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

FULL REPORT

A Micro-costing and Cost Consequence Analysis from a Randomized Controlled Trial Comparing Genome Sequencing to Exome Sequencing for Genetic Diagnosis

Authors

Vercancy Wu, MA
Clinical Research Project Manager, Child Health Evaluative Sciences, The Hospital for Sick Children, Toronto, Canada

Christian R. Marshall, PhD
Director, Genome Diagnostics, Department of Paediatric Laboratory Medicine, The Hospital for Sick Children, Toronto, Canada
Professor, Laboratory Medicine and Pathobiology, University of Toronto, Toronto, ON, Canada

Jackie Hwang, BSc MLT MBA
Senior manager, Genome Diagnostics, Department of Paediatric Laboratory Medicine, The Hospital for Sick Children, 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, ON, Canada

Report No. 2025-01

Date: 2025-04-30

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

Co-investigators

Robin Z. Hayeems ScM, PhD[1,3], Kate Tsiplova, MSc[1], Meredith K. Gillespie HBSc MSc[4,5], Anna Szuto HBSc MSc[6], Caitlin Chisholm HBSc MS[4], Dimitri J. Stavropoulos MSc PhD[2], Viji Venkataramanan MA[1], Bowen Xiao PhD[1], Gregory Costain MD PhD[6], Mélanie Beaulieu Bergeron, PhD[4], Sarah Sawyer MD PhD[4,5], Lynette Lau MSc[2], Lijia Huang PhD[4], Roberto Mendoza-Londono MD MSc[6], Martin J. Somerville MSc PhD[2], Kym M. Boycott MD PhD[4,5], and the GSO Study Team

Institution:

1. Program in Child Health Evaluative Sciences, Hospital for Sick Children Research Institute, Toronto, ON, Canada
2. Division of Genome Diagnostics, Department of Paediatric Laboratory Medicine, Hospital for Sick Children, Toronto, ON, Canada
3. Institute for Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada
4. Department of Genetics, Children's Hospital of Eastern Ontario, Ottawa, ON, Canada
5. CHEO Research Institute, University of Ottawa, Ottawa, ON, Canada
6. Division of Clinical and Metabolic Genetics, Hospital for Sick Children, Toronto, ON, Canada

GSO Study Team

Susan Aucoin, Liz Sinclair Bourque, Tamara Braid, Lucas Bronicki, Meredith Curtis, Sonny Elango, Taila Hartley, Edward J. (Ted) Higginbotham, Emma Hitchcock, Rebekah Jobling, Jennifer Keating, Reem Khan, Whiwon Lee, Michael Mackley, Diana Matviychuk, Olivia Moran, Jean McGowan-Jordan, Anna Pan, Myriam Poirier, E. Magda Price, Kenzie Pulsifer, Sean Simko, Audrey Schaffer, Brian Smith, Venuja Sriretnakumar, Wilson Sung, Syed Hassan Zaidi

Funding and Acknowledgements

Funding was provided by Genome Canada and the Ontario Genomics (OG-186) and the Ontario Ministry of Health. Wendy J. Ungar hold the Canada Research Chair in Economic Evaluation and Technology Assessment in Child Health. Robin Z. Hayeems holds the Canada Research Chair in Genomics and Health Policy. Kym M. Boycott holds the Canada Research Chair in Rare Disease Precision Health. We wish to thank Erin Hsue for assistance with quality control checks.

Ethics Declaration

Research ethics approval was obtained through Clinical Trials Ontario (CTO #3655). CTO is responsible for approving studies involving two or more health care institutions in Ontario. Individuals undergoing sequencing provided clinical consent.

Conflict of Interest Declaration

Wendy J. Ungar chairs the Ontario Genetics Advisory Committee and is a member of the Ontario Health Technology Assessment Committee and does not receive compensation for these roles. She has received grant funding from the Pharmaceutical Research and Manufacturers of America Foundation, honoraria from the Taiwan Center for Drug Evaluation, the Canadian Fertility & Andrology Society and the EuroQol Research Foundation and consulting income from Broadstreet HEOR. Christian R. Marshall received honoraria and travel funds from Illumina. Robin Z. Hayeems is an unpaid member of the Ontario Genetics Advisory Committee. Robin Z. Hayeems, Martin Somerville and Kym Boycott are unpaid members of the Provincial Genetics Advisory Committee. D. James Stavropoulos reported having stocks or stock options in Phenotips. Vercancy Wu, Jackie Hwang, Kate Tsiplova, Meredith K. Gillespie, Anna Szuto, Caitlin Chisholm, Viji Venkataramanan, Bowen Xiao, Gregory Costain, Mélanie Beaulieu Bergeron, Sarah Sawyer, Lynette Lau, Lijia Huang, and Roberto Mendoza-Londono declare no conflicts of interests.

