Prospective cohort studies in Canada and their role in assessing environmental risk factors

John McLaughlin, PhD, FCAHS
Professor, Dalla Lana School of Public Health, University of Toronto
Executive Director, Canadian Partnership for Tomorrow Project, U of T

For OEH Seminar, Toronto
January 10, 2020

Disclosure of Conflicts of Interest

None

Declaration:
• Professor McLaughlin is co-lead of program with funding awarded through a national competition to establish the scientific base for the Canadian Partnership for Tomorrow Project
Outline

1. Environmental epidemiology examples and designs
2. Cohort study history & Canadian opportunities
3. Canadian Partnership for Tomorrow Project
   • CPTP purpose, design & current status
4. CPTP for studies of environmental factors
5. CPTP partnerships and unique opportunities
6. Future directions and discussion

Environmental & Genetic Factors in Lung Cancer

Toronto-based case-control study: 445 cases & 948 controls

Results: Benzene - OR=1.8 (95% CI = 1.3 - 2.7) Nitrogen dioxide - OR=1.6 (95% CI = 1.2 - 2.1)

Environmental Factors

<table>
<thead>
<tr>
<th>Exposure</th>
<th>OR (95% CI)</th>
<th>Pub</th>
</tr>
</thead>
<tbody>
<tr>
<td>eg, ETS</td>
<td>1.4 (0.9-2.2)</td>
<td>Kim et al, Int J Ca 2014</td>
</tr>
</tbody>
</table>

Gene OR (95% CI)

<table>
<thead>
<tr>
<th>Gene</th>
<th>OR (95% CI)</th>
<th>Pub</th>
</tr>
</thead>
<tbody>
<tr>
<td>15q25</td>
<td>1.2 (1.1-1.3)</td>
<td>Hung et al, Nature 2009</td>
</tr>
</tbody>
</table>

Consider:

• Strength of association & causality?
• Modifiable risk factor?
• Effective intervention?
• Implementation & by whom?
Lung Cancer Standardized Incidence Ratios

– Advanced GIS analysis with Bayesian smoothing & covariate adjustment

Geospatial Analysis Project
(Cancer Research Society grant to McLaughlin, Holowaty, Norwood & Harris)

Population data resources for exposures & outcomes enable environmental studies

Previously presented - Cross-Canada Case-Control Study - An early report on Glyphosate
(published odds ratio (OR), relative risk (RR) and confidence interval (CI))

McDuffie et al. (2001), Cancer Epidemiol Biomarkers Prev 10:1155

- Recall of pesticide exposures in early-1990s and before
- 4 types of cancer & many pesticides
- Prior hypothesis re: NHL and Phenoxy-herbicides (e.g., 2,4-D)
- Number who ever used glyphosate

Cases = 51/517 (10%) vs. Controls = 133/1506 (9%)
- Ever used glyphosate - OR = 1.2 (95% CI = 0.8-1.7) (adjusted*)
- Longest use of glyphosate - OR = 2.1 (95% CI = 1.2-3.7)*

* Logistic regression model adjusted for age, province, medical risk factors (e.g., measles, mumps, family history, etc.)
Non-Hodgkin Lymphoma

- Annual Incidence rate = 20 per 100,000 (age-standardized, both sexes combined / Source: Canadian Cancer Statistics, 2016)
- Rare disease

How large a cohort needed to obtain 500 incident cases?

- Cohort size (approximately)
  - No. cases / Rate (cases per 100,000 persons per year)
  - 2,500,000 p-yrs (i.e., follow 100,000 for 25 yrs)
- Whereas our case-control study collected information on approximately 2,000 participants

### Study Design Options – e.g., NHL

<table>
<thead>
<tr>
<th>Study</th>
<th>OR</th>
<th># Cases</th>
<th>Exposure Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case-Control - NHL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- McDuffie 2001</td>
<td>2.1 (1.2-3.7)</td>
<td>517 (51 exposed)</td>
<td>All used self-report</td>
</tr>
<tr>
<td>- De Roos 2003</td>
<td>2.1 (1.1-4.0)</td>
<td>872</td>
<td>Validation with no evidence of recall bias *</td>
</tr>
<tr>
<td>- Eriksson 2008</td>
<td>2.0 (1.1-3.7)</td>
<td>910</td>
<td></td>
</tr>
<tr>
<td>Cohort</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- De Roos 2005</td>
<td>1.1 (0.7-1.9)</td>
<td>92 NHL (71 exposed)</td>
<td>Self-report, with validation in sub-study</td>
</tr>
<tr>
<td></td>
<td>2.6 (0.7-9.4)</td>
<td>32 Multiple Myeloma</td>
<td></td>
</tr>
</tbody>
</table>

