The Hospital for Sick Children
Technology Assessment at SickKids (TASK)

FULL REPORT

MICROCOSTING OF WHOLE GENOME SEQUENCING (WGS) OF TRIOS IN A HETEROGENEOUS PEDIATRIC CARDIAC POPULATION

Authors:

Jathishinie Jegathisawaran, MHEcon
Clinical Research Project Coordinator, Child Health Evaluative Sciences, The Hospital for Sick Children, Toronto, Canada

Kate Tsiplova, MSc
Research Project Manager, Child Health Evaluative Sciences, The Hospital for Sick Children, Toronto, Canada

Robin Hayeems, ScM, PhD
Scientist, Child Health Evaluative Sciences, The Hospital for Sick Children, Toronto, Canada
Associate Professor, Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, Canada

Wendy J. Ungar, MSc, PhD
Senior Scientist, Child Health Evaluative Sciences, The Hospital for Sick Children, Toronto, Canada
Professor, Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, Canada

Report No. 2020-01

Date: September 21, 2020

Available at: https://lab.research.sickkids.ca/task/reports-theses/

Co-investigators:

Eriskay Liston, MS, CGC
Genetic Counsellor, Cardiac Genome Clinic, Ted Rogers Centre for Heart Research, Toronto, Canada

Dr. Miriam Reuter, MD
Research Associate, Ted Rogers Centre for Heart Research, Cardiac Genome Clinic, The Hospital for Sick Children Research Institute, Toronto, Canada

Roozbeh Manshaei, PhD
Facility Manger, Scientific Lead, The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, Canada

Iris Cohn, MSc (Pharm)., R.Ph
Clinical Research Pharmacogenetics Advisor, Clinical Pharmacology and Toxicology & Translational Medicine, The Hospital for Sick Children, Toronto, Canada

Dr. Rebekah Jobling, MD
Staff Physician, Clinical and Metabolic Genetics, The Hospital for Sick Children, Toronto, Canada

Dr. Raymond Kim, MD, PhD
Scientific Lead, Ted Rogers Centre for Heart Research
Medical Geneticist, The Hospital for Sick Children, Toronto, Canada

Collaborator:

Dr. Seema Mital, MD, FACC, FAHA, FRCP(C)
Professor of Pediatrics, Head of Cardiovascular Research & Staff Cardiologist, Heart Function and Transplant Program, The Hospital for Sick Children, Toronto, Canada
Senior Scientist, Genetics & Genome Biology, SickKids Research Institute, Toronto, Canada

Acknowledgements:

This research was supported by a grant for The Hospital for Sick Children's (SickKids) Centre for Genetic Medicine and a grant from the PhRMA Foundation. Wendy J. Ungar is supported by the Canada Research Chair in Economic Evaluation and Technology Assessment in Child Health. We wish to thank Viji Venkataramanan for her knowledge expertise and Stephanie Luca for her assistance with quality checks.

List of Abbreviations

ACMG	American College of Medical Genetics and Genomics
ASD	Autism spectrum disorder
BWA	Burrows-Wheeler Aligner
CAD	Canadian dollar
CCA	Cost-consequence analysis
CGC	Cardiac Genome Clinic
CHD	Congenital heart defects
CI	Confidence interval
CIHI	Canadian Institute for Health Information
CMP	Cardiomyopathy
CNV	Copy number variant
DD	Developmental delay
DSA	Deterministic sensitivity analysis
GATK	Genome Analysis Toolkit
GE^3LS	Genomics and its ethical, economic, environmental, legal, and social aspects
HAS	HiSeq Analysis Software
HF	Heart failure
HTA	Health technology assessment
Indels	Insertions and deletions
MIS	Management Information Systems
MOH	Ontario Ministry of Health and Long-Term Care
PA	Probabilistic analysis
qPCR	Real-time polymerase chain reaction
SNV	Single nucleotide variant
SV	Structural variant
TPMT	Thiopurine S-methyltransferase
TRCHR	Ted Rogers Centre for Heart Research
WGS	Whole genome sequencing

Table of Contents

List of Tables iii

List of Figures iv

Executive Summary v

1 Introduction 1

1.1 Background 1

1.2 Study objectives 3

2 Methods 3

2.1 Study design 3

2.2 Microcost item identification 7

2.3 Microcost item valuation 8

2.3.1 Whole genome sequencing 8

2.4 Assumptions 17

2.5 Microcosting analysis 20

2.5.1 Probabilistic analysis 20

2.5.2 Sensitivity analysis 21

3 Results 22

3.1 Test costs per patient with a cardiac condition 22

3.2 Program costs for heterogeneous pediatric cardiac population 25

3.3 Deterministic sensitivity analysis 26

4 Discussion 29

5 Conclusion 31

References 32

List of Tables

Table 1. Categories of Resource Use 7

Table 2. Whole Genome Sequencing-trio (Illumina HiSeq X™) Parameter Estimates and Distributions Used in the Probabilistic Analysis 13

Table 3. Assumptions: Microcosting Analyses 19

Table 4. Estimated Annual Cost per Cardiac Genome Clinic Trio for Whole Genome Sequencing (Illumina HiSeq X™) 23

Table 5. Deterministic Sensitivity Analysis of Estimated Total Cost per Cardiac Genome Clinic Trio for Whole Genome Sequencing, Varying Overhead Cost Proportion. 28

List of Figures

Figure 1: Whole Genome Sequencing – trio: Process Flow for Platform HiSeq X™ 5
Figure 2: Coordination Process for Patient Recruitment 6
Figure 3. Proportion of Total Annual Cost for Cardiac Genome Clinic Whole Genome Sequencing - trio (Illumina HiSeq X™) by Cost Category, Year 1. 24
Figure 4. Present Value of Program Costs Over Five Years for Whole Genome Sequencing – trio (HiSeq X™). 25
Figure 5. Deterministic Sensitivity Analysis of the Effect of Increasing the Number of Whole Genome Sequencing-trio (HiSeq X™) Tests per Year for All Indications from 500 Trios (1500 samples) to 1000 Trios (3000 samples) on Trio Sample Costs in Year 1. 27

Executive Summary

Background

Multiple causes of heart failure (HF) in children and the challenges in treatment and management of care has increased demand for whole genome sequencing (WGS). WGS captures information that could help determine the cause of or risk factors associated with HF not only for patients but also for family members. Identifying a genetic cause or risk factor can, in turn, aid in clinical decisions related to screening, treatment and management. An economic evaluation of WGS technology requires a comprehensive and accurate estimation of all costs involved in the sequencing workflow. This would aid in policy and implementation decisions of this technology for the pediatric HF patient population.

Objectives

The objective of this study was to estimate costs per trio for WGS, including coding and non-coding regions, for a targeted patient population consisting of children with heterogeneous cardiac diseases including cardiomyopathies (CMP), congenital heart defects (CHD) and inherited cardiac arrhythmias enrolled in the cardiac genome clinic (CGC) at The Hospital for Sick Children (SickKids), Toronto, Canada from an institutional payer perspective over five years.

Methods

Using a bottom-up microcosting approach, the opportunity cost per trio excluding mark-ups, fees and charges for WGS-trios on the Illumina HiSeq X™ platform for pediatric patients with multiple cardiac diseases was estimated. This was done from an institutional payer perspective based on the diagnostic laboratory practices at SickKids. The cost per trio was determined for each year of a five-year program. Total program costs to service the CGC pediatric population were also estimated over five years. A probabilistic analysis (PA) was conducted to incorporate parameter uncertainty in the model. Three one-way deterministic sensitivity analyses (DSA) were conducted to examine the effects of changing the inputs for the overhead cost, the total volume of WGS tests in the institution, and excluding pharmacogenomics while other inputs remained the same.

Results

The cost per trio in Year 1 was $8053 (95% confidence interval [CI]: 7699, 8558) for WGS-trio (HiSeq X™). Reagent supply costs accounted for the largest proportion of costs (50%) followed by bioinformatics (25%). The total institutional program cost to offer WGS for CGC diagnosis over five years was $5.63

million (95% CI: 5.38, 5.98) based on 144 CGC trios per year. Varying the inputs in DSAs resulted in a minimal difference of under 5% in the overall costs per WGS-trio.

Conclusions

This study estimated the cost of trio WGS using a bottom-up microcosting approach. The study provides comprehensive cost data for use in future economic evaluations of genome sequencing in pediatric cardiac patients. It allows for a costing model that can be easily updated as technology evolves and adapted to other pediatric patient populations. Additional analyses are required to assess the clinical and economic impact of the WGS in this population.

1 Introduction

1.1 Background

Whole genome sequencing (WGS) is an emerging technology with potential for increased diagnostic accuracy, and improved management and care for multiple diseases. WGS, due to its comprehensive nature, provides detailed information about a patient's genome. It detects both small and large *de novo* and inherited variations in coding and noncoding regions of DNA, including copy number variants (CNVs) and small nucleotide variants (SNVs) (1, 2). Additionally, novel, causative mutations of rare or common Mendelian disorders have been identified through the use of this technology. WGS can generate findings unrelated to the purpose of the test, including secondary or incidental findings, that may predict risk for other conditions and have a significant impact on a patient's health (3). WGS can also identify pharmacogenomic variants associated with medication metabolism or sensitivities (4).

In Ontario, WGS has been primarily used in research settings to understand genetic causes and subsequent management strategies of diseases such as autism spectrum disorder (ASD) and other heterogeneous developmental delay (DD) disorders (5). Previous microcosting and cost-consequence analysis (CCA) of ASD and DD have facilitated further analyses and funding decisions in the province of Ontario (6). Given the utility of the information provided by WGS for both clinical and policy decisions, it is a worthwhile investment to study the use of WGS in other disease groups.

Hereditary heart failure (HF) comprises a group of diseases for which understanding the genetic etiology and burden of disease is essential for optimizing management and care (7). Cardiomyopathies (CMP), congenital heart defects (CHD) and inherited cardiac arrhythmias such as cardiac channelopathies can cause HF in both pediatric and adult populations. Genetic variants and epigenetic changes may explain clinical phenotypes that are challenging to diagnose, treat, and manage. The Cardiac Genome Clinic (CGC) within Ted Rogers Centre for Heart Research (TRCHR) was established in 2016 at The Hospital for Sick Children (SickKids) to investigate the genetic causes and provide treatment and management options for CMP, CHD and cardiac arrhythmias. The CGC represents a collaboration between the Division of Clinical and Metabolic Genetics and the Division of Cardiology. Part of CGC's mission is to address challenges of implementing clinical WGS for this population.