Acronyms and Abbreviations

ASD	Autism spectrum disorder
BoC	Bank of Canada
cap.	Capture
CAD	Canadian dollar
CCA	Cost-consequence analysis
CCMG	Canadian College of Medical Geneticists
CI	Confidence interval
CHEO	Children's Hospital of Eastern Ontario
CMA	Chromosomal microarray analysis
CNV	Copy number variant
Conf.	Confidential
DD	Developmental disabilities
DNA	Deoxyribonucleic acid
DPLM	Department of Pediatric Laboratory Medicine
ES	Exome sequencing
G.C.	Genetic Counsellor
GDL	Genome Diagnostic Laboratory
GS	Genome sequencing
GSO	Genome-wide Sequencing Ontario
GWS	Genome-wide sequencing
HTA	Health technology assessment
Hybe.	Hybridization
ID	Intellectual disabilities
lab.	Laboratory
min	Minute
MIS	Management Information Systems
MoH	Ministry of Health
NA	Not applicable
NYGH	North York General Hospital
PA	Probabilistic analysis
PCR	Polymerase chain reaction
qPCR	Real-time polymerase chain reaction
RCT	Randomized controlled trial
RDs	Rare diseases
SD	Standard deviation
SickKids	The Hospital for Sick Children
SIV	Sample identification verification
SOB	Schedule of Benefits
TASK	Technology Assessment at SickKids
TCAG	The Centre for Applied Genomics
Trun. Norm.	Truncated normal distribution
UHN	University Health Network
USD	United States dollar
VUS	Variant of uncertain significance

Table of Contents

List of Tables

List of Figures

Executive Summary

Background

Diagnosing rare diseases (RDs) is challenging due to their atypical and diverse symptoms, heterogeneity, and genetic complexity. Genome-wide sequencing, consisting of genome sequencing (GS) and exome sequencing (ES), has emerged as a promising strategy for achieving timely diagnosis of RDs, yet it is not currently routinely available as a clinical test across Canada. Subsequent to a positive funding recommendation by Ontario Health for clinical ES for unexplained developmental disabilities and multiple congenital anomalies, funding was received from Genome Canada and the Ontario Ministry of Health to establish the Genome-wide Sequencing Ontario pilot project. This pilot study, co-led by The Hospital for Sick Children (SickKids), Toronto, Canada, and the Children's Hospital of Eastern Ontario (CHEO), Ottawa, Canada, aimed to evaluate laboratory performance, effectiveness costs, and cost-effectiveness of GWS. The pilot project was designed to furnish additional evidence to inform implementation and a future funding decision regarding GS. The present technical report summarizes the economic component of the pilot study.

Objectives

The objectives of this study were to: (1) estimate the precise cost per trio for both GS and ES using a bottom-up micro-costing approach for a targeted patient population consisting of mostly children with suspected rare genetic conditions and their biological parents and (2) using data from a randomized controlled trial, conduct a cost-consequence analysis (CCA) to estimate the incremental cost of trio GS vs. trio ES per unit improvement in molecular diagnostic yield from an institutional payer perspective.

Methods

The study assessed cost per trio for GS and for ES (Illumina NovaSeq 6000) excluding mark-ups, fees, and charges. The estimation was conducted using a bottom-up micro-costing approach based on the laboratory workflow and the volumes for sequencing-related inputs provided by the Department of Paediatric Laboratory Medicine at SickKids. The total cost was decomposed into seven categories, including reagents, consumables, small and large equipment, shipping and ordering, software, labour and overhead. The analysis was conducted from an institutional payer perspective based on the harmonized diagnostic laboratory practices at SickKids and CHEO. The aggregated cost per trio for GS and ES were determined and the total program costs were estimated for each enrollment year. To

address parameter uncertainty in the model, a probabilistic analysis using Monte Carlo simulations was performed. A CCA was conducted to examine the incremental cost and incremental diagnostic yield of GS vs. ES.

Results

In a cohort of 653 families assessed over a two-year period, 324 trios were randomized to GS, and 329 trios were assigned to ES. The total costs per trio for GS and ES were CAD 4364.02 (95% CI 3984.94, 5013.67) and CAD 2888.79 (95% CI 2567.72, 3492.72) respectively. Reagents were the primary cost component for both strategies, accounting for 61% of the total expenditure for GS and 34% for ES. Software and labour were identified as the second and third highest cost components for GS (15% and 14%, respectively). In contrast, labour and consumables ranked as the second and third most substantial cost components for ES (22% and 18%, respectively). The incremental cost for GS compared to ES was CAD 1475.23, and the diagnostic yields for GS and ES were 32.72% and 35.87%, respectively. The difference between ES and GS diagnostic yield was 0.032 (95% CI: -0.041, 0.104, p-value 0.397).

Conclusions

This study furnished evidence of the cost and cost-effectiveness of trio GS vs. ES using a bottom-up micro-costing approach. GS was associated with higher costs and a similar diagnostic yield for this randomized population with RDs, based on the technical capabilities for sequencing current at the time of the study. The study provides comprehensive costs for future economic evaluations of alternative diagnostic pathways to inform future funding and implementation decisions and impetus for further evaluating variants uniquely detectable by GS.

1. Introduction

1.1 Background

A rare disease (RD) denotes a medical condition that affects a limited number of individuals in a given population (1). These diseases often arise from genetic mutations or variations that are uncommon in the general population (2). Genetic testing plays a pivotal role in diagnosing RDs by analyzing an individual's Deoxyribonucleic acid (DNA) to identify mutations or variations, thereby providing crucial insights into the condition's underlying cause (3). Accurate and early diagnosis through genetic testing is transformative for affected individuals and families, enabling informed treatment decisions, early access to targeted therapies, and enhanced family planning (4). Ongoing advancements in genomic technologies and collaborative efforts are driving progress in understanding and managing RDs, elevating the potential for a precision medicine approach.

The advent of genome-wide sequencing provides insight into the genetic underpinnings of various conditions. Amongst the methods employed, two prominent approaches are genome sequencing (GS) and exome sequencing (ES). The exome consists of less than 2% of a person's entire genome, but it contains approximately 85% of known disease-related variants (5). ES focuses exclusively on protein-coding regions, as mutations in these regions are most likely to be related to observed phenotypes. In contrast, GS detects both small and large *de novo* mutations, as well as inherited variations in coding and noncoding regions of DNA, including copy number variants (CNVs), small nucleotide variants and structural variations (6, 7). Additionally, novel, causative mutations of rare or common Mendelian disorders have been identified through the use of these technologies (8). Both GS and ES can generate findings unrelated to the purpose of the test, i.e., secondary findings, that may predict risk for other conditions and have a significant impact on a patient's health (9). GS can also identify pharmacogenomic variants associated with medication metabolism or sensitivities (10). This comprehensive method is proving instrumental in unravelling complex diseases and uncovering genetic diversity.