Radiation & Breast Cancer – A cohort example

Canadian Fluoroscopy Cohort Study

- Howe and McLaughlin. Radiation Research 1996
- 31,917 women
- treated for TB in Canadian institution between 1930-52
- 688 breast cancer deaths from 1950 to 1987
- Excess RR approximately constant from 5 to 39 yrs after exposure, with a possible decrease between 40 and 57 yrs
- Model estimates excess lifetime risk of breast cancer mortality after repeated, low-dose radiation exposures
- Relevant to risk assessment for routine mammographic screening

Cohort study history & Canadian opportunities
Prospective Cohort Studies – A selective history

• Frost (1935) introduced "cohort study" in 1935 to study TB in people born at different periods

• Framingham Heart Study – n = 5200 recruited in 1948-52, with 65 year follow-up and >3000 publications (Dawber et al)
  • Including odds ratio and early use of logit, leading to logistic regression (Cornfield 1951)

• Doll and Hill (1954) published 1st mortality follow-up from British Doctors Study – n=40,000 recruited in 1951 and followed for 50 years (Doll et al 2004)
  • Assessing environmental risk factors – e.g., Doll and Peto (1954) re: smoking

Prospective Cohort Studies in Canada – a brief historical selection

• Computerized record linkage (Newcombe, Science 1959)
• Many cohort eg’s across Canada – by population stratum, age (birth, older ages), occupation, geographic, ethno-cultural, disease-status, etc.
• Canadian Fluoroscopy Study – 120,000 adults fluoroscoped for TB therapy from 1930-52 with 60+ yr follow-up - linked to mortality and cancer incidence (Miller et al.)
• Canadian National Breast Screening Cohort Study (NBSS) – 90,000 women in RCT of breast cancer screening, ages 40–59 between 1980-85, 30 yr follow-up by record linkage (Miller et al.)

Anthony Miller
Professor Emeritus, DLSPH
and Order of Canada! (Dec. 2019)
Canadian Partnership for Tomorrow Project (CPTP) - Purpose, design & current status

As cohorts collect data on participants over time,
- some develop diseases
- some die and
- some remain disease free

The data from population laboratories are “invaluable for understanding gene-environment interactions in complex human disease.”

Cohorts contribute across the patient journey, delivering advances ranging from... – health promotion and disease prevention - to diagnosis and late outcomes

Purpose of Health Research - To assess...
- Disease causes
- Clinical disease occurrence
- Disease outcomes

Applications Advanced by Cohort Research
- Disease Prevention & Health Promotion
- Clinical research - detection & diagnosis
- Clinical outcomes & health services

Cohorts contribute to all domains of research

CPTP - A confederation of 6 regional cohorts
(covering 9 provinces, with final province in progress)

CPTP's new scientific base established at the University of Toronto in 2019 (with Ontario Institute for Cancer Research)
Entering a new era of scientific productivity, leadership, partnership, collaboration and productivity

CPTP’s leadership brings extensive experience in building large-scale research initiatives

- Build on & strengthen established structures
- Strong scientific and operational leadership
- Build on existing partnership
- Designed to strategically attract new partners

CPTP launched

2008

2014
Data access pilot successfully completed

2015
Access infrastructure in place

2015
Data made available to the scientific community

2017
Data linkages pilot project complete

2019
New scientific leadership established

Dr. John McLaughlin
Executive Director

Dr. Philip Awadalla
Scientific Director

Experienced in building and sustaining major research platforms used by researchers across Canada, including CPTP.

- Leadership Team:
  - Dr. Trevor Dummer, National Scientific Co-Director at UBC - with all Regional Directors

- Operational leads - across Canada

- National Strategic Advisory Council – with Funders and Hosts

- International Scientific Advisory Board
CPTP Purpose, Vision and Mission

**Purpose:** Enhance and accelerate research to prevent disease for a healthier Canada.

**Vision:** Improve population health through a better understanding of the causes of chronic disease and cancer.

**Mission:** To provide a national platform that supports high quality, innovative population health research in Canada and globally.

