MOHCCN GOLD COHORT STANDARDS POLICY V1.4

Table of Contents

1. Introduction

The Marathon of Hope Cancer Centres Network (MOHCCN) aims to create a "gold-standard" cohort of clinical cancer specimens with a well-annotated, uniformly generated, and consistently quality-controlled dataset (clinical and genomic) from 15,000 (15k) cases collected from across Canada over 5 years. Not only does the MOHCCN aim to build a pan-Canadian Cancer Network and to produce immediate clinical impact by identifying actionable targets through molecular profiling, but it also proposes to generate in-depth molecular profiling data from cancer patient cohorts to address important scientific questions. This richly clinically annotated molecular dataset, starting with standardized clinical information, treatment response data, and whole-genome and transcriptome profiles (WGTS), will serve as an invaluable resource for cancer biology discovery (Figure 1).

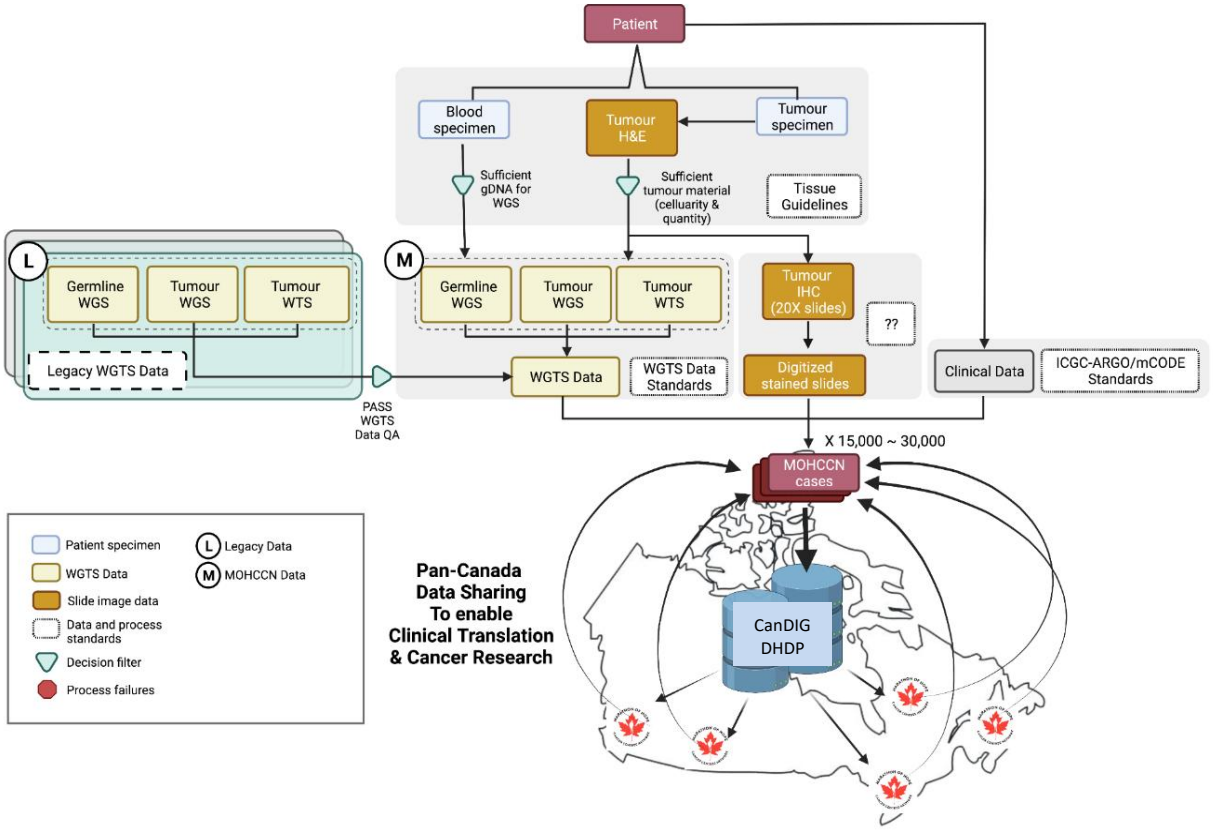

Figure 1. MOHCCN data roadmap to 15k gold standard cases.