Genomic testing began to etch a transformative trajectory within the healthcare landscape of Ontario beginning with chromosomal microarray analysis (CMA) in the mid-2000s. CMA identifies large structural chromosomal changes and can lead to a diagnosis in 8–15% of people tested, and has played a vital role in diagnosing conditions like autism spectrum disorder (ASD) (11). ES gained prominence in the late 2000s, led by The Centre for Applied Genomics (TCAG) at SickKids, for diagnosing rare genetic

disorders in protein-coding regions. Initiatives such as the SickKids Genome Clinic and the Rare Disease Network harnessed ES potential. The 2010s witnessed the rise of GS, supported by the Ontario Institute for Cancer Research and Genomics Ontario, exemplified by the Ontario Health Study, which aimed to collect population-level health and genetic data to elucidate gene-environment interactions and their role in chronic disease epidemiology (12, 13). Additionally, the Integrated Clinical Genomic Information System in partnership with the McLaughlin Centre (14) was established with the purpose of integrating clinical genomics into healthcare to advance precision oncology and personalized medicine. These programs enabled patients to access sequencing if they were enrolled in a research study. Clinical ES was initially available to Ontario patients only through an out-of-country special authorization mechanism with individual requests from specialists requiring approval by the Ontario Ministry of Health (MoH). Technology Assessment at SickKids (TASK) (15) initiated the first micro-costing model to compare the cost-effectiveness of CMA, GS and ES in pediatric patients with developmental disorders (16, 17). The models were updated in 2018 to include trio-based sequencing (proband and biological parents) (18) and from 2019 to 2022, genomic sequencing micro-costing was expanded to pediatric cardiomyopathy patients (19, 20).

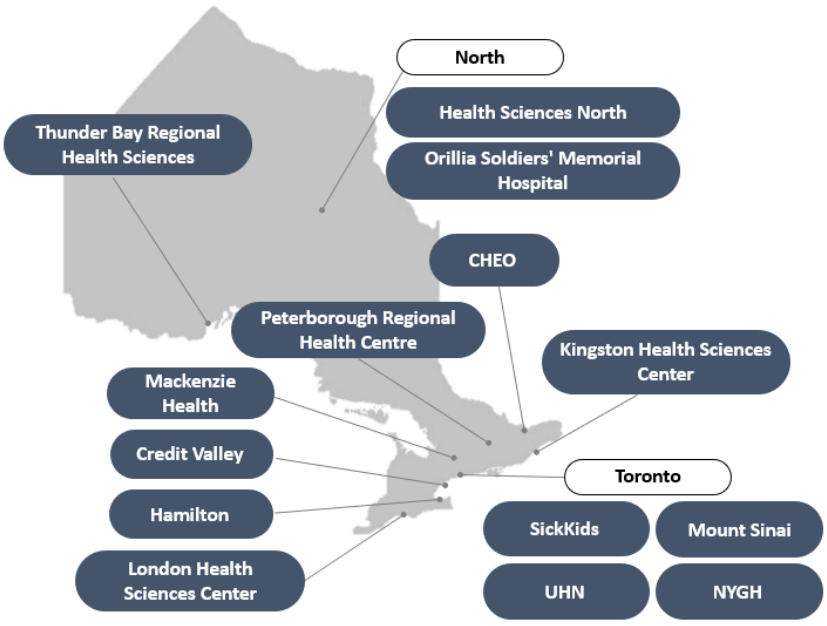
In 2019, in response to increasing demand that clinical genome-wide sequencing be made available as a provincially funded service, a comprehensive health technology assessment (HTA) was undertaken by Ontario Health to examine genome-wide sequencing strategies for unexplained developmental disabilities (DD) and multiple congenital anomalies. The HTA investigated clinical effectiveness, safety, cost-effectiveness, budget impact, patient and provider preferences, and ethical considerations (21). The cost-effectiveness analysis considered the value of GS or ES compared to standard care when positioned at various tiers in the diagnostic assessment pathway (22). After reviewing the evidence, a positive funding recommendation was made in 2020 to fund clinical ES for this patient population as a second-tier test following a negative or inconclusive CMA. As evidence for GS diagnostic yield and patient outcomes was less than that of ES at the time of the HTA, and given the evolving nature of the technology, the funding recommendation acknowledged that further evidence would be needed to revisit the funding recommendation for GS. The HTA report also acknowledged the significant implementation issues related to generating capacity in laboratories, medical genetics, genetic counselling, and other professions, as well as developing and offering an efficient sequencing service for the population of Ontario. This led to the development in 2021 of Genome-wide Sequencing Ontario (GSO), a two-year pilot project co-funded by Genome Canada and the Ontario MoH to establish and

offer ES as a clinical service to Ontario patients with RDs and to collect additional data regarding effectiveness and implementation indicators (23).