CPTP Data and Biosamples: Beginning in 2008

<table>
<thead>
<tr>
<th>Core questionnaire</th>
<th>DNA containing samples</th>
<th>Physical measures</th>
<th>Urine samples</th>
<th>Toenail clippings</th>
</tr>
</thead>
<tbody>
<tr>
<td>• &gt;300,000&lt;br&gt;• Demographics&lt;br&gt;• Lifestyle&lt;br&gt;• Risk factors&lt;br&gt;• Several others</td>
<td>• Venous blood collection (&gt;150,000)&lt;br&gt;• Blood spots (&gt;28,000)&lt;br&gt;• Saliva (&gt;8,000)</td>
<td>• Up to 90,000&lt;br&gt;• Height/weight&lt;br&gt;• Waist/hip circumference&lt;br&gt;• BMI&lt;br&gt;• Grip strength</td>
<td>• 101,000</td>
<td>• &gt;30,000</td>
</tr>
</tbody>
</table>
Enabling diverse health research based on linked data - e.g.
- causes of illness
- outcomes of disease
- response to clinical interventions
- community interventions
- detection of health related inequities
- health systems performance
- etc.

CPTP is linking personal, behavioural, environmental, health system and biological data to investigate cancer and chronic disease causes and determinants.

Linking to health data – A distinct Canadian advantage
An Active Partnership with Canadians: Recruited more than 320,000 participants (ages 30-74)

1 in every 100 Canadians participate

Alberta Tomorrow Project
29,800

Manitoba Tomorrow Project
41,374

Ontario Health Study
43,609

BC Generations Project
213,003

Atlantic PATH
36,003

Summary of CPTP participants and data available for research
T Dummer et al.
CMAJ, June 2018
CPTP follows both healthy and affected participants over time (i.e., prevalent & incident disease)

**Overall perception of health status in the CPTP cohort**

- Poor: 2%
- Fair: 9%
- Good: 33%
- Very Good: 40%
- Excellent: 16%

**Enriching the CPTP Datasets**

<table>
<thead>
<tr>
<th>Derived Data</th>
<th>Questionnaires</th>
<th>3rd Party Data Linkage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genotype (fall 2018)</td>
<td>CPTP core follow-up questionnaire</td>
<td>Canadian Urban Environmental Health Research Consortium (CANUE) (fall 2018)</td>
</tr>
<tr>
<td>Blood Analyte Data (fall 2018)</td>
<td>Occupational History Questionnaire</td>
<td>Cardiac, Vascular and Cognitive Dysfunction Alliance (CVCD Alliance) (spring 2019)</td>
</tr>
</tbody>
</table>
Access to CPTP Data by the Research Community

CPTP Data Portal – Online and Active

Website: www.partnershipfortomorrow.ca
Portal: portal.partnershipfortomorrow.ca
Facebook: www.facebook.com/cptproject
Twitter: @cptproject

CPTP for OEH research on environmental determinants of disease
Supporting studies of concern to local populations

Arsenic and Health (Atlantic)

ORIGINAL ARTICLE
Relationship between drinking water and toenail arsenic concentrations among a cohort of Nova Scotians
Zhike H. Yu, Trevor J.B. Dumme, Ameer Adams, John D. Munirambah and Louise Parker
Journal of Exposure Science and Environmental Epidemiology advance online publication, 31 December 2013.

Arsenic concentrations in toenails exposed to different levels of well water arsenic among Nova Scotians across tertiles of body fat mass index

Advancing evidence and knowledge translation relevant to local populations

Arsenic and Cancer (Atlantic)

Environment International
Volume 96, May 2014, Pages 118-123

What is the role of obesity in the aetiology of arsenic-related disease?
Zhike H. Yu*, Shunfeng Zhang, John D. Munirambah, Louise Parker, Trevor J.B. Dumme

Highlights
- An obesity-related diet score was developed using reduced rank regression
- Drinking water arsenic levels were comparable across quantiles of the diet score
- Individuals with obesity-related diets had lower toenail arsenic concentrations

Understanding the translation of scientific knowledge about arsenic risk exposure among private well water users in Nova Scotia

Highlights
- Scientific and public knowledge of arsenic risk in private wells is misaligned.
- Well users surveyed were confident in water quality without regular testing.
- Local and social networks are key to arsenic risk knowledge circulation.
- Stakeholder interaction highlights effective channels for risk communication.
CPTP Partnership:

Canadian Urban Environmental Health Research Consortium

Funded as CIHR Signature Initiative ($4M over 5 years)

Environmental data linked to CPTP

CANUE is led by:

• Dr. Jeffrey Brook (PI)
• With CPTP leaders and interdisciplinary, pan-Canadian team of research leaders and collaborators

- Every location in Canada described by complex set of environmental factors
- CANUE is building capacity to study how these multiple **environmental factors are linked to a wide range of health outcomes**
- Enable effective, evidence-based strategies for planning healthy cities and towns, today and in the future.