2. MOHCCN Case Definition and Data Requirements for the 15k Gold Cohort

A MOHCCN case is defined as a unique patient. The table below provides the patient specimen requirements (tumour samples and non-tumour samples, like blood or non-neoplastic tissue) to meet the minimum (Gold Standard) data standards. Guidelines to support collections of samples will be developed through the following MOHCCN Working Groups:

Complete Case Components	MOHCCN Working Group
Clinical data	Clinical Data Standards Subcommittee
Specimen/non-neoplastic materials (whole blood, buffy-coat, PBMC, extracted genomic DNA)	Biospecimens Working Group
Tumor materials (FFPE, FF, viably frozen, extracted genomic DNA and RNA). Digitized image of the pathology reviewed H&E	Biospecimens Working Group
WGTS data	Technology Working Group

Data Requirements for the 15k Gold Standard Case

	Required* (MOHCCN standards available)	Preferred** (MOHCCN standards needed)
Clinical data		
Cross-sectional imaging (CT/MRI/PET)		✓
MOHCCN Clinical Fields	✓	✓
Health technology assessment		✓
Disease-specific		✓
Tumor Tissue		
Pathology-reviewed cellularity (Digital H&E slide or clinical flow-cytometry)	✓	✓
Whole Genome (FFPE & frozen)	✓	✓
Whole Transcriptome (FFPE & frozen)	✓	✓
Multiplexed IHC (FFPE & frozen) [*Solid tumors only*]		✓
ATAC-seq (frozen)		✓
ChIP-seq (frozen)		✓
Bisulphite Whole Genome (frozen)		✓
Proteomic Assay (FFPE & frozen)		✓
Flow cytometry or CyTOF (frozen)		✓
Single cell RNAseq (viable/frozen)		✓
Spatial transcriptomics/proteomics profiling (FFPE/frozen)		✓

Blood	
PBMC Flow Cytometry (viable)	✓
Plasma cell-free whole genome	✓
Plasma cytokines/metabolites	✓
Others (stool, urine, pleural fluid, etc.)	
Microbiome analysis (stool)	✓

*Required: MOHCCN standards are available to support data collection.
**Preferred: MOHCCN standards to support data collection are in development.

Case Collection: Prospective and/or Retrospective

Patient specimens for WGTS and molecular profiling can be obtained via one of or a mixture of two methods: Prospective and/or Retrospective.

