Modeling chronic disease risk across equity factors using a population-based prediction model: the Chronic Disease Population Risk Tool (CDPoRT)

Kitty Chen, Kathy Kornas, Laura Rosella

ABSTRACT

Predicting chronic disease risk at a population level can help inform overall future chronic disease burden and opportunities for prevention. This study aimed to estimate the future burden of chronic disease in Ontario, Canada, using a population-level risk prediction algorithm and model interventions for equity-deserving groups who experience barriers to services and resources due to disadvantages and discrimination.

Methods

The validated Chronic Disease Population Risk Tool (CDPoRT) estimates the 10-year risk and incidence of major chronic diseases. CDPoRT was applied to data from the 2017/2018 Canadian Community Health Survey to predict baseline 10-year chronic disease estimates to 2027/2028 in the adult population of Ontario, Canada, and among equity-deserving groups. CDPoRT was used to model prevention scenarios of 2% and 5% risk reductions over 10 years targeting high-risk equity-deserving groups.

Results

Baseline chronic disease risk was highest among those with less than secondary school education (37.5%), severe food insecurity (19.5%), low income (21.2%) and extreme workplace stress (15.0%). CDPoRT predicted 1.42 million new chronic disease cases in Ontario from 2017/2018 to 2027/2028. Reducing chronic disease risk by 5% prevented 1500 cases among those with less than secondary school education, prevented 14,900 cases among those with low household income and prevented 2800 cases among food-insecure populations. Large reductions of 57 100 cases were found by applying a 5% risk reduction in individuals with quite a bit workplace stress.

Conclusion

Considerable reduction in chronic disease cases was predicted across equity-defined scenarios, suggesting the need for prevention strategies that consider upstream determinants affecting chronic disease risk.

BACKGROUND

The prevalence of chronic diseases is a significant global concern, particularly in Ontario, Canada, where non-communicable diseases, such as heart disease, diabetes, cancer or stroke, accounted for a large proportion of deaths in 2015. Predicting chronic disease incidence at a population level would be valuable for informing healthcare policymakers about future disease burden. Predictive models are useful for considering the multidimensionality of the risk distribution in the population. The use of algorithms for healthcare resource allocation plays a large role in population health management. Predictive models can also be used to inform target populations of their individual protective risk given demographic, behavioural or environmental conditions. The application of predictive models to inform interventions may amplify existing disparities if health equity considerations are not explicitly addressed in design and deployment. Predictive models that consider the upstream determinants could provide a transformative way of using population data and prediction models to support initiatives that tackle chronic disease burden and health inequities.

Reducing the chronic disease burden among equity-deserving groups, being those groups that experience barriers to opportunities and resources due to disadvantages and discrimination, is a fundamental principle of health equity. Equitable healthcare refers to access to healthcare services that are used based on an individual’s underlying status, regardless of socioeconomic status. Social determinants of health (SDH) that underlie inequities, such as access to affordable housing, safe neighbourhoods and quality education, are sometimes excluded from traditional interventions, which tend to focus on modifiable lifestyle factors, such as diet,
physical activity and the use of tobacco and alcohol.6 Targeting equity-deserving groups, within a proportionate universalism framework, can help address the health gradient through health equity and burden of disease approaches.8 By strategically balancing targeted and universal population health perspectives, resource allocation can be delivered with an intensity related to the level of social need within the population.9 Tools that can integrate SdoH and health equity principles into decision-making can promote and sustain equity in the health system.10

The Chronic Disease Population Risk Tool (CDPoRT) is a validated tool that forecasts population trajectories of chronic disease using routinely collected population data, including health and sociodemographic information.11 The objective of this study was to apply CDPoRT to the population living in Ontario, Canada, to estimate the future burden of chronic disease, identify risk among equity-deserving groups and demonstrate how the tool can be leveraged to address equity gaps in chronic disease prevention.