**CANUE Data Platform – Data Themes**

ANALYSIS-READY EXPOSURE DATA = easy for researchers to access and use

Source (with thanks): J. Brooks, D. Doiron, E. Seaton
Canadian Urban Environmental Health Research Consortium

CPTP and CANUE

Source (with thanks): J. Brooks, D. Doiron, E. Seaton

Canadian Urban Environmental Health Research Consortium

CANUE Data Linked to National CPTP Dataset

https://portal.partnershipfortomorrow.ca/

ENVIRONMENTAL EXPOSURE DATA

The CHRI-funded Canadian Urban Environmental Health Research Consortium (CANUE) collates and generates standardized area-level environmental data on air and noise pollution, land use, green/natural spaces, climate change/extreme weather, and socioeconomic conditions and links this data to existing Canadian cohort studies and administrative health records. An initial body of CANUE exposure datasets have been merged with the national harmonized CPTP dataset and are now available to researchers. These datasets include:

- Canadian Active Living Environments Database (Can-ALE)
- Material and Social Deprivation Index
- Normalized Difference Vegetation Index (NDVI) (i.e., “greenness” metrics)
- Annual average nitrogen dioxide (NO2) exposure
- Annual average ozone (O3) exposure
- Annual average fine particulate matter (PM2.5) exposure
- Annual average sulfur dioxide (SO2) exposure
- Weather and Climate metrics

Source (with thanks): J. Brooks, D. Doiron, E. Seaton

Canadian Urban Environmental Health Research Consortium
### CANUE Data Example – Air quality

**Sulphur Dioxide (SO₂)**
- Modeled by Environment Canada using OMI satellite data
- Annual average concentrations
- 30 km resolution

Source (with thanks): J. Brooks, D. Doiron, E. Seaton

### CANUE Data Example – Greenness

**Normalized Difference Vegetation Index (NDVI)**
- Processed from NASA satellite data using Google Earth Engine
- Annual average, annual maximum and growing season average at each postal code, and within set distance buffers
- 30 m resolution

Source (with thanks): J. Brooks, D. Doiron, E. Seaton
### CPTP Quintile distribution

(Provisional data; n = ~250,000)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
<th>Q5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Material deprivation</td>
<td>69,102</td>
<td>56,578</td>
<td>48,992</td>
<td>40,712</td>
<td>28,094</td>
</tr>
<tr>
<td>Social deprivation</td>
<td>50,494</td>
<td>49,717</td>
<td>47,873</td>
<td>48,163</td>
<td>47,231</td>
</tr>
<tr>
<td>Walkability</td>
<td>66,984</td>
<td>84,445</td>
<td>63,855</td>
<td>21,258</td>
<td>14,317</td>
</tr>
<tr>
<td>NO₂</td>
<td>54,685</td>
<td>52,760</td>
<td>46,326</td>
<td>43,022</td>
<td>47,187</td>
</tr>
<tr>
<td>SO₂</td>
<td>52,390</td>
<td>41,642</td>
<td>46,457</td>
<td>42,450</td>
<td>43,286</td>
</tr>
<tr>
<td>O₃</td>
<td>33,942</td>
<td>28,905</td>
<td>54,672</td>
<td>63,227</td>
<td>70,993</td>
</tr>
<tr>
<td>Greenness (year mean)</td>
<td>41,631</td>
<td>47,534</td>
<td>57,504</td>
<td>56,267</td>
<td>48,810</td>
</tr>
</tbody>
</table>

Source (with thanks): J. Brooks, D. Doiron, E. Seaton

### Residential History Construction: Social Data Linkage Environment (SDLE)

Individual identifiers and unique id

SDLE

Residential history created and exposure assigned by postal code/year

CANUE

Exposure and unique id

Source (with thanks): J. Brooks, D. Doiron, E. Seaton

Canadian Urban Environmental Health Research Consortium
For More Information about CPTP and CANUE:

**CPTP National Webinar**
- by Jeff Brook on February 13th at 12:00
- https://zoom.us/webinar/register/WN_FAo2lfKeSVynLe3noWe6ig

**CPTP Website and Data Portal**
- https://portal.partnershipfortomorrow.ca

**CANUE Website and Data Portal**
- https://www.canuedata.ca
Key partnership: Canadian Alliance for Healthy Hearts and Minds (CAHHM)

- Collected detailed information on vascular disease, cardiac disease and cognitive function using MRI scans
- Data collected from 10,000 Canadians through existing cohorts, including 1500 First Nations people living in Canada
- Unique data to evaluate the impact of diverse environmental determinants on cardiovascular health

The Alliance is co-led by Drs. Sonia Anand, Matthias Friedrich and the late Jack Tu.