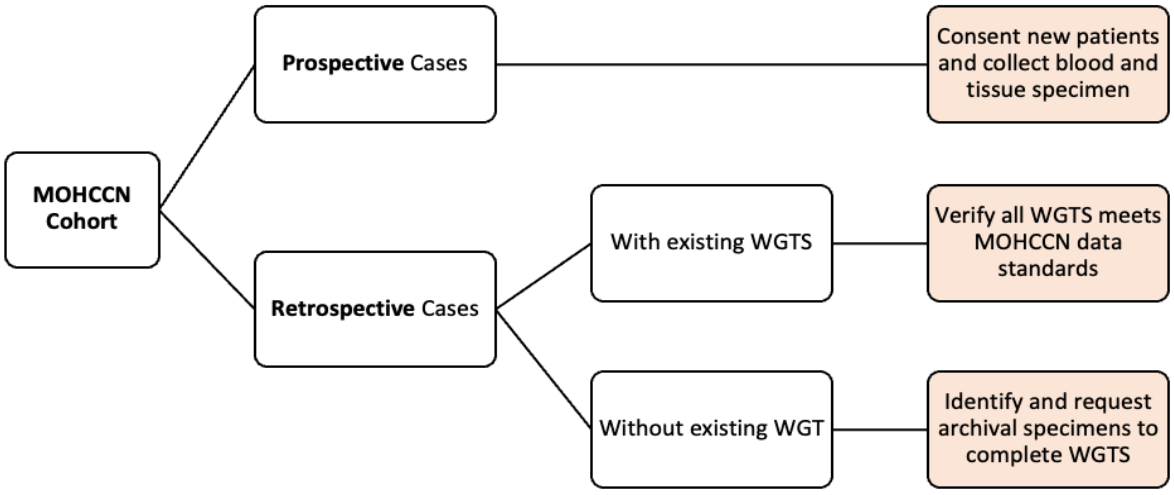

Figure 2. Prospective/retrospective cases.

Prospective specimen procurement is the set of procedures to collect and store new (since the time of study approval) patient specimens for WGTS. Solid tumor tissues can be collected via surgical resection or biopsies at baseline. If possible, prospective specimen handling protocols should be followed to ensure quality of specimens for WGTS success. By adhering to best-practices guidelines for specimen collection and processing for WGTS, prospective cases have the highest probability of success in meeting MOHCCN case standards. We recommend focusing the majority of efforts towards acquiring and enrolling patients/cases prospectively for the MOHCCN 15k gold cohort in Years 3 to 5.

Retrospective specimen procurement is the set of procedures to identify the most optimal available/archival (previously collected) patient specimen for WGTS. This procedure may also involve identifying and obtaining available partial germline WGS and or tumor WGTS data. As the specimens may originate from multiple biobanks and storage facilities, this is a labor-intensive and nuanced process. Cohorts

relying on retrospective specimen procurement encounter many unique challenges prohibiting successful WGTS data generation. To complete the MOHCCN 15k gold cohort, we recommend minimizing the number of retrospectively procured specimens or cases where possible.

3. MOHCCN Case Data Requirements and Quality Tiers

To find a more linear and straightforward operational process in the generation of the molecular data for MOHCCN cases, a Tier B of data has been defined as part of the standards to be approved as a "Gold Cohort" case given their scientific value.

While all projects will do their best efforts and aim for the coverages outlined in the WGTS guideline, data that does not meet the outlined thresholds will fall under Tier B and still be considered a MOHCCN case. Any data that falls below Tier B will not be considered a MOHCCN "Gold Cohort" case. Please refer to the chart below for threshold details.

The BAM files delivered should be considered "lossless". That is, the BAM should contain all the reads generated by the instrument. This should include mapped and unmapped reads and reads flagged as PCR or optical duplicates. In the absence of a lossless bam file, raw fastq files should be provided.

