ACCESS OFFICE HOURS

CanPath

Canadian Partnership for Tomorrow's Health

Discover the requirements for successful student-led projects

Welcome to CanPath Access Office Hours!

- **CanPath Access Process & FAQs** presented by <u>Nouar Elkhair</u>, Access Officer

- **Overview of CanPath Biosamples** presented by <u>Treena McDonald</u>, National Biosamples Coordinator

- **CanPath Data & Navigating the Portal** presented by <u>Sheraz Cheema</u>, Data Manager

CanPath

Access Process Overview

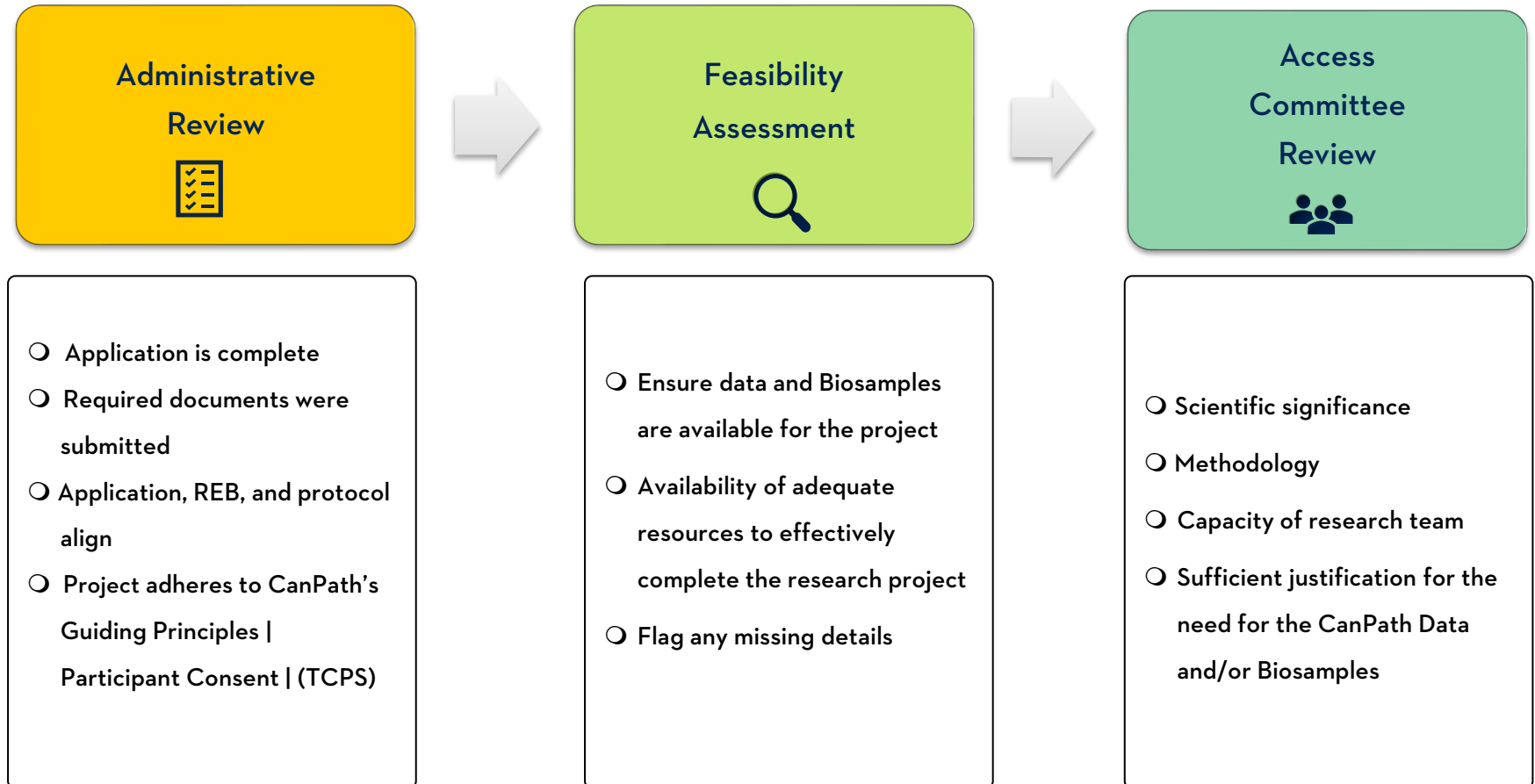Nouar Elkhair – CanPath Access Officer

What documents are required to access CanPath Data and Biosamples?

- Access Application Form (register on the [access portal](#))

- Research protocol (having received ethics approval)

- Approval letter by a Research Ethics Board

- 2-page CV of the principal applicant

- Contact information of your legal institutional representative

- Proof of scientific review or peer review, *if available*

- Evidence of funding, *if applicable*

CanPath | 3-Step Access Process

Administrative Review

- Application is complete
- Required documents were submitted
- Application, REB, and protocol align
- Project adheres to CanPath's Guiding Principles | Participant Consent | (TCPS)

Feasibility Assessment

- Ensure data and Biosamples are available for the project
- Availability of adequate resources to effectively complete the research project
- Flag any missing details

Access Committee Review

- Scientific significance
- Methodology
- Capacity of research team
- Sufficient justification for the need for the CanPath Data and/or Biosamples

CanPath

How quickly can I access Data or Biosamples?

Expedited vs. Full Access Review

	1 Administrative Assessment	**2** Feasibility Assessment	**3** Access Committee Review
EXPEDITED REVIEW	CanPath Access Office reviews application for completeness	**One Scientific Director** reviews the application	**One member** of the Access Committee reviews applications on a **rolling basis**