Key partnership: Pan-Canadian Real-world Health Data Network (PRHDN)

- PRHDN is a distributed data network that allows researchers and policy/decision makers across Canada to use linked and linkable administrative (real-world) data holdings and expertise in multi-province studies and initiatives without requiring that data leave provincial boundaries.
Cohort studies with comprehensive data and biospecimens:

- **CPTP** (Canadian Partnership for Tomorrow Project): > 320,000
- **CLSA** (Canadian Longitudinal Study of Aging): ~ 50,000
- **MIREC** (Maternal Infant Research on Environmental Chemicals): 2001
- **CHILD** (Canadian Healthy Infant Longitudinal Development): 3455

Examples of alternative platforms:
- Canadian Health Measure Survey (5 cross-sectional surveys): ~ 29,000
- Canadian Community Health Survey (X-S survey every 2 yrs): ~ 65,000

*Large size essential to assess complexity & rare exposures or outcomes*

CPTP singularly enables Canada to contribute to the network of internationally recognized large-scale initiatives.
Future Directions and Discussion

CPTP Innovations for Occupational Epidemiology

Evaluation of AI algorithm to Classify Occupational History (2019-20)

- PI - Dr. Ellen Sweeney, Director of Strategic Research Initiatives, Atlantic PATH
- Interdisciplinary team conducting project to enhance CPTP’s data platform
- Aim – to improve data quality by harmonizing data of Atlantic Path’s and Alberta’s Tomorrow Project’s 111,000 occupational history questionnaires
- Automatic Semantic Occupational Coding algorithm (Bao, Baker, Adisesh) to analyze and harmonize open-text data, occupational history information across two cohorts, and evaluated for utility in other regions.
The **Exposome** - Three Domains of Environmental Exposures (with examples)

![Diagram showing three domains: General external, Internal, Specific external](image)


---

**Exposomics (Internal) & Detection of Biological Response**

- More traditional measures of **EXTERNAL** environment
- New possibilities to assess **INTERNAL** environment & biological response

- High throughput "genomics" tools to characterize:
  - Genome – Transcriptome – Proteome – Metabolome ...
  - Characterize gene-environment interactions
  - Detect biological effects and susceptibilities
CPTP enables studies of gene-environment interactions

Recombination affects accumulation of damaging and disease-associated mutations in human populations

CPTP enables studies of gene-environment interactions, of direct relevance and interest to communities

Gene-by-environment interactions in urban populations modulate risk phenotypes

Genetic study of Quebec residents finds air pollution trumps ancestry

"That's really what precision health is about," Dr. Awadalla said. "You want to capture these things before people are in the doctor's office and having to be treated."
Integrating Internal and External Exposomes
- Epigenomic Assessment of Gene-Environment Patterns in Quebec

Awadalla et al (Nature Comms - 2017) - Differentially expressed genes (DEGs) are associated with local ambient air pollution.

Opportunity for advancing environment and health research in Canada

CANUE – Canada’s external exposome data for individual-level assessments, with CPTP

New exposomics initiatives
- Complete exposome assessments
- Integration with big health data
- Enhance expertise & advanced analytical methods
Discussion

CPTP makes “Big Data” and Cohort approaches possible to detect and manage environmental effects on health across Canada.

_What could you do with ongoing data on more than 300,000 participants across Canada?_

- Analysis of existing data, and collaborations to expand on and enrich data – e.g., advanced exposure assessment, exposomics...
- A platform for your grants, to support analytics and grad students – e.g., CIHR calls for proposals.
- A base and starting point for major initiatives – e.g., as with “Alliance” (CAHHM), CANUE, CFI, New Frontiers...
- Collaborations to resolve gaps and limitations in Canada’s population data capacity – e.g., sustained cohort follow-up complements Canada’s existing platforms for monitoring, characterizing and managing environmental health effects.

Thanks to all CPTP participants and supporters across the 6 regional cohorts who generously donate their time, information and biological samples. **CPTP is a success because of participants’ ongoing commitment.**
Thanks to all sponsors and partners – Current & Future, and
Join us at portal.partnershipfortomorrow.ca