A. Molecular Requirements

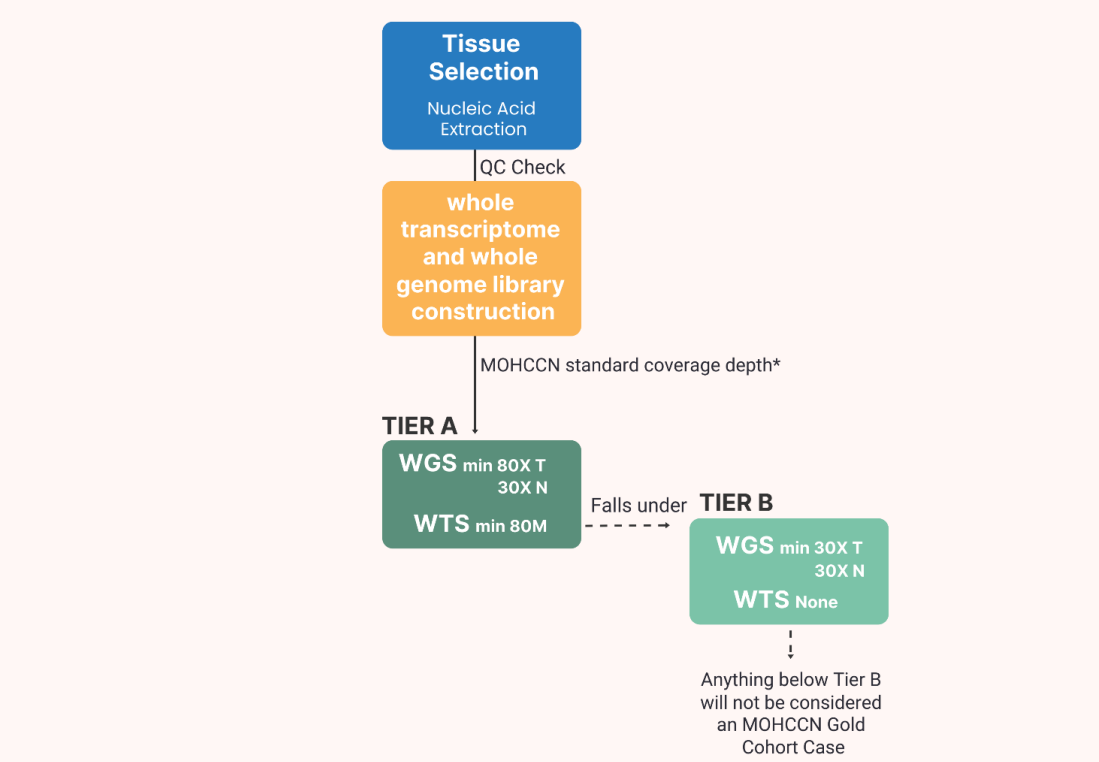

Figure 3. WGTS General Workflow for generation of a MOHCCN Gold Cohort Case.
*It is recommended to overshoot the MOHCCN standards coverage to ensure a higher probability of achieving Tier A. This policy will be revised to include more details on best practices to achieve the gold standards once that information is collected and harmonized across centres.

Tier A

Case Technology	Standards
Whole Genome Sequencing (WGS)	Minimum 30X normal and 80X tumour coverage (WGTS Guideline)
Whole Transcriptome Sequencing (WTS)	Minimum 80M pairs (WGTS Guideline)

Tier B

Case Technology	Standards
Whole Genome Sequencing (WGS)	Minimum 30X normal and 30X tumour coverage
Whole Transcriptome Sequencing (WTS)	No RNA

B. Sequencing Metadata Requirements

Upon successful completion of sequencing of a case, the cohort or the sequencing centre shall submit to CanDIG the information outlined in the MOHCCN Sequencing Metadata Policy. The MOHCCN will follow the standards used by the International Nucleotide Sequence Database Collaboration (INSDC).

C. Clinical Requirements

Upon successful completion of sequencing of a case, the cohort shall submit to CanDIG the complete clinical data within six months. The complete clinical data shall consist of the mandatory and conditional/optional fields described in the most current MOHCCN Clinical Data Model (CDM). The complete clinical data may be submitted in a staggered fashion if desired, provided that submissions are completed within six months. If a good faith effort has been made to obtain a mandatory clinical field and has failed, the field may be submitted as “Not Available.”

Follow-Up

Subsequent to submission of "complete" clinical data, clinical follow-up information, including but not limited to a patient's vital and disease status, shall be submitted on an annual basis for a minimum of five years, subject to available funding.

D. H&E Requirements

The cohort shall confirm the existence of H&E slides and document the parameters of H&E imaging, including the magnification obtained and model number of the instrument. This information shall be submitted with the cohort metadata.

Magnification

It is preferred to have H&E images scanned at 40x magnification, for image analysis work. For images collected prior to December 2025, 20x magnification will be accepted.

If the institution retains the physical slides for future availability and image storage space is a concern, scanning at 20x magnification is sufficient. Images at 40x may be requested and generated in the future.