METHODS

Study population

This study focused on the population living in Ontario, Canada, aged 20 years and older in the 2017/2018 cycle of the Canadian Community Health Survey (CCHS) public-use microdata. The CCHS is a cross-sectional survey, representative of 98% of the Canadian population, that collects self-reported health information from Canadians aged 12 years and up. The survey covers the 10 provinces and the 3 territories, excluding persons living on reserves and settlements, full-time members of the Canadian Forces and institutionalised persons. Data collection was conducted using computer-assisted software or telephone interviewing. Detailed information about the CCHS survey methodology is available elsewhere.12,13

Using the 2017/2018 CCHS, we excluded those with existing chronic conditions such as congestive heart failure, chronic obstructive pulmonary disease, diabetes, lung cancer, myocardial infarction or stroke, aged 19 years and under, and living outside of Ontario, resulting in a total of 20282 respondents and 7 494 556 individuals when weighted.

Estimating 10-year chronic disease risk

CDPoRT is a validated model that can use risk factor information in population survey data to accurately predict the first incidence of six major chronic diseases (myocardial infarction, chronic obstructive pulmonary disease, congestive heart failure, diabetes, stroke and lung cancer) over a 10-year period. We applied CDPoRT to the CCHS to calculate the 10-year baseline risk and incidence of chronic disease in the Ontario adult population and among population groups.

Based on a Weibull accelerated failure time survival model, sex-specific prognosis models were developed using six cycles of the CCHS from 2001 to 2014 and were validated in two cohorts: internal validation using CCHS linked to health administrative data in Ontario, Canada (n = 11 874), and external validation using similar data in Manitoba, Canada (n = 13 244). CCHS respondents were followed from the CCHS interview date for 10 years to identify the first incidence of one of the six major chronic diseases using health administration data. The model estimates risk according to their survey responses that capture demographics, health behaviours and socioeconomic characteristics. Definitions for each CDPoRT input variable are described in online supplemental appendix A. The validated CDPoRT model had demonstrated sufficient overall predictive performance and a good discrimination (Harrell C-index: 0.780 (female) and 0.767 (male)).13 We have previously validated this approach for estimating population risk for a number of health outcomes.14,15 A detailed description of the model development and validation can be found in previous studies.11,16

Fruit and vegetable consumption was an optional CCHS module in the 2017/2018 cycle and was not collected in Ontario. To include the maximum number of respondents in the analysis, missing values from the CDPoRT predictor variables were imputed using multiple imputation based on Canadian provinces outside of Ontario and based on previous CCHS cycles.

We applied CDPoRT to our sample to estimate each individual’s 10-year chronic disease risk based on survey responses. We used centred adjusted coefficient values to account for population differences and adjust for baseline risk factors. The mean chronic disease risk estimate across the study population respondents was used to calculate the population-level risk of chronic disease at baseline and after each scenario as described below. To determine the expected number of new chronic disease cases overall and among subgroups, we multiplied the population size by the average risk. The algorithm formula is also described in online supplemental appendix B.

Equity factors

Population subgroups were defined according to equity factors available in the CCHS, which included age categories (20–34 years, 35–44 years, 45–54 years, 55–64 years, 65–74 years, 75–84 years, and 85+ years), sex (female or male), household income quintile (Q1, low; Q2, low-middle; Q3, middle; Q4, middle-high; and Q5, highest), ethnicity (white and non-white), education categories (less than secondary school and secondary school graduate), workplace stress (none, minimal, a bit of stress, quite a bit and extreme stress), household food security (food secure, moderately food insecure and severely food insecure), immigrant status with time since immigration (Canadian born; immigrant, 0–9 years; and immigrant, 10+ years) and household size (1, 2, 3, 4 and 5+ persons). Using Health Canada’s food security status model, ‘food secure’ refers to no indication of income-related food access, ‘moderately food insecure’ suggests compromises in food quality or quantity and ‘severely food insecure’ indicates reduced food intake and disrupted eating patterns.17