1.2 Genome-wide Sequencing Ontario

GSO is a strategic clinical collaboration between SickKids and the Children's Hospital of Eastern Ontario (CHEO) focusing on RD diagnostics (24). Under a hub-and-spoke model, both SickKids and CHEO served as analysis hubs, with SickKids solely undertaking DNA sequencing on all samples collected in Ontario. Eligible patients residing in Ontario who met the genome-wide sequencing (GWS) inclusion criteria established by the MoH and the Canadian College of Medical Geneticists (CCMG) were enrolled in GSO from 18 sites, including medical centers and clinics specialized in RDs across Ontario. Samples were collected from eligible patients and sent from local clinics or specimen collection centers within the CHEO and SickKids communities based on their geographic locations (Figure 1). The CHEO catchment area included northern Ontario, Peterborough, Ottawa, Kingston, Durham, London, and Windsor, while SickKids covered western and other locations in the province. Testing volumes were distributed between CHEO and SickKids in a 40/60 split.

Figure 1. GSO patient recruitment map



Abbreviations: CHEO, Children's Hospital of Eastern Ontario; UHN, University Health Network; NYGH, North York General Hospital

Through the pilot study, GSO built capacity for clinical-grade GWS, specifically targeting patients with suspected rare genetic conditions. The success of the pilot study led to sustained MoH funding for clinical ES. GSO's ongoing work continues to inform the MoH's decisions regarding the evolution and establishment of new indications for clinical GWS in Ontario.

1.3 Rationale

The majority of RDs affect children and most of these diseases are caused by underlying genetic conditions (25). Although individually rare, the more than 7,000 rare genetic diseases are collectively common, affecting 1 in 12, or approximately 3 million Canadians (26, 27). RDs have devastating impacts on health and wellbeing, and more than one-third of affected families lack a genetic molecular diagnosis despite extensive investigation (25). These families often spend more than five years on an odyssey of specialist visits and invasive testing in search of a diagnosis (28). This "diagnostic odyssey" is lengthy, costly, and often futile (29). ES and GS have the potential to significantly reduce the time to diagnosis and improve diagnostic yield, thereby enabling patients to access available treatments earlier.

A systematic review and meta-analysis of 37 studies reported molecular diagnosis rates ranging from 24-68% and *de novo* variant diagnosis rates between 18% to 70% using whole genome sequencing or whole exome sequencing in children with rare genetic diseases (30, 31). Amongst 37 studies, only one was a randomized controlled trial (RCT) comparing rapid whole genome sequencing to standard genetic testing in inpatient trios (infants aged <4 months in a regional neonatal intensive care unit and pediatric intensive care unit with illness of unknown etiology [NSIGHT-2 trial]) (32). That study was limited to acutely ill infants and did not consider cost-effectiveness.

Despite pseudoeconomic claims regarding "falling costs," both GS and ES continue to be costly testing options. A systematic review of economic evaluations reported cost estimates for GS in cancer or RDs that ranged between United States dollar (USD) 2094 (Canadian dollar [CAD] 2722) and USD 9706 (CAD 12,618), and for ES that ranged between USD 716 (CAD 931) and USD 4817 (CAD 6262) per patient (33). Previous micro-costing and cost-consequence analysis (CCA) of genomic testing strategies in ASD and DD populations (16-18, 34), as well as micro-costing of GS in a heterogeneous cardiac population (19, 20) from our group provided early evidence regarding cost-effectiveness of next generation sequencing technologies for the province of Ontario. These technologies continue to evolve rapidly and sequencing

workflows have improved upon previous procedures (16, 18, 20) with regard to operating platforms, use of software, use of cloud-based resources, and labour components. Due to the lack of high-quality evidence of clinical effectiveness from RCTs and the large variation in cost estimates, the value for money of GS in comparison to ES remains uncertain in RD diagnostics. Such evidence can be used to inform future funding recommendations for Ontario and other jurisdictions.

1.4 Study objectives

The objectives of this study were to:

1. Estimate the precise cost per trio for both GS and ES using a bottom-up micro-costing approach for a targeted patient population with suspected rare genetic conditions and their biological parents.
2. Using data from an RCT, conduct a CCA to estimate the incremental cost of trio GS vs. trio ES per unit improvement in molecular diagnostic yield from an institutional payer perspective.

2. Methods

In this section, the study design is presented followed by patient eligibility and enrollment, and delineation of the sample and data collection flow. The sequencing workflow is subsequently outlined, emphasizing the methodology and workflow stages involved in both genome and exome DNA sequencing procedures. The methods for costing, including cost item identification, measurement, and valuation are presented. Methods used to cost GS and ES are described within the major categories including labour, supplies (reagents and consumables), equipment, shipping and ordering, and informatics as demonstrated in Table 1. Assumptions required to conduct the analysis are listed. The cost analysis, probabilistic analysis (PA) and CCA are then described.

2.1 Study design

As a part of the GSO implementation project, a micro-costing and CCA of GS compared to ES were conducted to establish the costs and to evaluate the cost-effectiveness of two sequencing strategies for Ontario patients with RDs. This study was embedded in an RCT in patients with suspected rare genetic conditions. An unblinded, stratified, permuted block randomization was applied after specimen collection to randomize patients to GS or ES. Due to the wide heterogeneity of patient phenotypes and

clinical characteristics, randomization was stratified according to the following criteria to ensure balance in those variables that might confound cost and/or diagnostic yield:

1. Phenotype categories: syndromic Intellectual disabilities (ID)/developmental disabilities (DD), multiple congenital anomalies without ID/DD, multisystem disorder without ID/DD, single system disorder without ID/DD and isolated ID/DD;
2. Prior testing categories: none, CMA, gene panel testing and unknown;
3. Clinical sites: SickKids and CHEO

Patients randomized into each of the above categories were assigned to receive GS or ES at a 1:1 ratio. A block size was either 2, 4, or 6 depending on the expected recruitment in each stratum, resulting in a total of 60 blocks comprising all cases.