FULL ACCESS COMMITTEE REVIEW	CanPath Access Office reviews application for completeness	**All Scientific Directors** review the application	**All members** of the Access Committee review applications during the **bi-monthly teleconference**

Expedited reviews: can be completed in 4–6 weeks

Full Access review: timelines vary

Which applications are expedited?

Expedited Review

- **Harmonized dataset** requests (e.g., baseline, follow-up, CANUE environmental exposure datasets, COVID-19 datasets)
  - Note: Harmonized datasets include data from 5 regions – BC Generations Project (BC), Alberta's Tomorrow Project (AB), CARTaGENE (QC), Ontario Health Study (ON), and Atlantic PATH (NB, NS, PEI, & NL)
  - COVID Serology data also includes data from the Manitoba Tomorrow Project

Full Access Review

- **Biosample** requests

- **Data Linkage** linking CanPath data to administrative health data (e.g., ICES, AbSPOR, PopDataBC)

- **Regional data** requests other than the harmonized datasets

CanPath

What does the Access Review Process involve?

- The CanPath review process is iterative ➔ … Questions … Modifications … Clarifications … ➔ Timely response is key!

- Following Access Committee Approval:

  Sign Access Agreement ➔ **Invoicing** ➔ **Data/Biosamples Release**

- Annual Report, REB Updates, and Member Updates are <u>ongoing</u>

- Once a project is <u>complete:</u>

  Return Derived Data ➔ **Return/Destroy Data/Biosamples**

CanPath

How much will it cost to access data?

- Cost recovery fees for data-only applications for academic researchers are available on the CanPath website.

- For biosamples and industry-based research requests, contact access@canpath.ca

	Cost for the first dataset	Cost for each subsequent dataset
Researcher	$3,000	$1,000
Early-Career Researcher† *Canadian Institution*	$750	$250
Early-Career Researcher† *Non-Canadian Institution*	$1,800	$600
Graduate Student or Postdoctoral Fellow ‡	$250	$190

CanPath

FAQ

- **Before applying, can I get estimates of the sample sizes for my project?**
  - Yes, CanPath provides preliminary aggregate data for Data and Biosamples meeting your selection criteria.