Interventions

Four intervention scenarios were developed to examine equity changes and their estimated impact on chronic disease risk in the following population groups: individuals with low household incomes (ie, Q1 and Q2), individuals with less than secondary school education, individuals who are moderately or severely food insecure and individuals who reported quite a bit and extreme workplace stress. These equity-deserving groups were selected because income, education, food security and workplace stress are important SdoH factors that can be targeted with community-wide interventions such as educational programmes, income supplements and workplace wellness programmes. We applied a relative risk reduction of 2% and 5% for each target group to the baseline chronic disease risk to quantify the expected impact of intervening on equity factors on chronic disease outcomes. Previous research has used similar risk reduction values of 2% to 10% to model population-level interventions.18–20 The relative risk reductions were chosen for a modest effect achievable in population-level interventions.19–21

Chen K, et al. J Epidemiol Community Health 2024;0:1–6. doi:10.1136/jech-2023-221080
**Statistical analysis**

Descriptive statistics were calculated for sociodemographic factors and health behaviours by sex. CDPoRT estimates each individual’s 10-year risk of developing a major chronic disease. Weighted analysis was performed using the sampling weights provided by Statistics Canada to generate the predicted number of chronic disease cases in population subgroups. Absolute risk reduction from baseline was calculated by subtracting the risk reduction scenario of 2% and 5% from the baseline risk. Similarly, the number of new chronic disease cases prevented was calculated by subtracting the number of new cases estimated in the scenario from the number of new cases estimated from baseline. All statistical analyses were performed using R, and the CDPoRT code can be found on GitHub.22

**Research ethics approval**

The study was conducted according to the principles of the Declaration of Helsinki and approved by the University of Toronto Health Sciences Ethics Board, protocol #31967.

**RESULTS**

Of the 113290 respondents in the 2017/2018 CCHS cycle, 20282 were eligible for the analysis after excluding participants living outside of the province of Ontario (n=79779), participants under the age of 20 years (n=3072) and participants who had an existing chronic disease (n=56425).

The weighted population of this sample represented 7 494 556 Ontarians (49.5% male and 50.5% female) (table 1). The largest age group was 20–34 years old (31.0%). The study sample was mostly of white ethnicity (74.0%), have an education greater than secondary school (92.8%), live in a two-person household (33.0%), are food secure (92.7%), are Canadian born (67.5%) and experience a bit of workplace stress (44.0%).

CDPoRT predicted 1420000 new chronic disease cases in Ontario from 2017/2018 to 2027/2028 (table 2). The expected burden was higher among males than females (755 thousand cases vs 660 thousand cases). The age group 55–64 years had the highest predicted chronic disease cases (339 thousand cases), and ages 85+ years had the highest chronic disease risk (62.5% risk). White respondents were expected to experience a greater chronic disease burden than visible minority (1.15 million cases vs 1.07 million cases). Those born in Canada had a higher expected burden of chronic disease, followed by long-term immigrants (with 10+ years in Canada) and immigrants with <10 years in Canada (944 thousand cases, 425 thousand cases and 47 thousand cases, respectively). Long-term immigrants had a greater predicted risk of chronic disease when compared with Canadian born (21.8% vs 18.7%). Two-person households had the greatest predicted burden (24.2%, 598 thousand cases), and one-person households had a greater chronic disease risk (25.2%, 277 thousand cases). Respondents with secondary school education or higher were expected to have a greater chronic disease burden but a lower chronic disease risk (17.5%, 1.21 million cases) than those with less than secondary school education (13.5%, 288 thousand cases). Severely food-insecure populations had a higher risk of chronic disease and a lower expected burden (19.5%, 36 thousand cases) compared with food-secure individuals (18.9%, 1.31 million cases). Lowest-income to low-middle-income brackets had the highest chronic disease risk (21.2% and 21.1%, respectively). Those who reported quite a bit and extreme workplace stress had the highest risk of chronic disease (13.5% and 15.0%, respectively).