2.2 Patient eligibility and recruitment

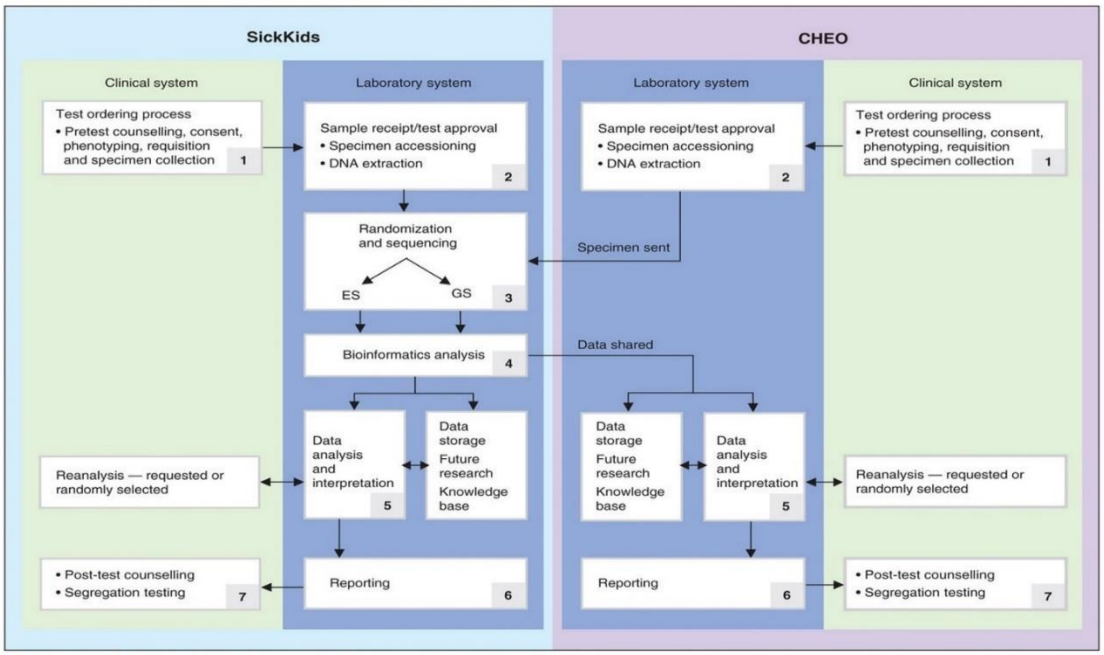
Patients eligible for GSO and this RCT resided in Ontario and met the eligibility criteria for GWS as established by the MoH and the CCMG position statement (35). The enrolled patient must have had a comprehensive baseline genetics evaluation, including family history assessment, phenotyping diagnosis, pretest genetic counselling and consent, and, where indicated, chromosome microarray and targeted genetic testing, including biochemical testing; a genetic etiology as the most likely explanation for the phenotype; and a clinical presentation that included any one of the following: moderate to severe developmental or functional impairment, multisystem involvement, a progressive clinical course that could not be explained by another cause or a by a differential diagnosis that included 2 or more conditions that would require evaluation by separate gene panels; and blood samples available from the index patient and both parents (i.e., trios).

The target population included adults and pediatric patients under 18 years of age, and their biological parents (trios). During the ordering process, prospective trios were recruited and evaluated eligibility from 14 medical genetics centers across Ontario illustrated in Figure 1. Blood specimen samples were collected for eligible trios and shipped to laboratories at SickKids. After pre-sequencing preparation, samples were then randomized to undergo either GS or ES. Randomization was carried out by the SickKids laboratory after sample collection. Cases with uncertain eligibility were reviewed by a joint committee of clinical and laboratory experts. For cases deemed ineligible, physicians could choose to appeal to the joint committee for re-evaluation.

2.3 Sample and data flow

Figure 2 illustrates the harmonized, multi-institutional model established to deliver GS and ES. In the test ordering process, venipuncture was completed at medical clinics to get blood drawn from eligible patients who had completed pretest counselling, consent, phenotyping and requisition. Based on geographic location, samples were then delivered to laboratories at CHEO or SickKids for specimen accessioning and DNA extraction. Subsequently, in the analytic phase, all specimens were shipped to the Genome Diagnostic Laboratory (GDL) at SickKids and underwent library preparation in anticipation of sequencing. Samples were randomly assigned to undergo either GS or ES. Once sequencing was completed, the output underwent informatics analysis and results of individual cases were shared between SickKids and CHEO. In the post-analytical phase, a genome analyst or clinical geneticist analyzed and interpreted the results, which were then conveyed to the genetics practitioner. Cases with unknown or complex findings were discussed by panel experts in a case conference.

Figure 2. Genome-wide sequencing Ontario harmonized, multi-institutional model for delivering genome-wide sequencing



Abbreviations: CHEO, Children's Hospital of Eastern Ontario; DNA, Deoxyribonucleic acid; ES, exome sequencing; GS genome sequencing; SickKids, Hospital for Sick Children