The CGC is conducting prospective clinical research investigating causal variants of pediatric cardiac conditions (discovery research) as well as health services research aimed at facilitating implementation of clinical WGS in pediatric HF patients. The facility is using the infrastructure and principles of the SickKids Genome Clinic that was established in 2016 (8). CGC inclusion criteria are: a) children clinically identified to have a cardiomyopathy (CMP), congenital heart defect (CHD) or cardiac arrhythmia with a suspected genetic etiology b) children who had conventional targeted genetic testing related to one or more of these conditions (panel testing or microarray) for which the results were negative; c) new patients identified to have one of these conditions for whom WGS was indicated as a first line genetic test; d) patients for whom a genetic etiology has not been established or a genetic etiology has been established but with a wide phenotypic variability in the family. Eligible patients receive cardiac and systemic phenotyping (i.e. precise and comprehensive analysis of phenotypic abnormalities) and are offered WGS where inclusion criteria are met. Trio-based WGS is then conducted (i.e. proband plus biological parents undergo WGS). This sequencing method enhances the speed and likelihood of accurate diagnosis by decreasing the number of candidate variants (9) that the analyst needs to adjudicate. Diagnostic rate can also be improved by tailored, comprehensive manual medical review that relies on a frequently updated gene/phenotype database rather than depending on a pre-set phenotype driven gene list. This minimises the chances of missing *de novo* mutations (10).

Health technology assessment (HTA) is a standard process for evaluating emerging health care technologies including diagnostic tests. As a component of the GE^{3}LS (genomics and its ethical, economic, environmental, legal, and social aspects) domain, HTA of genomic sequencing technologies is essential to generating high quality evidence to support policies that are equitable and that maximize health benefits to the population. An economic evaluation is a core part of HTA and compares the costs and consequences of new technologies to standard care to quantify the additional costs associated with the technology per unit of added benefit, thus providing insight into whether these technologies add value for money and are appropriate to adopt into clinical practice (11, 12).

In precision medicine, optimizing diagnosis, and management and care choices depend not only on achieving good clinical outcomes but also on economic value. Policy and reimbursement decisions regarding WGS as a health technology should take into consideration both clinical and economic evidence. While the laboratory costs of sequencing have decreased dramatically in recent years (13, 14), there is a paucity of studies that comprehensively estimate actual test costs. Full economic evaluation of

WGS technology that weigh the incremental costs of WGS against its incremental benefits to patients require accurate estimations of all costs involved in the workflow (15).

1.2 Study objectives

The objective of this study was to estimate the precise cost per trio for WGS using a microcosting approach for a targeted patient population consisting of children with CHD, cardiac arrhythmia or CMP. In the microcosting approach, the volume of use and unit price of each resource use component was estimated (16) and the entire workflow process of a genetic test was tracked. This microcosting study did not include consideration of diagnostic yield or other lab performance metrics.

2 Methods

2.1 Study design

WGS consisted of multiple workflow components. Figure 1 illustrates all of the components of this technical pathway from specimen preparation to clinical interpretation, including case review and confirmatory testing.

Using a bottom-up microcosting approach, the opportunity cost per trio excluding mark-ups, fees and charges for WGS tests for patients with a range of cardiac conditions were estimated for each component in the workflow process. This was done from an institutional payer perspective based on the diagnostic laboratory practices at SickKids, Toronto, Canada. In addition to WGS, analysis of pharmacogenomics variants was performed. This analysis followed the clinical guidelines annotations from the Clinical Pharmacogenetics Implementation Consortium (CPIC), the Royal Dutch Association for the Advancement of Pharmacy - Pharmacogenetics Working Group (DPWG), the Canadian Pharmacogenomics Network for Drug Safety (CPNDS) and other professional societies (17) and was conducted in probands (patients) in contrast to WGS analysis which was done on trios. Only the variants with level 1A clinical annotation published on the curated PharmGKB database were analyzed (18). Pharmacogenomics testing to validate WGS data was done outside of SickKids. Secondary variants as per American College of Medical Genetics and Genomics (ACMG) guidelines (19) were identified and confirmatory testing was done by a lab outside of SickKids.

The total cost per trio was determined for each year of a five-year program. Total SickKids program costs to service the pediatric cardiac patient population were also estimated over five years.

Figure 1: Whole Genome Sequencing – trio: Process Flow for Platform HiSeq X™

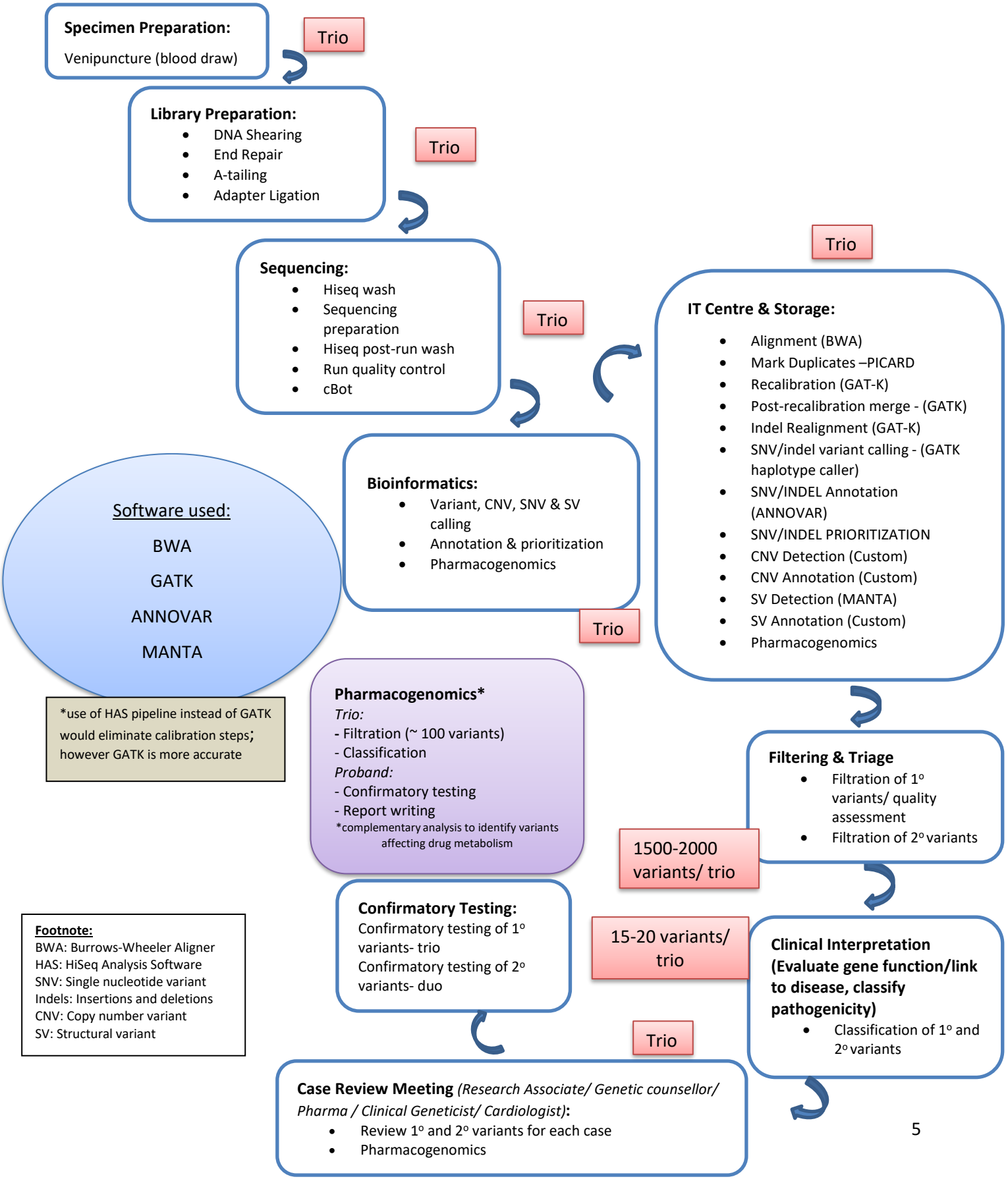
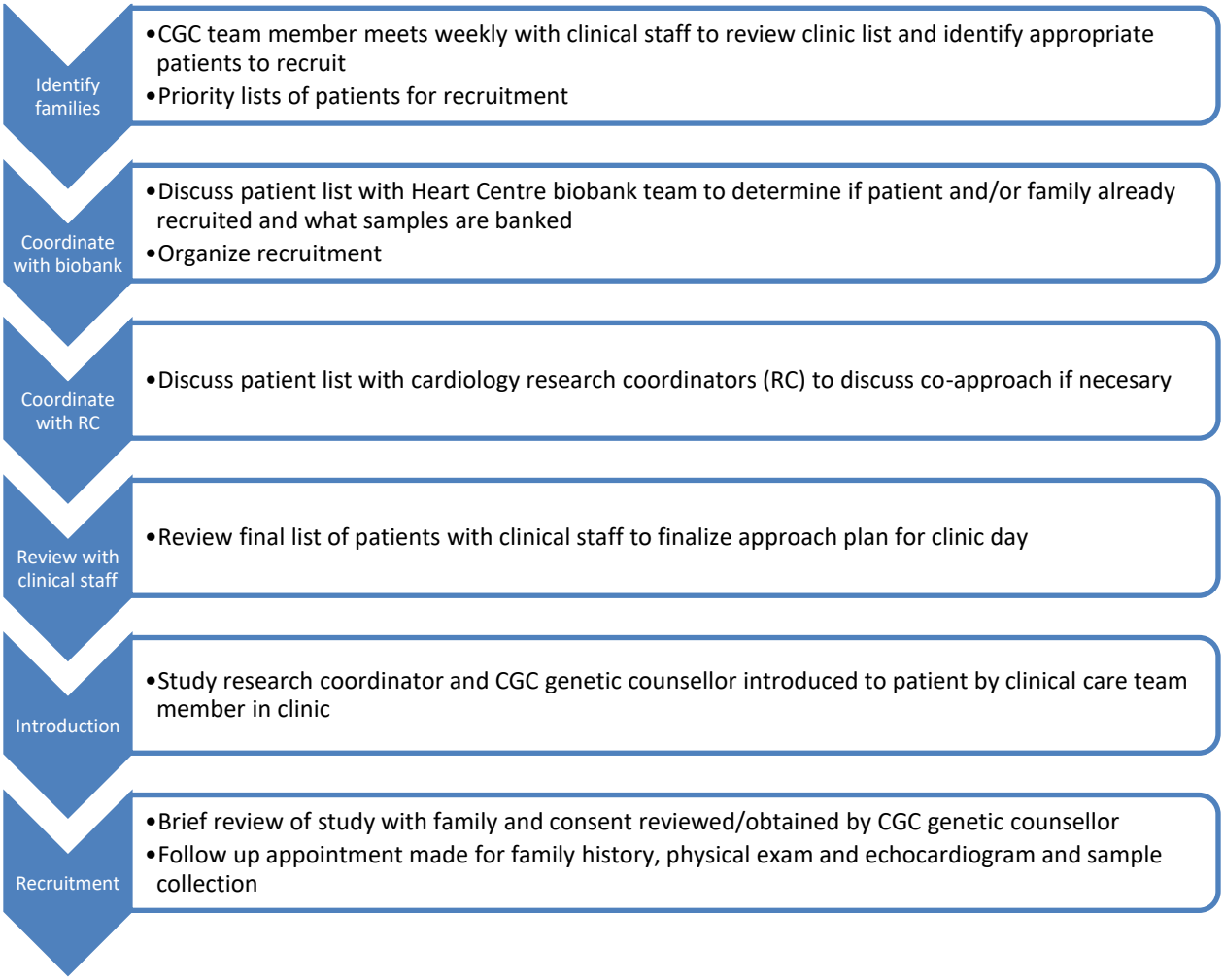