Digital File Format

An open format for digital files is requested, so the files can be used by most slide image processing/viewing software academically. Examples of open format file types are:

- .svs
- .ndpi
- .tiff.

Other Guidance Related to H&E

- The image must be free of any potential patient identifiers, including any macro data in the file such as barcodes and label snapshot.
- Image quality should be adequate for use in research publications (e.g. journal articles). Avoid blurry or non-interpretable images.
- If H&E slides include tumour markings in ink, it is recommended to scan both a marked-up and unmarked slide. However, if only a marked-up slide is available, this is acceptable.

Document Revision History

Developed by	Reviewed by	Endorsed by	Effective Date	Policy Version
BWG	Steering Committee	Network Council	February 27, 2026	V1.4
TWG & DPSC	Steering Committee	Network Council	August 1, 2025	V1.3
DPSC	Steering Committee	Network Council	October 3, 2024	V1.2
TWG	Network Council and Exec Ad hoc	Network Council	February 1, 2024	V1.1
TWG	Steering Committee	Network Council	October 6, 2022	V1

Authors

Name	Institution	Title
Marco Marra (TWG Co-Chair)	BC Cancer	Distinguished Scientist / Professor, Medical Genetics and Michael Smith Labs at UBC
Trevor Pugh (TWG Co-Chair, BWG Co-Chair)	U of T	Director / Senior Investigator
Ian Watson (TWG Co-Chair)	McGill	Associate Professor
Sorana Morrissy	U of Calgary	Assistant Professor
Thomas Belbin	MUN	Associate Professor
Lincoln Stein (DPSC Co-Chair)	OICR	Head, Adaptive Oncology
Steven Jones (DPSC Co-Chair)	BCGSC	Director of Bioinformatics
Guillaume Bourque	McGill	Professor
Jeffrey Bruce	UHN	Scientific Associate
Jennifer Chan	U of Calgary	Director / Associate Professor
Karen Cranston	CanDIG	Technical Project Manager
Daniel Gaston	Dalhousie	Lead, Bioinformatician
Benjamin Haibe-Kains	U of Toronto	Associate Professor
Martin Hirst	UBC	Senior Scientist
Anne-Marie Mes-Masson (BWG Co-Chair)	U of Montreal	Associate Scientific Director
Jessica Nelson	BCGSC	Projects Team Leader
Carolyn-Ann Robinson	U of Calgary	Senior Research Associate
Enrique Sanz-Garcia	UHN	Assistant Professor / Staff Medical Oncologist
Lillian Siu	UHN	Senior Scientist
Tran Truong	UHN	Director of Data & Technology
Ian Watson	McGill	Associate Professor
Ma'n Zawati	McGill	Assistant Professor
Madeleine Arseneault	McGill	Research Assistant
Veronique Barres	CHUM	
Jonathan Bush	UBC/BCCH	Clinical Associate Professor / Investigator

Christine Caron	CHUM	Project Manager
Stephanie Crapoulet	Vitalite NB	Lead
Sidney Croul	NSHA	Lead, Pathologist
Viktor Deineko	Northern Health	Research Program Manager
Kathryn Graham	Impact Action Lab	Executive Director
Carrie Hirst	BCGSC	
Sachin Katyal	CCMB	Associate Professor / Senior Scientist
Ilinca Lungu	OICR	Manager, Tissue Portal / Research Technician
Liliane Meunier	CHUM	Research Assistant / Co-responsible for the Molecular Pathology Platform
Andy Mungall	BCGSC	Group Leader, Biospecimen and library cores / Staff Scientist
Danielle Simonot	AHS	Biobank Lead
Thomas Sontag	HSJ	Biobank Manager
Emily Van de Laar	UHN	Data Automation Architect
Suzanne Vercauteren	UBC/BCCH	Clinical Professor / Investigator
Ben Wang	UHN	Head, Immune Profiling Team