2.4 Sequencing workflow

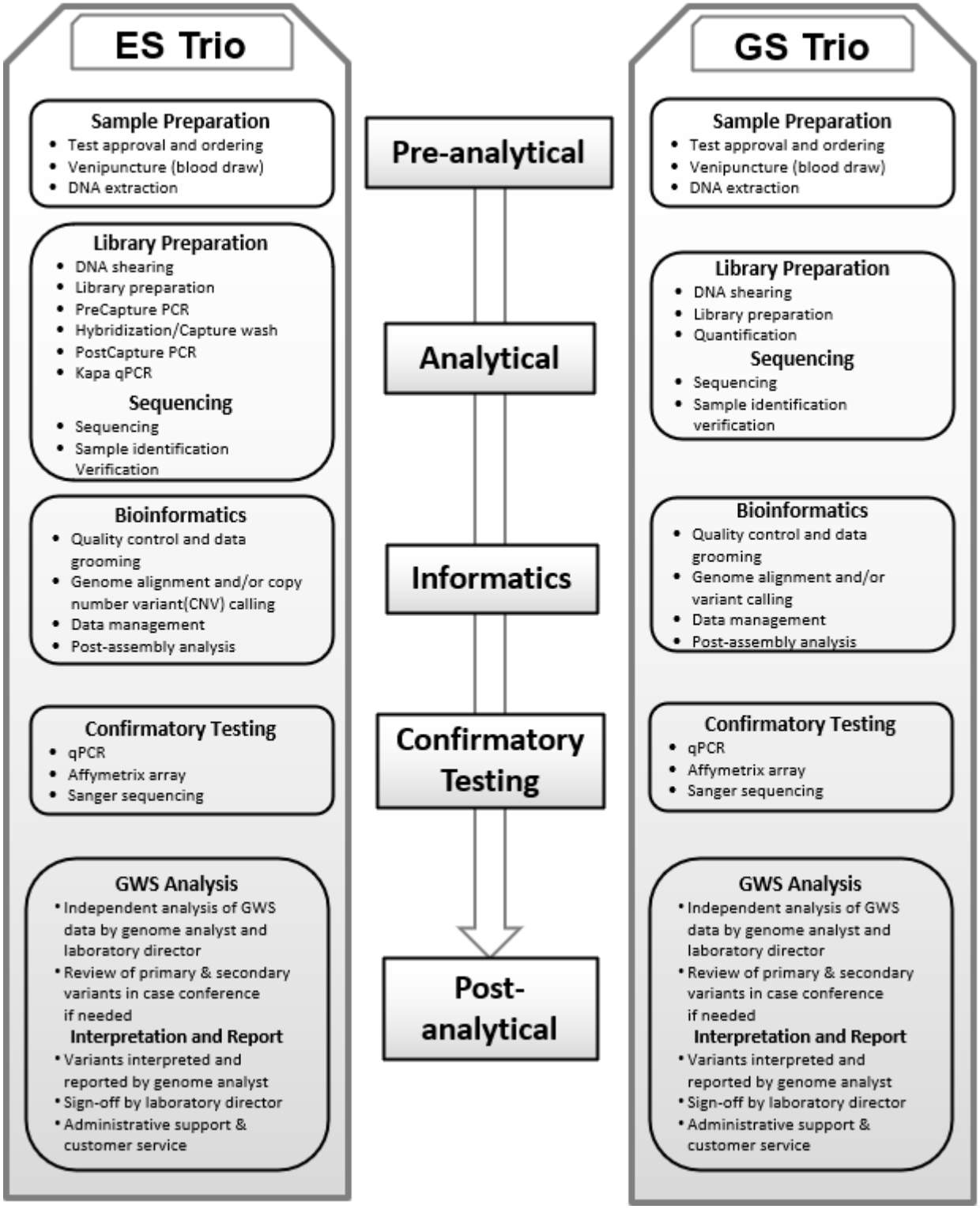
Both GS and ES technologies were implemented by GDL within the Department of Pediatric Laboratory Medicine (DPLM) at SickKids. Sequencing consisted of multiple workflow stages; each stage contained

various work steps as demonstrated in Figure 3. The sequencing platform used in this study (Illumina NovaSeq 6000) has significantly improved efficiency compared to previous platforms and can be used for both GS and ES. Figure 3 demonstrates how the workflow stages were queued in the technical pathway from sample preparation to the clinical interpretation of for both GS and ES. The workflow consisted of five major stages for both GS and ES: pre-analytical phase; analytical phase; informatics; confirmatory testing and post-analytical phase. Within each workflow stage, there were multiple workflow steps as indicated in Figure 3.

Costing data were collected for each workflow step by the GDL at SickKids. Clinical data were collected and audited by a data audit team on a monthly basis using a centralized REDCap database. There were eight data collection instruments designed to capture the entire data collection process, including patient characteristics, review process, sample accessioning, randomization, findings, unconfirmed findings, director's form, and segregation analysis. In our previous studies (16, 18, 20), labour time for analyzing sequencing data was not captured. In this study, the analysis time spent by the genome analyst and laboratory director was collected prospectively for each case to evaluate their contribution to the overall working efficiency and cost. Additionally, this approach allowed for a more comprehensive understanding of the labour resources required at different steps of the sequencing process.

After completion of the data collection, cases were evaluated by experts including clinical geneticists and laboratory directors, and findings were categorized as positively diagnostic, partially diagnostic, potentially diagnostic, no diagnostic findings, and non-diagnostic. Cases with complex interpretation underwent case conferences with a panel that included a clinical geneticist, a genetic analyst and a laboratory director.

Figure 3. Workflow of genome and exome sequencing using NovaSeq 6000 sequencing system



Abbreviations: ES, exome sequencing; DNA, Deoxyribonucleic acid; GS, genome sequencing; GWS, genome-wide sequencing; PCR, polymerase chain reaction; qPCR, quantitative polymerase chain reaction.