Figure 2 outlines the clinical research patient recruitment process at the CGC. Over the last 2.5 years, the study recruited 120 families with children who have CHD, cardiac arrhythmias or CMP. Although microcosting was performed within the context of clinical research, as the technology is in the process of translation, estimates are presented as a reasonable representation of costs in clinical practice.

Figure 2: Coordination Process for Patient Recruitment



2.2 Microcost item identification

The major cost categories were labour, small and large equipment, supplies, confirmatory testing and bioinformatics. The last category reflects the large computing component of WGS. A list of major categories and sub-categories for each technology is presented in Table 1. Each of the sub-categories was further broken down into individual microcost items according to SickKids laboratory operating procedures which are described in detail in Table 1.

Table 1. Categories of Resource Use

Major Category	Minor Category
Labour	Specimen preparation
	Library preparation
	Sequencing
	Bioinformatics
	IT centre & storage
	Filtration and triage
	Clinical interpretation
	Case review meeting
	Confirmatory testing
	Pharmacogenomics
Supplies	Sample handling
	Consumables
	Reagents
Confirmatory Testing	Sanger sequencing of primary and secondary variants
	Agena MassARRAY® of pharmacogenomics variants
Bioinformatics	Bioinformatics file storage
	Bioinformatics computational use
Small Equipment	Small equipment
Large Equipment	Sequencing equipment
	Equipment contract

Abbreviations: WGS, Whole genome sequencing

2.3 Microcost item valuation

2.3.1 Whole genome sequencing

The present analysis included estimates for a trio WGS done on the HiSeq X™ platform. This sequencer can sequence 16 samples per run to achieve a 30-45X read depth. The Illumina HiSeq X™ requires a large initial investment, with relatively low supply costs for the machine. Table 2 contains resource use and price data for HiSeq X™, for trio testing. The specimen preparation, library preparation and sequencing took at place at The Centre for Applied Genomics (TCAG), SickKids. Bioinformatics analysis, filtering & triage and clinical interpretation were conducted by a TRCHR bioinfomatician.

The resource use and unit price data for each input were provided by the laboratory staff, industry or extracted from published or grey literature such as the Canadian Institute for Health Information's (CIHI) *Standards for Management Information Systems in Canadian Health Service Organizations* ("MIS Standards") (20). Where possible and appropriate, a range encompassing all plausible values of an input's resource use and unit price was provided in addition to a point estimate. Costs for each input were calculated by multiplying resource use by unit price. For labour, time in minutes for each task was multiplied by wage rates. Price estimates from different reporting years were used for costing of the individual items (2014 to 2019).

2.3.1.1 Labour

Total minutes for each input in the specimen preparation, library preparation and sequencing categories were determined for a single sample. These values were tripled for trios since the number of samples processed per run is three times that of a proband. The labour time per sample for each input in library preparation and sequencing categories was calculated by dividing the total time per task by the number of samples processed per run. Due to automation, HiSeq X™ can process 48 samples during the library preparation and can sequence 16 samples per run.

Labour resource use and prices were estimated for the analysis of sequenced data and the maintenance of the high performance computing cluster at the SickKids' Centre for Computational Medicine. The overall output range for one HiSeq X™ instrument is 60-75 genomes per month. The resource use per sample for variant, CNV, SNV and SV calling and annotation and prioritization was calculated by dividing the labour time by the average of 67.5 genomes per month. Based on expert opinion, 1.25 FTE units of labour is required to process this range of genome output per month (personal communication, R.

Manshaei). This labour time was assumed to include sample logistics management (i.e. starting computing jobs, tracking samples, transferring data) as well as data processing (i.e. periodic updates to the annotation pipeline). The resource use per sample for bioinformatics was calculated by dividing the labour time by the average output per month. For pharmacogenomics analysis, the estimated labour time required was fixed at 30 minutes.

Bioinformatics maintenance components for HiSeq X™ pipelines steps included: alignment (Burrows-Wheeler Aligner) (BWA), mark duplicates (PICARD), recalibration (GATK), post-recalibration merge (GATK), indel realignment (GATK), SNV/indel variant calling (GATK HaplotypeCaller), SNV/indel annotation (ANNOVAR), SNV/INDEL prioritization, CNV detection (custom), CNV annotation (custom), SV detection (MANTA), SV annotation (custom) and pharmacogenomics analysis. The calculation time and the number of nodes required for each step in the bioinformatics pipeline were obtained from the TRCHR bioinformatician. One hour of labour was assumed to be required to support one node per year (13). The bioinformatics maintenance labour resource use in minutes was estimated by multiplying the calculation time by the time needed to support the required number of nodes. Ranges for labour volume use were provided.

An extensive list of genetic variants provided by the bioinformatician was screened by the research associate in the filtering and triage step. In this step, approximately 1500-2000 variants were examined to flag 15-20 variants of interest for both primary and secondary variants. Time required for filtration of primary variants was 90 minutes and for secondary variants, the requirement was 10 minutes (personal communication, M. Reuter). For pharmacogenomics, the required time for filtration was 10-25 minutes for a list of 100 variants (personal communication, I. Cohn).

The 15-20 filtered variants were further investigated and interpreted for effect on gene function, possible link to disease and classification of pathogenicity by the genome analyst using in silico analyses. The same variants were also classified as secondary variants if they met pre-classified ACMG criteria (21). In a subsequent bi-weekly case review meeting, the trio's primary and secondary variants were discussed. Four key personnel, namely the genome analyst, genetic counsellor, clinical geneticist and cardiologist contributed to discussion of the results regarding the trio in question. Average time required for this discussion was 15 minutes per family of trio for all staff. Confirmatory testing of

primary and secondary variants through Sanger sequencing was performed in 23.3% (11.2 families) of 48 families (144 samples) for whom trio WGS is done in a year (personal communication, E. Liston).

A similar process was undertaken for pharmacogenomics analysis, which is at present, in the CGC, only conducted for probands. Following the bioinformatics step, analysis was done by the clinical research pharmacogenetics advisor to predict specific haplotypes and phenotypes. Filtration of approximately 100 cardiac variants of the probands tested to date required an average of 17.5 minutes. Majority of the variants (87.5%) required 22 minutes of classification time, whereas more complicated variants (12.5%) needed up to 60 minutes (personal communication, I.Cohn). A staff pharmacist also spent fifteen minutes per family to discuss the results of pharmacogenomics analysis at the case review meeting. Confirmatory testing was done in 90% of the variants. As a last step, a report was written up by the staff pharmacist.

Hospital and lab employees involved in WGS testing included nurses, lab technicians, lab technologists, bioinformatics analysts, high performance computing staff, genetic counsellor, research associate, clinical geneticist, staff cardiologist and staff pharmacist. Labour prices reported were from 2019; if 2019 estimates were not found, previously available salaries were adjusted with a yearly increase of 1.75%. Benefits at SickKids were calculated by applying 26% to the hourly wage. For the clinical geneticist and staff cardiologist, a per minute cost was determined from the Ontario Schedule of Benefits (ODB) (22). Salaries of other staff members were obtained by either an informal survey of lab staff, reported salaries from SickKids or an employment website (23). Because of the confidential nature of the salary information, reporting of unit prices (wages) for labour items has been suppressed. For most of the inputs, the salary range was based on the SickKids salary scale, lab staff survey or expert opinion within the lab. When it was not possible to obtain a salary range from these sources, salaries were assumed to vary by 20% from its point estimates. There were no ranges for the hourly wages obtained from the ODB. All prices were reported in 2019 Canadian dollars (CAD).

2.3.1.2 Equipment

The large equipment costs were estimated for the HiSeq X™ sequencing platform of Illumina (San Diego, USA) and included the cost of the platform, its maintenance contract and Bioanalyzer and TapeStation instruments made by Agilent Technologies Inc. (Santa Clara, USA). The price for one HiSeq X™ instrument was based on the assumption that five sequencers were purchased at SickKids (Table 2). The

maintenance contract was 10.35% of the cost of the sequencer per year. A price range was provided for a Bioanalyzer and TapeStation by the manufacturer and TCAG lab manager.

Small equipment consisted of the tube microcentrifuge, plate microcentrifuge, thermomixer, vortex, pipette sets, magnet particle concentrator, and thermocycler. Small equipment prices were estimated by the TCAG lab manager and need to be replaced every five years. For this model with a time horizon of five years, one replacement was required. Since thermocyclers and pipette sets are replaced every 2.5 years, their resource use was calculated using the following formula: 2/all tests per year for all indications. This was to indicate that they were replaced twice in five years. The price ranges for large equipment and for some of the small equipment were based on the expert opinion of the TCAG lab manager. For items without price ranges, unit prices were assumed to vary by 10% from their point estimates. To account for price and currency fluctuations, unit prices of small equipment were assumed to vary by 10% from their point estimates. Unit prices of platforms and maintenance contracts of large equipment were given ranges by experts or were varied by 10%. The sample costs for CGC patients for large and small equipment was determined by allocating the proportion of use by CGC patients of all patients with all indications in the institution.

