001
25000_59999 * south_asian	1.20	0.51, 2.85	0.7
under_25000 * south_asian	2.00	0.93, 4.30	0.077
Race * Health Region			
arab_middle_eastern * Central	0.66	0.26, 1.70	0.4
black * Central	0.44	0.19, 0.98	0.046
east_asian_pacific_islander * Central	0.98	0.38, 2.53	>0.9
indigenous * Central	0.63	0.22, 1.79	0.4
latin_american * Central	0.67	0.23, 1.96	0.5
mixed * Central	0.73	0.24, 2.22	0.6
other * Central	0.80	0.36, 1.78	0.6
south_asian * Central	0.54	0.25, 1.20	0.13
arab_middle_eastern * East	0.43	0.13, 1.45	0.2
black * East	0.83	0.34, 2.04	0.7
east_asian_pacific_islander * East	0.86	0.29, 2.56	0.8
indigenous * East	0.69	0.23, 2.08	0.5
latin_american * East	1.03	0.32, 3.34	>0.9
mixed * East	0.91	0.28, 3.03	0.9
other * East	1.05	0.39, 2.83	>0.9
south_asian * East	0.52	0.19, 1.45	0.2
arab_middle_eastern * West	1.00	0.37, 2.73	>0.9
black * West	0.76	0.32, 1.80	0.5
east_asian_pacific_islander * West	0.52	0.20, 1.34	0.2
indigenous * West	0.39	0.14, 1.09	0.073
latin_american * West	0.94	0.32, 2.72	>0.9
mixed * West	0.37	0.12, 1.16	0.089
other * West	0.41	0.18, 0.93	0.032
south_asian * West	0.41	0.18, 0.95	0.037

[1] OR = Odds Ratio, CI = Confidence Interval

The Case of Biomedical Data

Longitudinal Data

 Gerretsen, Philip et al. (Nov. 2021). "Individual determinants of COVID-19 vaccine hesitancy". In: *PLOS ONE* 16.11. Ed. by Leeberk Raja Inbaraj, e0258462. DOI: 10.1371/journal.pone.0258462. URL: https://doi.org/10.1371/journal.pone.0258462.

 Nafilyan, Vahe et al. (July 2021). "Sociodemographic inequality in COVID-19 vaccination coverage among elderly adults in England: a national linked data study". In: *BMJ Open* 11.7, e053402. DOI: 10.1136/bmjopen-2021-053402. URL: https://doi.org/10.1136/bmjopen-2021-053402.