2.5 Costing

A bottom-up micro-costing approach was used to estimate the cost per trio (proband plus biologic parents) for each component at each workflow stage for GS and ES. Costs, excluding mark-ups and charges, were estimated from an institutional payer perspective based on the diagnostic laboratory practices at SickKids. The costs incurred in each workflow stage were calculated for all items used within the laboratory for each sample sequenced. This includes costs of labour, reagents, consumables, small and large equipment, shipping and ordering, software and overhead as listed in Table 1. As some sequencing workflow steps were run in batches, micro-costs per sample were adjusted to costs per trio when appropriate. The detailed costing model was constructed in Microsoft Excel (version 2023). Data were then exported to the statistical software package R (Version 4.3.1) for a PA. Methods for item identification, measurement, and valuation are explained below.

2.5.1 Micro-cost item identification

Major cost categories for both strategies were labour, reagents, consumables, equipment, shipping and test ordering, software and overhead. A list of major and minor categories for cost analysis in the following sections is presented in Table 1. Each minor category represented a distinct workflow step. Within each workflow step, the cost per sample per item was calculated based on the resource utilization and unit price associated with each item. Subsequently, the cost per sample per workflow stage was calculated by aggregating the cost of all items within that particular workflow step.

The workflows for GS and ES were similar. However, additional steps were involved during library preparation for ES as it required more complex procedures to achieve targeted DNA extraction from coding regions.

A labour cost was applied for sample preparation, DNA extraction, library preparation, sequencing, confirmatory testing, informatics, clinical interpretation, case conference and administrative tasks.

Costs of supplies were divided into consumables and reagents. Consumables refer to expendable single-use materials (e.g., pipette tips). Consumables and reagents were consumed during sample preparation, library preparation, sequencing and confirmatory testing. A 5% contingency for failed runs or the need to top up cases to get more coverage was embedded into supply costs for ES and GS. Two types of confirmatory testing were used for both GS and ES: the quantitative polymerase chain reaction

(qPCR)/Affymetrix array and Sanger sequencing. The qPCR/Affymetrix array was used for confirming the presence of DNA deletions (a segment of DNA is missing or deleted from a specific region of a chromosome) or duplications (the presence of extra copies of a particular DNA segment within a chromosome, resulting in an increased number of that segment). Sanger sequencing was used to determine the precise order of nucleotide bases within a DNA molecule and to verify the presence of specific genetic variants suspected to be associated with a particular disorder.

Informatics is a rapidly maturing aspect of sequencing and is relevant to data management, storage, informatics analysis, variant annotation, and interpretation. Sequencing output data were analyzed using Dragen and Emedgene software Illumina, San Diego, US) which streamlined the analysis and annotation of genetic variants to improve efficiency and accuracy in interpretation. Some of the workflow steps involved using software or databases stored on servers in the United States, such as Illumina ICA active storage and Isilon storage. Informatics was included in various analysis steps in the analytical and post-analytical stages.

Small and large equipment items were identified as machinery or tools utilized within the sequencing workflow. Small equipment typically encompassed instruments and tools, such as metal racks, which were used for sample preparation and library preparation. Large equipment usually referred to sequencing platforms used for preparing or conducting the actual sequencing of the DNA samples and generating the raw data for downstream analysis and included the NovaSeq6000 (Illumina), TapeStation (Agilent, Santa Clara, US), Bravo Robot (Agilent), Quantstudio (Thermo Fisher, Waltham, US) and ultrasonicator (Covaris, Woburn, US). Each large equipment machine required maintenance on a yearly basis covered by a service contract. Service contracts were for one year for all large equipment except the Bravo Robot, which was serviced under a three-year contract. Unlike consumables or reagents, equipment was reusable during genome and ES. They were relevant for multiple stages including pre-analytical, analytical, informatics and confirmatory testing.

Table 1. Categories of resource use for GS and ES per trio

Major Category	**Minor Category** *(GS and ES)*
Labour	Sample preparation
	Library preparation
	Sequencing
	Informatics
	Confirmatory testing
	Case conference
	Administrative
Consumables	Sample preparation
	Library preparation
	Sequencing
	Confirmatory testing
Reagents	Sample preparation
	Library preparation
	Sequencing
	Confirmatory testing
Equipment and Service Contracts	Large equipment
	Small equipment
	Service contract
Shipping & Ordering	Specimen receipt
	Specimen accessioning
	Test approval and ordering
Informatics	Analysis software and storage
	Informatics common equipment
	Informatics software licensing
Overhead	Labour
	Equipment
	Software

Abbreviations: ES, exome sequencing; GS, genome sequencing.

2.5.2 Micro-cost item measurement and valuation

The volume of use and unit price of each item are described below by major category for GS and ES. All cost items were micro-costed during the study using either 2022 or 2023 CAD. Because most large equipment were purchased prior to 2022, a 2023 expected value for the cost of each piece of large equipment and associated contract were determined based on the shelf-life/term and the year of purchase. When unit prices were in USD, they were converted to CAD using a currency exchange rate of 1.3 USD/CAD (June 13, 2022) from the Bank of Canada (BoC) (36). Since multiple sub-stages (workflow steps) were involved in each workflow stage, a cost per trio was first determined for each workflow step as described below. The overall cost per trio was then calculated by summing the costs across all

workflow stages. The quantity of resource use and unit price data for each input was provided by laboratory staff at DPLM, manufacturers or extracted from public sources such as the 2022 Ontario Schedule of Benefits (SOB) (37). In addition to obtaining resource use and unit price point estimates, ranges comprising reasonable variation for each input's volume and unit price were generated. Table 2 provides the resource volume for each item used in the GS and ES workflows. Instead of a point estimate, each parameter was assigned a distribution and a range to capture uncertainty. For each item used in both GS and ES, a price, along with its distribution and range, was determined as listed in Table 3.