2.3.1.3 Supplies

Supplies included costs of shipping a sample to TCAG laboratory, Illumina Nano DNA library preparation reagents, other library preparation consumables and reagents and HiSeq X™ sequencing reagents (Table 2). Ranges for resource use were not assigned, as it was assumed that one unit of supplies was required per sample. This resource use was tripled to account for the estimation of trios. To account for price and currency fluctuations, unit prices of shipping and handling, library preparation and sequencing reagents were assumed to vary by 10% from their point estimates.

2.3.1.4 Confirmatory testing

Confirmatory testing for both primary and secondary variants was completed through Sanger sequencing. One variant from one gene was sent for validation. Since the inception of the project in November 2016 until July 2019, 120 families were evaluated. Of these 120, samples from 28 families were sent for confirmatory testing of both primary and secondary variants. Therefore, on an annualized basis, of the 48 enrolled families, 11.2 families (23.3%) received validated results for primary or

secondary variants (personal communication, E. Liston). Confirmatory testing of pharmacogenomics variants has been contracted out to an external lab by SickKids where the testing is performed on the Agena MassARRAY® platform and by next generation sequencing (NGS). MassARRAY is completed for the selected genetic variants of the *CYP2C19, CYP2C9*, *CYP2D6, CYP3A5 and VKORC1* genes. NGS is performed on thiopurine S-methyltransferase (TPMT) genetic variants. Of the ninety probands who have had pharmacogenomic analysis in one year, 74.4% have been validated.

2.3.1.5 Bioinformatics

The costs calculated in this category were for bioinformatics data file storage including trimmed fastq and recalibrated, locally re-aligned BAM files, and computational use for both WGS and pharmacogenomics analysis. Software costs were not included as GATK is an open source software with no licensing fee. The resource use for the data file storage depended on file size and length of storage time and was calculated in gigabytes per year. The resource use (CPU per hour) for each step was calculated by multiplying the number of computing jobs by the number of CPUs (cores) by the time (in hours) required to complete the job and by 25% extra time allocated as waiting period for the saturated nodes to become available. This extra time was varied as 0% and 50% for lower and upper bounds, respectively. The resource use estimates, along with ranges, were obtained from the bioinformatics manager. Unit prices for storage and computational use had a 10% range.

The cost was estimated for storage use and the computation use for the pipeline steps specified in Section 2.3.1.1. Prices for computational use were based on TCAG's purchase of 72 compute nodes (20 cores, 40 threads, 256 GB RAM) for processing WGS samples on HiSeq X™. The price of each node was $26,804 CAD over five years, including warranty.

Table 2. Whole Genome Sequencing-trio (Illumina HiSeq X™) Parameter Estimates and Distributions Used in the Probabilistic Analysis

Cost Items	Quantity of Use per Sample		Unit Price (CAD)	
	Estimate	Distribution	Estimate	Distribution
LABOUR				
Specimen Preparation (Units: minutes)				
Pediatric venipuncture	22.8	Fixed	Conf.	Trun. Normal μ,σ=Conf.
Packaging with testing documentation	3.0	Fixed	Conf.	Trun. Normal μ,σ=Conf.
Service recipient primary registration	5.4	Fixed	Conf.	Trun. Normal μ,σ=Conf.
Printing and sorting of specimen labels	1.2	Fixed	Conf.	Trun. Normal μ,σ=Conf.
Creation of recipient folder	15.0	Fixed	Conf.	Trun. Normal μ,σ=Conf.
Packaging with testing documentation	3.0	Fixed	Conf.	Trun. Normal μ,σ=Conf.
Service recipient limited registration	5.4	Fixed	Conf.	Trun. Normal μ,σ=Conf.
Phenotips Entry	17.5	Trun. Normal μ=17.5,σ=0.83;	Conf.	Trun Normal μ,σ=Conf.
Library Preparation (Units: minutes)				
DNA quantification	1.3	Fixed	Conf.	Trun. Normal μ,σ=Conf.
Pre-prep reagents	1.3	Fixed	Conf.	Trun. Normal μ,σ=Conf.
Shearing	1.3	Fixed	Conf.	Trun. Normal μ,σ=Conf.
Purification	2.5	Fixed	Conf.	Trun. Normal μ,σ=Conf.
End repair	2.5	Fixed	Conf.	Trun. Normal μ,σ=Conf.
A-tailing	2.5	Fixed	Conf.	Trun. Normal μ,σ=Conf.
Adapter ligation	2.8	Fixed	Conf.	Trun. Normal μ,σ=Conf.
Sequencing (Units: minutes)				
HiSeq wash	5.6	Fixed	Conf.	Trun. Normal μ,σ=Conf.
Sequencing prep	5.6	Fixed	Conf.	Trun. Normal μ,σ=Conf.
HiSeq post-run wash	8.4	Fixed	Conf.	Trun. Normal μ,σ=Conf.
Run quality control	2.8	Fixed	Conf.	Trun. Normal μ,σ=Conf.
cBot	5.6	Fixed	Conf.	Trun. Normal μ,σ=Conf.

Bioinformatics (Units: minutes)				
Variant, CNV, SNV & SV calling (total time per month/samples per month)	124.44	Total time fixed (8400 minutes); Samples per month: Trun. Normal μ=67.5,σ=4.67	Conf.	Trun. Normal μ,σ=Conf.
Annotation & prioritization (total time per month/samples per month)	31.11	Total time fixed (2100 minutes); Samples per month: Trun. Normal μ=67.5,σ=1.17	Conf.	Trun. Normal μ,σ=Conf.
Pharmacogenomics analysis	30	Fixed	Conf	Trun. Normal μ,σ=Conf
IT Centre & Storage (Units: minutes)				
Alignment (BWA)	0.2466	Trun. Normal μ=0.25,σ=0.02054	Conf.	Trun. Normal μ,σ=Conf.
Mark Duplicates – PICARD	0.0051	Trun. Normal μ=0.01,σ=0.000428	Conf.	Trun. Normal μ,σ=Conf.
Recalibration – GAT-K	0.0822	Trun. Normal μ=0.08,σ=0.00685	Conf.	Trun. Normal μ,σ=Conf.
Post-recalibration merge (GAT-K)	0.0010	Trun. Normal μ=0.0010,σ=0.0000856	Conf.	Trun. Normal μ,σ=Conf.
Indel realignment (GAT-K)	0.0616	Trun. Normal μ=0.06,σ=0.00514	Conf.	Trun. Normal μ,σ=Conf.
SNV/indel variant calling (GAT-K)	0.1233	Trun. Normal μ=0.12,σ=0.01027	Conf.	Trun. Normal μ,σ=Conf.
SNV/indel Annotation (ANNOVAR)	0.0062	Trun. Normal μ=0.01,σ=0.000513	Conf.	Trun. Normal μ,σ=Conf.
SNV/indel Prioritization	0.0003	Trun. Normal μ=0.0003,σ=0.0000214	Conf.	Trun. Normal μ,σ=Conf.
CNV detection (Custom)	0.0154	Trun. Normal μ=0.02,σ=0.00128	Conf.	Trun. Normal μ,σ=Conf.
CNV annotation (Custom)	0.000086	Trun. Normal μ=0.0000856,σ=0.00000713	Conf.	Trun. Normal μ,σ=Conf.
SV detection (MANTA)	0.0205	Trun. Normal μ=0.02,σ=0.00171	Conf.	Trun. Normal μ,σ=Conf.
SV Annotation (Custom)	0.0000856	Trun. Normal μ=0.0000856,σ=0.00000713	Conf.	Trun. Normal μ,σ=Conf.
Pharmacogenomics Analysis	0.001027	Trun. Normal μ=0.0000856,σ=0.00000713	Conf.	Trun. Normal μ,σ=Conf.
Filtering & Triage (Units: minutes)				
Filtration of primary variants	90.0	Fixed	Conf.	Trun. Normal μ,σ=Conf.
Filtration of secondary variants	10.0	Fixed	Conf.	Trun. Normal μ,σ=Conf.
Clinical Interpretation (Units: minutes)				
Classification of primary variants	120	Log. Normal μ=4.34,σ^2=0.955	Conf.	Trun. Normal μ,σ=Conf.
Classification of secondary variants	12.5	Trun. Normal μ=12.5,σ=0.83	Conf.	Trun. Normal μ,σ=Conf.