Table 1 presents the 10-year chronic disease risk and new cases for each intervention scenario. A 2% and 5% in chronic disease risk for each population with the specified characteristic would reduce the overall risk for the group with less than secondary school education by 0.1% and 0.3%, respectively, preventing 330 and 1500 cases in the 10-year period. Similarly, low-income
DISCUSSION

This study demonstrates the application of a validated risk prediction tool (CDPoRT) to forecast the 10-year risk distribution of major chronic diseases across various equity factors. Modelling small population shifts in key equity factors demonstrated meaningful reductions in chronic disease burden within a 10-year period. Integrating SdOH factors in predictive models can help inform health system planning and resource allocation in a population. 21

We examined several SdOH factors and found that baseline chronic disease risk was highest among those with less than secondary school education, severe food insecurity, low income and extreme workplace stress. Consistent with existing literature, our study aligns with the idea that low education levels and poor economic conditions may synergistically increase the likelihood of multiple chronic diseases. 24 25 The mechanisms in which socioeconomic status influences chronic disease outcomes likely involve multiple interacting factors occurring throughout the life course. Understanding these dynamics can help facilitate short-term and long-term health and social care planning to accommodate life course changes.

The results of the study demonstrated that new immigrants have a lower risk of chronic disease relative to Canadian born and long-term immigrants. Literature suggests that foreign born, especially new arrivals, have better health than their Canadian-born counterparts. 26 This has been termed the ‘healthy immigrant effect’ in which immigrants are generally healthier than native born within 5–20 years of arrival due to the immigration policies in Canada that select young, healthy working adults. 26 This health advantage is lost with increasing duration of residency in Canada, showing a higher prevalence of chronic conditions among foreign-born populations. 27 Health within the immigrant population likely represents a combination of social, political, economic and cultural factors within the host society. Positioning the immigration experience as an SdOH can help better understand their health gradient for specific chronic conditions and guide culturally sensitive interventions. Given the limited data available on immigration in this study, future research could use migration data sources to explore differential chronic disease outcomes by sex and migration status.

We found that chronic disease risk varied across household size, with the highest risk observed among one-person and two-person households. The quality of life of families can offer the clinician a unique insight into issues such as family dynamics and how treatment decisions affect the patient’s close social group of partners and family. 28 29 Our results highlight that further attention to age distributions and household dynamics could be important for further study.

Our findings suggested that low food security was predictive of chronic disease risk, consistent with previous research, although the magnitude of risk across food security levels has been shown to vary across studies and study population. 30 Food security status was a stronger predictor of chronic illness even more than income. In a study in the USA, food security was found to be significantly associated with 10 chronic diseases studied, while income was statistically associated with only 3 of the 10 chronic diseases. 31 In a Canadian context, food security has been independently associated with an increased risk of diabetes after adjusting for socioeconomic and biological factors. 32 Examining the food security spectrum is important

and low-middle-income populations would experience a 0.2% and 0.6% decrease, which represents 3820 and 14 900 cases prevented. Food-insecure populations had a 0.2% and 0.5% decrease, respectively, preventing 820 and 2800 chronic disease cases. Applying a 2% and 5% reduction in individuals with quite a bit of workplace stress would decrease the risk by 1.9% (25 100 new cases prevented) and 4.3% (57 100 new cases prevented).

Of these equity factors, populations in these scenarios had 1–3 risk factors (online supplemental appendix C).

and low-middle-income populations would experience a 0.2% and 0.6% decrease, which represents 3820 and 14 900 cases prevented. Food-insecure populations had a 0.2% and 0.5% decrease, respectively, preventing 820 and 2800 chronic disease cases. Applying a 2% and 5% reduction in individuals with quite a bit of workplace stress would decrease the risk by 1.9% (25 100 new cases prevented) and 4.3% (57 100 new cases prevented).
for understanding the link between chronic illness and potential economic hardship, which can be contextualised into future chronic disease interventions.