- **Can I get a Cost Estimate | Letter of Support before applying?**
  - Yes, complete the form (available on the CanPath Portal)
  - Send it to access@canpath.ca (~10 business days)

- **Any tips on filling in the application form/required documents?**
  - Review Application Tips (available on the CanPath portal)

Tips for successful student led applications

- Your research supervisor must be listed as the Principal applicant, and you must be part of the research team members

- Names/affiliations on application and REB must match

- The protocol and application must include clear selection criteria or exclusion and inclusion criteria (avoid "All ")

- Include sufficient justification for the data and biosamples requested (the request must align with research objectives)

- Choosing a dataset on the application means you will receive the full dataset for the cohorts providing that data.

- You can choose specific variables; however, this is not required.

Biosamples Overview

Treena McDonald – CanPath Biosamples Coordinator

Biosamples | What's available?

- Blood derivatives: serum, plasma, DNA, red blood cells, whole blood in DMSO (>145K participants donated one or more blood derivative)

- Urine (>90K participants)

- Toenails (Atlantic PATH only, >30K participants)

- Saliva: DNA

Access to Biosamples

- *Biosample Access Guidelines* and *Biosample Feasibility Review* are available from the CanPath portal.

- Asks of >500 uL for non-DNA samples and >2 ug for DNA samples are considered a *large ask* and will require extraordinary scientific justification to be considered.

- Full Access Committee review is required for submissions seeking biosamples.

- Approved biosamples may be conditionally released in a staggered format.

Data Overview

Sheraz Cheema – CanPath Data Manager

Datasets | What's available?

CORE

- 2008 – present
- 14 harmonized datasets
- Nearly 4k variables
- Up to 330k participants

GENOTYPE

- 50k+ participants
- 820k SNPs (Affymetrix UK Biobank Axiom® 2.0 gene chip)

ENVIRONMENTAL EXPOSURE

- Merged to core datasets (using 6-digit postal codes)
- 27 unique datasets
- Up to 330k participants

CAHHM

- 2013 – 2019
- MRI; Cognitive; administrative (ER)
- 4.7k variables
- 7.7k participants

CanPath

CanPath Datasets | Core

- Baseline Health and Risk Factor Questionnaire

- Baseline Health and Risk Factor Questionnaire – Additional Diseases

- Baseline Mental Health

- Baseline Physical Measures

- Follow-Up 1 Health and Risk Factor Questionnaire

- COVID-19 (2020) Questionnaire

- Baseline COVID-19 Antibody Questionnaire & Results

- Follow-up 1 COVID-19 Antibody Questionnaire & Results

CanPath Student Dataset

- **Purpose:** create individual-level dataset preserving statistical patterns (i.e. relationships between variables)

- **Synthetic dataset:** a random sample of CanPath data where participant information has been replaced and rearranged

  - Created using the R Package 'synthpop'

  - This package was explicitly designed to generate synthetic versions of longitudinal survey data

Advantages of the CanPath Student Dataset

- ✓ Large sample size

- ✓ Real-world population-level Canadian data
- ✓ Variety of areas of information allowing for a wide range of research topics

- ✓ No cost to faculty
- ✓ Potential for students to apply for real CanPath data to publish their findings

CanPath

Recent Applications of the Student Dataset

"This data challenge, developed using CanPath student data, is an amazing training opportunity to effectively introduce advanced causal inference techniques and how they can be applied to complex health studies using real-world data,"

Prof. Kuan Liu, *AI4PH* mentor and Assistant Professor, University of Toronto

Example studies from *AI4PH*:
- Association between depression and cardiovascular disease (CVD)
- Relationship between walkable neighbourhoods and physical activity
- Risk of developing Type II diabetes compared in White vs. Black adults

CanPath

CanPath.ca

CanPath

Canadian Partnership
for Tomorrow's Health