Case Review Meeting (Units: minutes)				
Research Associate	60	Trun. Normal μ=15,σ=1.67	Conf.	Trun. Normal μ,σ=Conf.
Genetic Counsellor	60	Trun. Normal μ=15,σ=1.67	Conf.	Trun. Normal μ,σ=Conf.
Clinical Geneticist	60	Trun. Normal μ=15,σ=1.67	Conf.	Trun. Normal μ,σ=Conf.
Cardiologist	60	Trun. Normal μ=15,σ=1.67	Conf.	Trun. Normal μ,σ=Conf.
Confirmatory Testing (Units: minutes)				
Primary variants	4.67	Fixed; Beta α=33.12,β=110.88	Conf.	Trun. Normal μ,σ=Conf.
Secondary variants	4.67	Fixed; Beta α=33.12,β=110.88	Conf.	Trun. Normal μ,σ=Conf.
Pharmacogenomics (Units: minutes)				
Filtration	17.5	Trun. Normal μ=17.5,σ=2.5;	Conf.	Trun. Normal μ,σ=Conf.
Classification	22.81	Trun. Normal μ=22.81,σ=1.23;	Conf.	Trun. Normal μ,σ=Conf.
Case Review Meeting	60	Fixed	Conf.	Trun. Normal μ,σ=Conf.
Confirmatory Testing	62.5	Trun. Normal μ=62.5,σ=19.17;	Conf.	Trun. Normal μ,σ=Conf.
Report Writing	15.075	Trun. Normal μ=15.07,σ=1.67;	Conf.	Trun. Normal μ,σ=Conf.
LARGE EQUIPMENT				
Illumina HiSeq X™	1/all tests	Fixed	1150000	Trun. Normal μ=1150000,σ=38333
1-year service contract	1/all tests	Fixed	119025	Trun. Normal μ=119025,σ=3968
Agilent BioAnalyzer/Tape station	1/all tests	Fixed	38500	Trun. Normal μ=38500,σ=1500
SMALL EQUIPMENT				
Tube microcentrifuge	1/all tests	Fixed	2250	Trun. Normal μ=2250,σ=83.3
Plate microcentrifuge	1/all tests	Fixed	5000	Trun. Normal μ=5000,σ=166.7
Thermomixer	1/all tests	Fixed	5000	Trun. Normal μ=5000,σ=166.7
Vortex	1/all tests	Fixed	450	Trun. Normal μ=450,σ=16.7
Pipette sets	2/all tests	Fixed	1600	Trun. Normal μ=1600,σ=101.2
Magnet particle concentrator for tubes	1/all tests	Fixed	700	Trun. Normal μ=700,σ=23.3
Thermocyclers	2/all tests	Fixed	3000	Trun. Normal μ=3000,σ=101.2
SUPPLIES (Units: counts)				
Shipping & Handling	1	Fixed	37.61	Trun. Normal μ=37.61,σ=1.25
Illumina Nano DNA library prep	3	Fixed	30.0	Trun. Normal μ=30.0,σ=1.0
Other library prep consumables	3	Fixed	50.0	Trun. Normal μ=50,σ=1.67
Sequencing reagents	3	Fixed	1290	Trun. Normal μ=1290,σ=43.0

CONFIRMATORY TESTING (proportion of patients)				
Sanger sequencing (primary & secondary)	0.23	Beta α=33.12,β=110.88	375.0	Trun. Normal μ=375.0,σ=41.67
Pharmacogenomics (MassARRAY & NGS)	0.74	Beta α=67,β=23	147.0	Trun. Normal μ=147.0,σ=0
BIONFORMATICS				
Bioinformatics File Storage (Units: GB per year)				
Trimmed fastq	720.0	Trun. Normal μ=720.0,σ=0	0.40	Trun. Normal μ=0.40,σ=0.013
final rem-dup, recalibrated, locally re-aligned BAM file	1530.0	Trun. Normal μ=1530.0,σ=0	0.40	Trun. Normal μ=0.40,σ=0.013
Bioinformatics Computation Use (Units: CPU time per hour)				
Alignment (BWA)	855	Trun. Normal μ=855,σ=57	0.612	Trun. Normal μ=0.612,σ=0.0204
Mark Duplicates – PICARD	37.5	Trun. Normal μ=37.5,σ=2.5	0.612	Trun. Normal μ=0.612,σ=0.0204
Recalibration (GAT-K)	172.5	Trun. Normal μ=172.5,σ=11.5	0.612	Trun. Normal μ=0.612,σ=0.0204
Post-recalibration merge (GAT-K)	7.5	Trun. Normal μ=7.5,σ=0.5	0.612	Trun. Normal μ=0.612,σ=0.0204
Indel Realignment (GAT-K)	258.75	Trun. Normal μ=258.75,σ=17.25	0.612	Trun. Normal μ=0.612,σ=0.0204
SNV/indel variant calling (GAT-K)	258.75	Trun. Normal μ=258.75,σ=17.25	0.612	Trun. Normal μ=0.612,σ=0.0204
SNV/INDEL Annotation	45	Trun. Normal μ=45,σ=3	0.612	Trun. Normal μ=0.612,σ=0.0204
SNV/INDEL Prioritization	3.75	Trun. Normal μ=3.75,σ=0.25	0.612	Trun. Normal μ=0.612,σ=0.0204
CNV Detection	112.5	Trun. Normal μ=112.5, σ=7.5	0.612	Trun. Normal μ=0.612,σ=0.0204
CNV Annotation	0.3125	Trun. Normal μ=0.3125,σ=0.02083	0.612	Trun. Normal μ=0.612,σ=0.0204
SV Detection (MANTA)	75	Trun. Normal μ=75,σ=5	0.612	Trun. Normal μ=0.612,σ=0.0204
SV Annotation (Custom)	0.3125	Trun. Normal μ=0.3125,σ=0.02083	0.612	Trun. Normal μ=0.612,σ=0.0204
Pharmacogenomics Analysis	11.25	Trun. Normal μ=11.25,σ=0.75	0.612	Trun. Normal μ=0.612,σ=0.0204

Abbreviations: WGS, Whole genome sequencing; qPCR, Real-time polymerase chain reaction; SNV, Single nucleotide variant; Conf., Confidential; Trun. Normal, Truncated normal. 'All tests' indicates the total volume of tests performed in the institution for all indications.

2.4 Assumptions

The assumptions of the microcosting model are summarized in Table 3. A time horizon of five years was selected based on the estimated useful lifetime of the small and large equipment. Exceptions were thermocyclers and pipette sets which are replaced every two and a half years. Future costs were discounted using a discount rate of 1.5% with the assumption that costs were incurred at the end of the year. Small and large equipment items were depreciated using a straight-line depreciation method. An opportunity cost of 1.5% was added to the cost of large equipment, such as sequencing machines and their maintenance contracts. The opportunity cost refers to the next best use of funds invested in equipment and is approximated by the return on the undepreciated value of equipment at each time point (24). Resource use and unit prices were assumed to remain the same from year to year. The following cost items were patient population specific: disease specific test volume, number of primary variants, number of secondary variants, genome output per month, filtration time for primary and secondary variants, interpretation time for primary and secondary variants, case review meeting, confirmatory testing and pharmacogenomics analysis. It was assumed that all other cost items did not depend on the patient population.

Some of the labour steps in the workflow process were compared to the laboratory labour components in the 2016 MIS Standards. The MIS Standards provide a standardized framework for collecting and reporting financial and statistical data on day-to-day operations of health service organizations (20). The MIS Standards provide average time required for standard laboratory activities, although time per activity may vary between institutions. The labour resource use, calculated as time in minutes per each step, was obtained from MIS Standards or directly from the relevant laboratories at SickKids.

HiSeq X™ was used for the WGS of trios. The equipment resource use per sample depends on the total number of tests done in the institution for all patients. As overall test volume increases, the equipment resource use and therefore equipment cost per CGC patient decreases. The number of trio tests conducted for this patient population was 144 per year for children and their biological parents based on the prevalence of this group of cardiac conditions (personal communication, E. Liston). Based on the approximate volume of all clinical whole genomes indicated at SickKids, it was further assumed that the annual number of WGS-trio tests for all indications could vary from 500 trios (1500 samples) to 1000 trios (3000 samples) per sequencer and was assumed to be 500 trios in the reference case.

Bioinformatics included multiple sub-categories. The labour cost associated with a bioinformatician's time to perform sample logistics and data processing was estimated. Equipment and labour costs associated with purchasing and maintaining computing nodes were also costed. Periodic validation, quality control and pipeline updating and testing were not included.

Confirmatory testing was conducted for primary and secondary variants as well as for the variants discovered during pharmacogenomics analysis. Labour costs for filtration, interpretation, case review meeting, confirmatory testing and reporting writing were estimated.

Overhead cost comprised administrative and infrastructure costs; these costs were added to labour, large and small equipment and bioinformatics costs. Overhead was not applied to supplies or confirmatory testing as supplies are bought at retail prices that include markups that cover overhead costs of the vendor. A query to the Ontario Ministry of Health's (MOH) Case Costing Initiative (OCCI) returned an estimate of overhead costs for acute inpatients of the top 50 Case Mix Groups (CMG), top 50 diagnoses and top 50 procedures for all age groups and for all case types in all hospitals in Ontario. The average Ontario overhead cost in 2016/2017 was 22.3% with a range of 15.8 to 35.1%. The hospital-specific overhead cost for SickKids was 31.6%. Based on this information, the reference overhead cost was assumed to be 22.3%, with a range of 15.8% to 31.6%.

Training and start-up costs were not included. These costs are incurred prior to offering the service and can be substantial, depending on the institution. The costs of pre-test and post-test-counselling, and any variant discovery research and development were also excluded.

Table 3. Assumptions: Microcosting Analyses

WGS Test	Description
General	• Costs were applied at the end of each year • Volumes of resource use and prices per unit did not change over 5 years • 22.3% overhead cost was assumed, ranging from 10 to 30% • Overhead cost was applied to labour, small and large equipment, and bioinformatics
Equipment	• Large equipment's useful lifetime was 5 years • Small equipment's useful lifetime was 5 years except thermocyclers and pipette sets which were replaced every 2.5 years • Large and small equipment cost were amortized over 5 years • 1.5% opportunity cost was applied to depreciation of large equipment only • The total capacity in the institution for trios with all indications was a maximum of 3000 cases per year per sequencer
Labour	• Filtration time to analyze ~1500-2000 variants in order to flag 15-20 variants of interest was 90 minutes for primary variants and 10 minutes for secondary variants • Interpretation time for primary variants was 120 minutes to look through 15-20 variants for gene function and classification as pathogenic/benign/VUS. For secondary variants, it was 10-15 minutes to examine as per the pre-classified ACMG criteria. These variants were not always the same as the primary variants • Case review meeting to discuss the trio took 60 minutes (15 minutes per family). The key personnel involved were the Research Associate, Genetic Counsellor, Clinical Geneticist and Cardiologist
Confirmatory testing	• Confirmatory testing was done for both primary and secondary variants. One variant in one gene was sent in and it took 20 minutes per proband. In one year, of the existing sample of 48 families (144 samples), 23.33% were validated • Confirmatory testing included Sanger Sequencing for WGS and Agena MassARRAY® spectrometry/qPCR and WGS for pharmacogenomics
Pharmaco-genomics	• Pharmacogenomics analysis was only completed for probands and consisted of filtration (10-25 minutes for 100 variants), classification (15-20 minutes per proband), case review meeting (15 minutes per proband), confirmatory testing (average of 62.5 minutes per proband) and report writing (30 minutes per proband)
Bioinformatics	• High performance computing cluster maintenance time was 1 hour/ year/node • The maximum number of tests were run each time during batch runs (i.e., a slide that can run 3 cases per test was not used to run a single case) • Each compute node had a warranty of 5 years (3 years with purchase and 2 years of extra purchase of warranty)

Abbreviations: WGS, Whole genome sequencing; qPCR, Real-time polymerase chain reaction; VUS, variant of uncertain significance; ACMG, American College of Medical Genetics.