Higher levels of workplace stress were found to predict a high number of incident chronic disease cases compared with low stress. Growing evidence suggests that psychosocial work stress is associated with a higher risk of various health problems. The job strain model explains that the combination of high job demands and low job control can compound work stress. The effort-reward imbalance model describes workplace stress as a combination of high efforts and low rewards in terms of salary, appreciation and job security. Work stress conceptualised according to two theoretical models has previously predicted incident chronic diseases in large-scale cohort studies.

There are few practical tools to support equity considerations in public health. Compared with existing chronic disease models, such as the Coronary Heart Disease Policy Model, CDPoRT is applied to population-survey data, which allows for consideration of SdoH, such as food security, which is not typically available in the health administrative data sources used in other chronic disease models. Our results demonstrated the importance of targeted interventions for marginalised populations. The social gradient in chronic disease has been well documented in research and suggests that direct targeting of socioeconomic barriers can moderate the efficiency of self-management support interventions. The approach presented can also be used in combination with more sophisticated microsimulation approaches; this model can further quantify the health and economic impact of various strategies. The modelling approach and findings of this study may be useful for informing important priority groups and targeted community-wide strategies for chronic disease prevention.

The results should be interpreted with consideration of the study’s limitations. Prediction of baseline chronic disease risk does not account for dynamic changes during follow-up such as mortality, shifts in population structure or migration. This issue can be ameliorated through the use of a simulation framework, as done for other conditions; however, this may limit the regular application of the approach in public health due to the need for additional data and capacity. Scenarios should also be interpreted with caution as stand-alone interventions as nuanced relationships across equity factors can impact chronic disease risk. Although the CCHS is a representative sample of the Canadian population, approximately 2% of the population, including institutionalised individuals living on Indigenous reserves, full-time residents of the Canadian Forces and people living in certain remote regions, were excluded from the sampling frame. Therefore, the results of this study may not be generalisable to these important equity-deserving groups. In particular, those living on reserves and certain remote regions present unique challenges to chronic disease, mainly those created by limited healthcare services and physical and emotional isolation. In addition, SdoH variables included in the study were not exhaustive as some factors (ie, urban vs rural areas and transportation types) were not available in the 2017/2018 CCHS. We chose to impute fruit and vegetable data to mitigate potential bias arising from excluding individuals with missing diet data. While acknowledging the potential for misclassification in self-reported data and the limited questions in the CCHS diet assessment, a scoping review on diet quality assessment tools identified the usefulness of CCHS diet variables in ascertaining dietary quality at the population level. These limitations are inherent in the use of Statistics Canada survey data and were beyond the control of this study.

**CONCLUSION**

Using a validated population risk tool, we demonstrated that a modest population shift in equity factors could result in a meaningful reduction of chronic disease burden. Our findings can provide important empirical data on upstream structural factors to support the equitable implementation of chronic disease prevention strategies.
Properly cited, appropriate credit is given, any changes made indicated, and the use and license their derivative works on different terms, provided the original work is permitted others to distribute, remix, adapt, build upon this work non-commercially, and/or omissions arising from translation and adaptation or otherwise.

Has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have repository. Canadian Community Health Survey Public Use Microdata File is publicly available in an open-access repository (https://ojes.ca/en).

This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use and license their derivative works on different terms, provided the original work is permitted others to distribute, remix, adapt, build upon this work non-commercially, and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID IDs Kitty Chen http://orcid.org/0009-0007-0624-1867 Laura C Rosella http://orcid.org/0000-0003-4867-869X

REFERENCES
29 Hjelmsga A, Deas I. The equity implications of an expanded health and wellbeing role for housing associations. Public Health Pract (Oxf) 2023;5:103555.
37 Hardman R, Begg S, Spelten E. What impact do chronic disease self-management support interventions have on health inequity gaps related to socioeconomic status: a systematic review. BMC Health Serv Res 2020;20:150.