2.5 Microcosting analysis

Costs per trio were calculated and aggregated by category and by year over the five-year time horizon on WGS (HiSeq X™). Total program costs over five years were also estimated on the same platform. The model was built on Microsoft Excel. Both probabilistic analysis (PA) and deterministic sensitivity analyses (DSAs) were run on R program for statistical computing and graphics (25).

2.5.1 Probabilistic analysis

For each input's resource use and unit price, a range and probability distribution were established in consultation with experts. Probability distributions were defined for inputs which were either proportions or for which upper and lower bound were provided in addition to a point estimate (Table 2). The source for some estimates was often the same expert. Since no evidence existed for any specific form of correlation, all input distributions were assumed to be independent. To propagate variance in the model, 10,000 values were drawn from each input's distribution (i.e. confidence intervals using Monte Carlo replications). Point estimates of inputs with fixed values, i.e. inputs for which ranges were not provided, were repeated 10,000 times.

Most input parameters were described by a truncated normal distribution where a point estimate corresponded to the mean of the normal distribution and lower and upper bounds corresponded to 99.7% confidence interval (i.e. upper and lower bounds were assumed to lie within three standards deviations from the mean):

$$X \sim N(\mu, \sigma^2),$$

where X is a resource use or unit price input, bounded at zero, $0 < X < \infty$, μ corresponds to the point estimate of X, $\sigma = \frac{u-l}{6}$, u is the upper bound and l is the lower bound of a range divided by 6 to make the bounds close to ranges, so most of the data are within the ranges. The 99.7% confidence level was chosen to convey a level of confidence in choosing the upper and lower bounds for an input. The normal distribution was truncated at zero since resource use and prices cannot be negative. This method applied to equipment, supplies and bioinformatics. All labour steps were also modelled using a truncated normal distribution, as stated above, with the exception of the following: the proportion of patients for whom primary and secondary variants were validated (confirmatory testing) for which a beta distribution was used, and clinical interpretation of primary variants, for which a log normal distribution was used.

The resource use for Sanger sequencing was quantified as the proportion of cases in which confirmatory testing was done for both primary and secondary variants (Table 2). At the individual case level, the confirmatory testing can be described by a binomial distribution. In order to represent uncertainty in the proportion of confirmatory tests, the beta distribution, a conjugate to the binomial distribution, was used (26):

$$X \sim Beta(\alpha, \beta),$$

where X is a resource use parameter for confirmatory testing, α is the number of confirmatory tests and β is the total number of tests less the number of confirmatory tests. Since the proportion of confirmatory testing was provided by an expert, that proportion was applied to the total number of tests to obtain the number of confirmatory tests.

The resource use for clinical interpretation was identified as minutes spent by a genome analyst classifying the variants as pathogenic/benign/variant of uncertain significance. On average, a research associate spent two hours per trio, but the time varied between ten minutes to ten hours. Since the distribution was substantially skewed, the log-normal distribution was used:

$$\mathrm{X} \sim \mathrm{Lognormal}\ (\mu, \sigma^2),$$

where X is a resource use parameter for clinical interpretation, μ is estimated by taking a mean of log-transformed minimum (10 minutes) and maximum (600 minutes) resource use values, σ^2 is estimated by $\frac{\log(600)-\mu}{F_X^{-1}(k)}$ where F_X^{-1} is the inverse of the cumulative distribution function and *k* is the percentile, which was set to 98.4 in order for the mean of X to approximate the average clinical interpretation time of 120 minutes.

2.5.2 Sensitivity analysis

An assessment of uncertainty is an essential part of an economic analysis (24, 26, 27). DSA was conducted for selected parameters that were highly uncertain or expected to vary substantially between institutions. For the parameters that were varied, reference level values were repeated 10,000 times. The DSAs permitted an examination of how changing the values of highly uncertain inputs one at a time affected the results.

2.5.2.1 Deterministic sensitivity analysis

Three one-way DSAs were conducted to examine the effects of changing the inputs while other input parameters remained the same: i) the overhead cost; ii) the total volume of tests in the institution; and iii) omission of pharmacogenomics. The reference overhead cost was set at 22.3%. In the DSA, the overhead cost was varied from 10 to 30%. For the WGS-trio test, the reference case total volume of tests in the institution was set at 1500 (equivalent to 500 trios). As the new sequencing technologies are implemented, the volume of referrals for testing is expected to increase. In order to examine how the cost per CGC patient for equipment changes with an increasing number of tests across the institution, the number of WGS-trios was varied from 500 to 1000 (1500 samples to 3000 samples). Omitting the pharmacogenomics component was also tested in a DSA to assess the cost difference between WGS-trio test which includes pharmacogenomics component and the test which does not.

3 Results

3.1 Test costs per patient with a cardiac condition

The results of WGS-trio (HiSeq X™) microcosting analysis are shown in Table 4. The total estimated costs per trio for each year of the five-year program are shown, as well as costs for major cost categories. Figure 3 shows the distribution of the cost per trio by cost category. The results were based on reference values for overhead costs (22.3%) and the number of total tests done per year for all indications (1500).

The total cost per WGS-trio in Year 1 was $8053.10 (95% CI: 7699.30, 8558.10). Supplies made up 50.8% of the total costs whereas bioinformatics accounted for 24.8% of the costs. Labour, overhead, confirmatory testing and large equipment accounted for 11.2%, 8.5%, 2.4% and 2.3%, respectively. Small equipment had a very small contribution to the overall cost with a proportion of 0.037% (Figure 3).

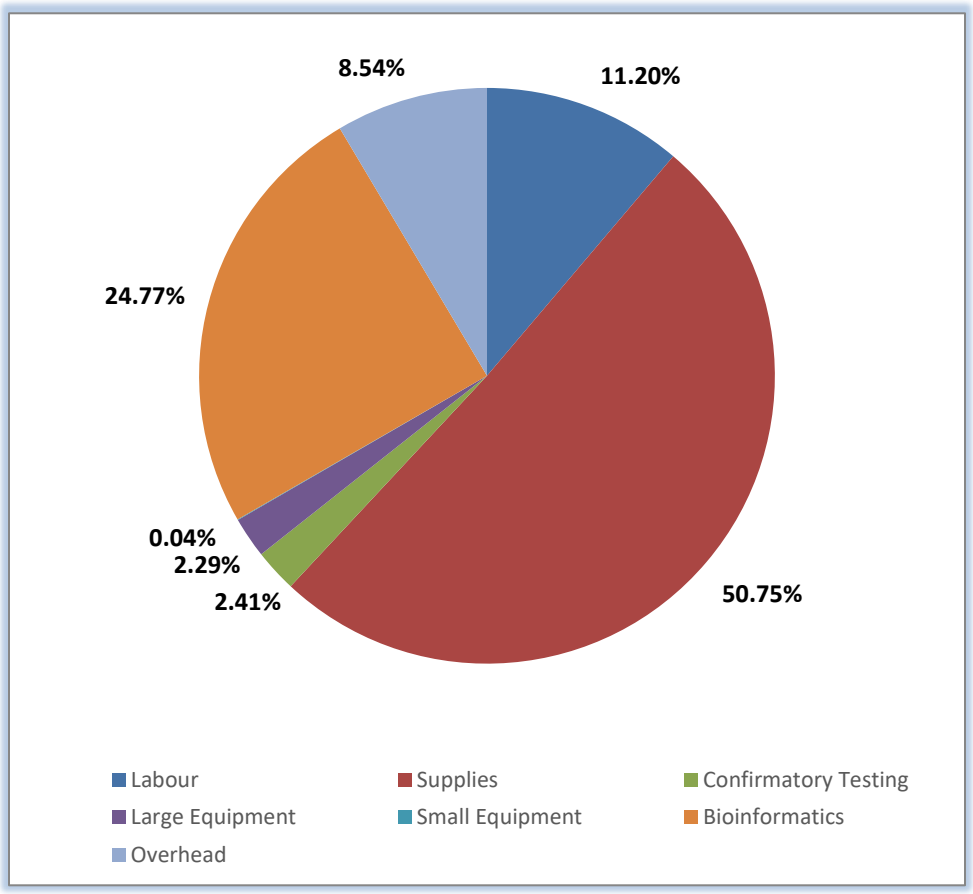
Table 4. Estimated Annual Cost per Cardiac Genome Clinic Trio for Whole Genome Sequencing (Illumina HiSeq X™)

Cost Category	Year 1 (95% CI)	Year 2 (95% CI)	Year 3 (95%CI)	Year 4 (95% CI)	Year 5 (95% CI)
Labour	900.50 (760.20, 1261.20)	887.20 (749.00, 1242.50)	874.10 (737.90, 1224.20)	861.10 (727.00, 1206.10)	848.40 (716.30, 1188.30)
Large Equipment	184.70 (174.00, 195.60)	179.40 (169.00, 190.00)	174.30 (164.10, 184.60)	169.20 (159.40, 179.20)	164.30 (154.70, 174.00)
Small Equipment	3.00 (2.90, 3.10)	2.90 (2.80, 3.00)	2.90 (2.80, 3.00)	2.80 (2.70, 2.90)	2.80 (2.70, 2.90)
Supplies	4088.00 (3840.60, 4333.20)	4027.60 (3783.90, 4269.20)	3968.00 (3727.90, 4206.10)	3909.40 (3672.80, 4143.90)	3851.60 (3618.60, 4082.70)
Confirmatory Testing	194.40 (149.40, 250.00)	191.50 (147.20, 246.30)	188.70 (145.00, 242.70)	185.90 (142.90, 239.10)	183.10 (140.80, 235.60)
Bioinformatics	1995.00 (1900.60, 2091.70)	1965.50 (1872.50, 2060.80)	1936.50 (1844.90, 2030.40)	1907.90 (1817.60, 2000.40)	1879.70 (1790.70, 1970.80)
Overhead	687.50 (646.90, 768.80)	676.80 (636.80, 756.90)	666.30 (626.80, 745.20)	655.90 (617.00, 733.70)	645.60 (607.30, 722.30)
Total	8053.10 (7699.30, 8558.10)	7931.00 (7582.40, 8428.40)	7810.70 (7467.30, 8300.80)	7692.30 (7353.90, 8175.10)	7575.60 (7242.20, 8051.40)

Estimates are given in 2019 Canadian dollars (CAD). Confidence intervals (CI) are based on 10,000 Monte Carlo replications. The results were based on reference levels for overhead costs of 22.3% and 1500 total tests done for all indications per year.

Abbreviations: CGC, Cardiac Genome Clinic; WGS, Whole genome sequencing.

Figure 3. Proportion of Total Annual Cost for Cardiac Genome Clinic Whole Genome Sequencing - trio (Illumina HiSeq X™) by Cost Category, Year 1.

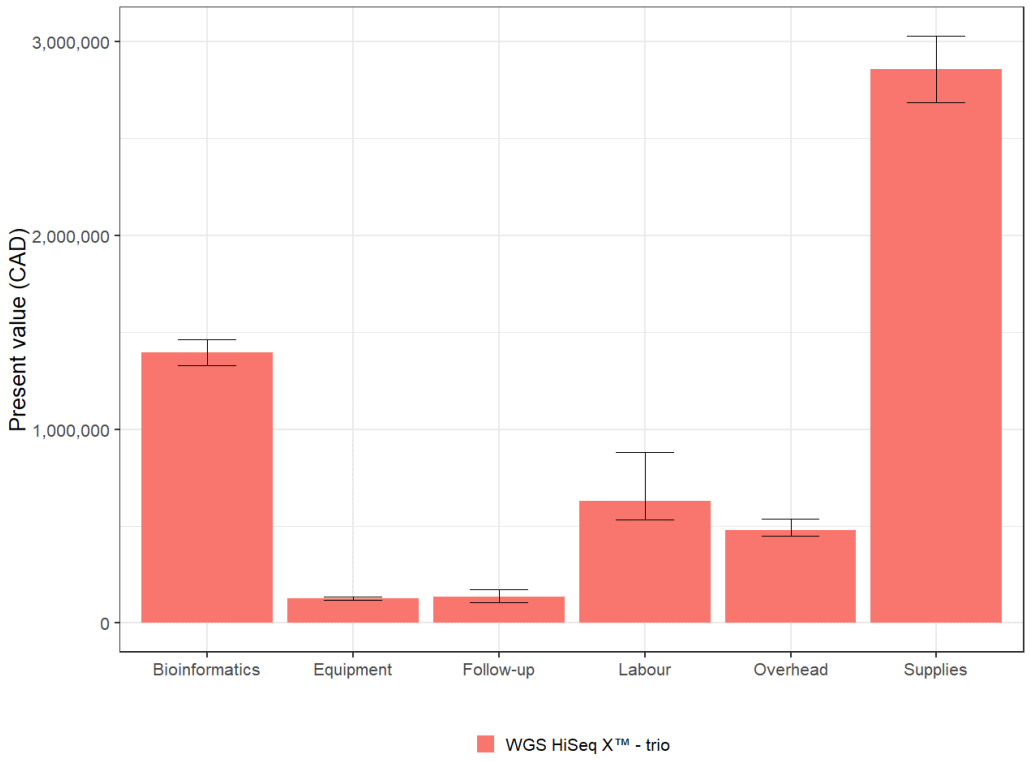


Estimates are given in 2019 Canadian dollars (CAD).
Abbreviations: WGS, Whole genome sequencing.

3.2 Program costs for heterogeneous pediatric cardiac population

The estimated total institutional program cost for the WGS-trio tests over the five-year period (present value) based on 144 CGC cases (48 families) per year on the HiSeq X™ platform was $5.63 million (95% CI: 5.38, 5.98). Figure 4 shows the present value of program costs for each cost component of the trio test. Equipment component includes the cost of both small and large equipment. The program cost of supplies was the largest among the six cost components.

Figure 4. Present Value of Program Costs Over Five Years for Whole Genome Sequencing – trio (HiSeq X™).



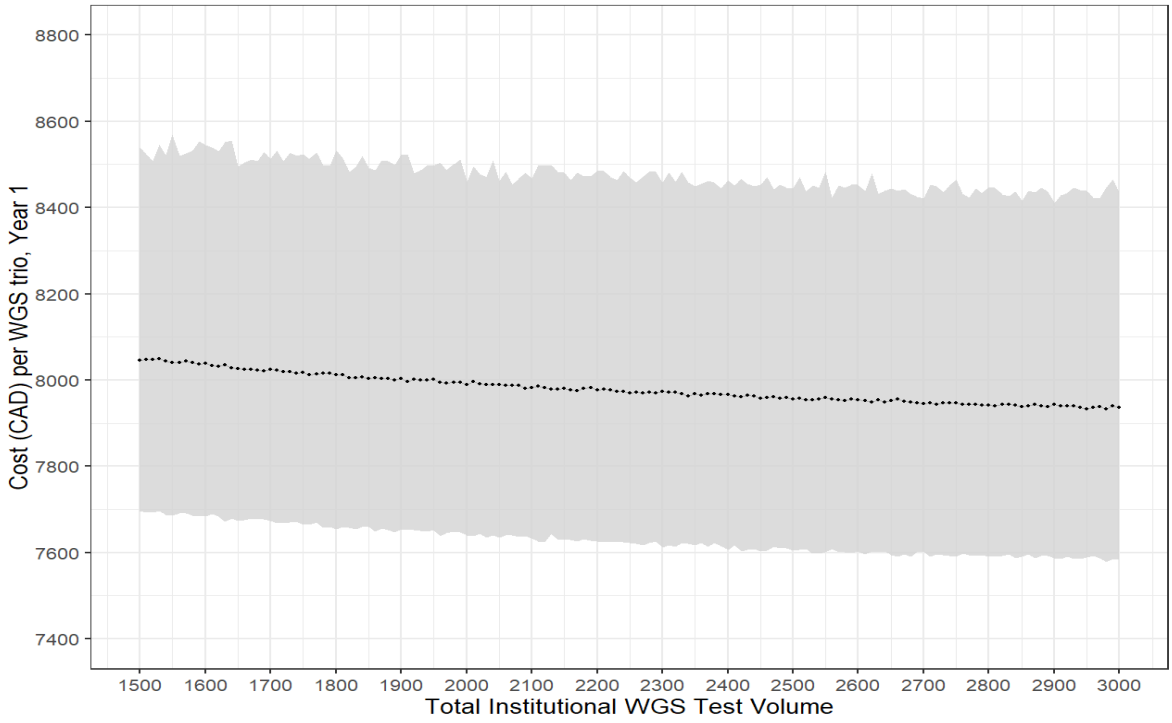
Estimates are given in 2019 Canadian dollars (CAD). Program costs are based on 144 CGC cases (48 families) annually for WGS-trio tests. Confidence bands are based on 10,000 Monte Carlo replications.
Abbreviations: CGC, Cardiac Genome Clinic; WGS, Whole genome sequencing.

3.3 Deterministic sensitivity analysis

Figure 5 shows the effect of increasing the number of annual tests for all indications on WGS-trio costs in Year 1 on the HiSeq X™ platform. Increasing the number of tests for all indications from 1500 to 3000 reduced the costs of WGS-trios by 1.45%. The economies of scale realized was minimal and the savings were attributed to the decrease in the equipment costs (small and large) and corresponding overhead costs. This is because equipment costs accounted for only 2.3% of the total cost, while supplies and bioinformatics costs which increased as the number of samples increased, accounted for most of the total cost.

Table 5 is a summary of the DSA which varied the overhead cost. The results were fairly robust to changes in overhead cost assumptions. Increasing the overhead cost to 30% led to a modest 2.86% increase for WGS-trio (HiSeq X™). Decreasing the overhead cost to 10% led to a 4.71% decrease for WGS-trio (HiSeq X™). Similarly, omitting the pharmacogenomics analysis resulted in a cost savings of $389.20 (4.83% reduction).

Figure 5. Deterministic Sensitivity Analysis of the Effect of Increasing the Number of Whole Genome Sequencing-trio (HiSeq X™) Tests per Year for All Indications from 500 Trios (1500 samples) to 1000 Trios (3000 samples) on Trio Sample Costs in Year 1.



Costs are reported in 2019 CAD. Confidence bands are based on 10,000 Monte Carlo replications.
Abbreviations: CGC, Cardiac Genome Clinic; WGS, Whole genome sequencing.

Table 5. Deterministic Sensitivity Analysis of Estimated Total Cost per Cardiac Genome Clinic Trio for Whole Genome Sequencing, Varying Overhead Cost Proportion.

Overhead cost	Year 1 (95% CI)	Year 2 (95% CI)	Year 3 (95%CI)	Year 4 (95% CI)	Year 5 (95% CI)
WGS, HiSeq X™ - trio					
10%	7673.90 (7336.30, 8142.00)	7557.70 (7225.10, 8018.90)	7443.20 (7115.70, 7897.60)	7330.50 (7007.70, 7778.20)	7219.50 (6901.30, 7660.60)
30%	8290.50 (7926.00, 8821.80)	8164.70 (7805.50, 8688.30)	8040.80 (7686.90, 8556.50)	7918.70 (7570.10, 8426.60)	7798.50 (7455.10, 8298.80)

Estimates are given in 2019 Canadian dollars (CAD). Confidence intervals (CI) are based on 10,000 Monte Carlo replications.
Abbreviations: CGC, Cardiac Genome Clinic; WGS, Whole genome sequencing

4 Discussion

In this study, the trio and program costs of WGS genetic tests for children with cardiac disorders in the CGC were estimated. Primary determinants of costs were supplies and bioinformatics, which accounted for 75% of the total cost. This is due to the greater consumption of costly reagents required for sequencing trios. Similarly, computing demands are much higher for trios. Labour cost was comprised of ten components. Of these components, bioinformatics, clinical interpretation, case review and pharmacogenomics analysis were the most costly.

The present microcosting model incorporated a pharmacogenomics component that provides information on how specific genetic variants may be responsible for metabolism of certain classes of medications. Furthermore, pharmacogenomics provides information to suggest alternate medications for the disease/condition of interest. The pharmacogenomics analysis also looks into additional haplotypes and phenotypes (4). The present model included ten cardiac genes for which there may be medication implications and the cost of this component was $389 per proband. If the list of the genes is expanded, then the cost of pharmacogenomics analysis is likely to increase.

While WGS was the technology of choice, WES may be a cost-effective alternative. Comparing the microcosts of WES for cardiac diseases to the costs of WGS or to other genetic testing strategies, including serial testing options, will provide further evidence to inform funding decisions and efficient implementation. Such an approach was previously undertaken to illustrate the costs involved in genetic testing of ASD (28).

The study has several strengths. All stages and costs involved in the workflow of WGS-trio were accounted for using the microcosting approach generating fully comprehensive per trio and program cost estimates of WGS. Uncertainty associated with parameter estimates was captured in the PA using Monte Carlo simulations. Parameters that were highly uncertain or expected to vary substantially between institutions were varied in DSA demonstrating robustness of the results to changes in assumptions. Predicting costs and volumes of use before a technology has been clinically established presents certain challenges. This study showed how economies of scale can be realized to reduce the trio costs as the volume of total WGS tests increases, in advance of full implementation. The level of this

economic efficiency may be different between a proband and trio sequencing as the trio cost is substantially more expensive compared to the proband cost. The study also showed where cost savings can be realized. Omitting pharmacogenomics analysis is a potential option if it is not expected to be of value for certain indications, if there is limited funding, or if the labour, supplies and the computing demands for this analysis are deemed to be too expensive. The additional cost involved in identifying and confirming secondary variants is a small component of the total cost and therefore unlikely to be a major cost contributor. Identification of secondary variants however has potentially significant downstream effects on use of health resource and the potential for additional health benefits which require further study. Although the estimates in this report are for a pediatric cardiac patient population, the microcosting model was deliberately constructed to be flexible and easily adaptable to other patient populations by changing the resource use items and the volume of testing in the institution. The present model was adapted from a microcosting model for ASD (28, 29).

This study has several limitations. For most of the price parameters, a range of 10% was not based on an expert opinion, but instead chosen to reflect potential price and currency fluctuations. Nevertheless, this range was within the variation for other parameters reported by experts. A five-year time horizon was chosen based on a projected shelf-life for the sequencing equipment, and because procurement decisions for large equipment can be based on a five-year budget plan. In reality, the life cycle for sequencers may be shorter due to rapid evolution of the sequencing hardware and software combined with the frequency of usage. A shorter life cycle would result in higher costs due to a shorter period of amortization. This evaluation did not capture any outcomes such as diagnostic yield or change in clinical management. When measured and captured in future studies, these outcome data would enable a cost-effectiveness analysis (CEA) or cost-consequence analysis (CCA). Another limitation is that this study modelled a trio involving the proband and both biological parents as recommended in clinical practice and does not include estimates for sequencing a duo or proband which may also occur. Furthermore, clinical consultations, genetic counselling and delivery of results to patients and parents were not captured.

Costing of WGS will likely prove to be condition-specific. However, this heterogeneous pediatric cardiac disease population represents a patient group with serious, potentially fatal conditions that impose significant burdens on families and the health care system. Findings from this study may be relevant to

designing future evaluations in other patient populations and in other jurisdictions. Patients and parents who were modeled in this study may not be fully representative of the population as a whole. However, establishing the ability to track health care resource use and costs is an essential first step towards generating more generalizable data as it allows researchers to better understand the resource use implications of genetic and genomic testing and facilitates future comparative economic evaluation to inform future clinical, policy and funding decision-making.

5 Conclusion

WGS could aid researchers, clinicians, patients with CMP, CHD and cardiac arrhythmias and their families in the diagnosis and medical management of these diseases. An economic evaluation of genomic sequencing technologies, such as WGS, requires a comprehensive and accurate estimation of all costs involved in the sequencing workflow. In this study, the total cost per WGS-trio was $8053.10 (95% CI: 7699.30, 8558.10). Supply costs accounted for the largest proportion of costs followed by the costs of bioinformatics. The present value of five-year institutional program was found to be $5.63 million (95% CI: 5.38, 5.98).

This testing strategy is relatively new in the pediatric population, especially in this sub-group of patients. Economic evaluation of WGS is therefore paramount. Based on the DSA conducted in this study, the economies of scale achieved are minimal; however, costs savings can be achieved through the reduction of the costs of reagents and computational requirements.

In future, potential comparative analyses would enable the evaluation of the clinical and personal utility of WGS to patients, families, health care professionals and policy makers in addition to the assessment of differences in cost efficiencies generated. This study provides comprehensive cost data for use in future economic evaluations of clinical GS in various cardiac populations, and allows for a costing model that can be easily adapted to other pediatric patient populations in different health systems.

References

1. Hauser NS, Solomon BD, Vilboux T, Khromykh A, Baveja R, Bodian DL. Experience with genomic sequencing in pediatric patients with congenital cardiac defects in a large community hospital. Mol Genet Genomic Med. 2018;6(2):200-12.
2. Frank M, Prenzler A, Eils R, Graf von der Schulenburg JM. Genome sequencing: a systematic review of health economic evidence. Health Econ Rev. 2013;3:29.
3. Green RC, Berg JS, Grody WW, Kalia SS, Korf BR, Martin CL, et al. ACMG Recommendations for Reporting of Incidental Findings in Clinical Exome and Genome Sequencing. Genet Med. 2013;15(7):565-74.
4. Cohn I, Paton TA, Marshall CR, Basran R, Stavropoulos DJ, Ray PN, et al. Genome sequencing as a platform for pharmacogenetic genotyping: a pediatric cohort study. Npj Genomic Medicine. 2017;2(1):19.
5. Tsiplova K, Zur RM, Marshall CR, Stavropoulos DJ, Pereira SL, Merico D, et al. A microcosting and cost–consequence analysis of clinical genomic testing strategies in autism spectrum disorder. Genet Med. 2017;19(11):1268-75.
6. Health Quality Ontario. Genome-Wide Sequencing for Unexplained Developmental Disabilities or Multiple Congenital Anomalies: A Health Technology Assessment. Ontario Health Technology Assessment Series. 2020;20(11):1-178.
7. Jacoby D, McKenna WJ. Genetics of inherited cardiomyopathy. Eur Heart J. 2012;33(3):296-304.
8. Bowdin SC, Hayeems RZ, Monfared N, Cohn RD, Meyn MS. The SickKids Genome Clinic: developing and evaluating a pediatric model for individualized genomic medicine. Clin Genet. 2016;89(1):10-9.
9. Wright CF, Fitzgerald TW, Jones WD, Clayton S, McRae JF, van Kogelenberg M, et al. Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. Lancet. 2015;385(9975):1305-14.
10. Farwell KD, Shahmirzadi L, El-Khechen D, Powis Z, Chao EC, Tippin Davis B, et al. Enhanced utility of family-centered diagnostic exome sequencing with inheritance model–based analysis: results from 500 unselected families with undiagnosed genetic conditions. Genet Med. 2014;17:578.
11. Caulfield T, Evans J, McGuire A, McCabe C, Bubela T, Cook-Deegan R, et al. Reflections on the cost of "low-cost" whole genome sequencing: framing the health policy debate. PLoS Biol. 2013;11(11):e1001699.
12. Phillips KA, Trosman JR, Kelley RK, Pletcher MJ, Douglas MP, Weldon CB. Genomic sequencing: assessing the health care system, policy, and big-data implications. Health Aff 2014;33(7):1246-53.
13. Wetterstrand KA. DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP)2014.
14. Wright C, Burton H, Hall A, Moorthie S, Pokorska-Bocci A, Sagoo G, et al. Next steps in the sequence. The implications of whole genome sequencing for health in the UK. Cambridge: PHG Foundation; 2011.
15. Jegathisawaran J, Tsiplova K, Hayeems R, Ungar WJ. Determining accurate costs for genomic sequencing technologies—a necessary prerequisite. Journal of Community Genetics. 2020;11(2):235-8.
16. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the Economic Evaluation of Health Care Programmes. 3rd ed. Oxford: Oxford University Press; 2005.
17. Caudle KE, Klein TE, Hoffman JM, Muller DJ, Whirl-Carrillo M, Gong L, et al. Incorporation of Pharmacogenomics into Routine Clinical Practice: the Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline Development Process. Current Drug Metabolism. 2014;15(2):209-17.

18. Whirl-Carrillo M, McDonagh E, Hebert J, Gong L, Sangkuhl K, Thorn C, et al. Pharmacogenomics Knowledge for Personalized Medicine. Clinical Pharmacology & Therapeutics. 2012;92(4):414-7.
19. Kalia SS, Adelman K, Bale SJ, Chung WK, Eng C, Evans JP, et al. Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics. Genet Med. 2017;19(2):249-55.
20. Canadian Institute for Health Information. MIS Standards 2013. Ottawa, ON, Canada: CIHI; 2013.
21. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in Medicine. 2015;17(5):405-23.
22. Ontario Ministry of Health. Ontario Schedule of Benefits, Physician Services Under the Health Insurance Act 2019.
23. Glassdoor 2019 [Available from: https://www.glassdoor.ca/index.htm.
24. Blons H, Rouleau E, Charrier N, Chatellier G, Côté J-F, Pages J-C, et al. Performance and Cost Efficiency of KRAS Mutation Testing for Metastatic Colorectal Cancer in Routine Diagnosis: The MOKAECM Study, a Nationwide Experience. PLoS One. 2013;8(7).
25. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2016.
26. Briggs AS, Mark; Claxton, Karl. Decision modelling for health economic evaluation. New York, United States: Oxford University Press; 2006.
27. Briggs AH, Weinstein MC, Fenwick EAL, Karnon J, Sculpher MJ, Paltiel AD. Model Parameter Estimation and Uncertainty Analysis: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group–6. Medical Decision Making. 2012;32(5):722-32.
28. Jegathisawaran J, Tsiplova K, Marshall CR, Stavropoulos DJ, Pereira S, Thiruvahindrapuram B, et al. A microcosting and cost-consequence analysis of genomic testing strategies (including trios) in austim spectrum disorder: an update Toronto, Canada: The Hospital for Sick Children; 2018 [Available from: http://lab.research.sickkids.ca/task/reports-theses/.
29. Tsiplova K, Zur RM, Ungar WJ, Marshall CR, Stavropoulos DJ, Pereira S, et al. A microcosting and cost-consequence analysis of genomic testing strategies in autism spectrum disorder Toronto, Canada: The Hospital For Sick Children; 2016 [Available from: http://lab.research.sickkids.ca/task/reports-theses/.