Table 2. Resource volumes of use point estimates, distributions, and ranges for GS and ES cost items

Items	Volume of Use per Batch (min)		
	Point Estimate	**Distribution**	**Range**
Labour			
Sample Preparation (Unit: min)			
DNA Extraction method 1 **(95%)** QIAsymphony platform	90	Fixed	NA
DNA Extraction method 2 **(5%)**			
Puregene	30	Fixed	NA
QIAamp DNA Mini kit	35	Fixed	NA
Specimen receipt	1.65	Fixed	NA
Specimen accessioning	3.85	Fixed	NA
Test approval and ordering	60	Log Normal ~ (4.055, 0.273)	[20, 120]
Library Preparation (Unit: min) - GS			
Covaris	60	Fixed	NA
TapeStation	15	Fixed	NA
Build worksheets/pull samples	30	Fixed	NA
Library preparation	60	Fixed	NA
Master mix preparation	30	Fixed	NA
Kapa preparation	50	Fixed	NA
HSD1000	20	Fixed	NA
Library Preparation (Unit: min) - ES			
Covaris	90	Fixed	NA
Build worksheets/pull samples	30	Fixed	NA
Library prep Bravo	45	Fixed	NA
Pre-cap PCR	30	Fixed	NA
Pre-cap ampure	20	Fixed	NA
D1000 TapeStation	30	Fixed	NA
Hybe.- cap. + wash	120	Fixed	NA
Post-cap PCR	40	Fixed	NA
Final ampure	20	Fixed	NA
HSD1000	20	Fixed	NA

Kapa prep	50	Fixed	NA
Sequencing (Unit: min)			
Calculations	60	Fixed	NA
Pooling	45	Fixed	NA
Dilution/PCR set up	40	Fixed	NA
Extension PCR	20	Fixed	NA
Analysis	30	Fixed	NA
Technical supervision	20	Fixed	NA
Informatics (Unit: min)			
Senior bioinformatician	15	Log Normal ~ [2.70, 0.1342]	[1, 20]
Confirmatory testing (Unit: min) - GS			
Deletion			
PCR setup	(50%) 35	Beta, α=3.26, β=322.74	NA
Analysis	(50%) 15	Beta, α=3.26, β=322.74	NA
Duplication			
DNA plate preparation to PCR steps	(50%) 900	Beta, α=3.26, β=322.74	NA
Purification	(50%) 120	Beta, α=3.26, β=322.74	NA
Confirmatory testing (Unit: min) - ES			
Deletion			
PCR setup	(50%) 35	Beta, α=13.06, β=639.94	NA
Analysis	(50%) 15	Beta, α=13.06, β=639.94	NA
Duplication			
DNA plate preparation to PCR steps	(50%) 900	Beta, α=13.06, β=639.94	NA
Purification	(50%) 120	Beta, α=13.06, β=639.94	NA
Sanger Sequencing			
PCR setup, electrophoresis, sequencing set up, sequencing clean up	30	Beta, α=32.65, β=620.35	NA
Analysis (first and second check)	10	Beta, α=32.65, β=620.35	NA
Primer design	45	Beta, α=32.65, β=620.35	NA
Confirmatory testing (Unit: proportion of patients) - GS			
Deletion	0.01	Beta, α=3.26, β=322.74	NA
Duplication	0.01	Beta, α=3.26, β=322.74	NA

Sanger sequencing	0.05	Beta, α=32.65, β=620.35	NA
Confirmatory testing (Unit: proportion of patients) - ES			
Deletion	0.02	Beta, α=13.06, β=639.94	NA
Duplication	0.02	Beta, α=13.06, β=639.95	NA
Sanger sequencing	0.05	Beta, α=32.65, β=620.35	NA
Interpretation & reporting (Unit: min) - GS			
Genome analysis	201	Log Normal ~ [5.19, 0.49]	[50, 780]
Laboratory director (Medical/Scientific review)	88	Log Normal ~ [4.31, 0.51]	[29, 360]
Customer service	5	Fixed	NA
Administrative support	15	Fixed	NA
Case conference			
Genome analysis	10	Log. Normal ~ [1.99, 0.19]	[5, 10]
Clinical geneticist	10	Fixed	NA
Laboratory director	10	Log. Normal ~ [1.99, 0.19]	[5, 10]
Interpretation & reporting (Unit: min) - ES			
Genome analysis	193	Log Normal ~ [5.13, 0.50]	[50, 720]
Laboratory director (Medical/scientific review)	81	Log Normal ~ [4.29, 0.47]	[30, 330]
Customer service	5	Fixed	NA
Administrative support	15	Fixed	NA
Case conference			
Genome analysis	10	Log. Normal ~ [1.99, 0.19]	[5, 10]
Clinical geneticist	10	Fixed	NA
Laboratory director	10	Log. Normal ~ [1.99, 0.19]	[5, 10]
Supplies (units: uses per trio)			
Reagents			
Illumina NovaSeq S1/S4 reagent kits	3	Fixed	NA
Confirmatory testing - deletion reagents	1	Fixed	NA
Confirmatory testing - duplication reagents	1	Fixed	NA
Other reagents	3	Fixed	NA
Consumables			
Confirmatory testing - deletion consumables	1	Fixed	NA
Confirmatory testing - duplication consumables	1	Fixed	NA

Other consumables (e.g., plates, TapeStation consumables)	3	Fixed	NA
Equipment (Units: uses per trio)			
Large equipment			
BioAnalyzer/TapeStation	3	Fixed	NA
BioAnalyzer/TapeStation service contract	3	Fixed	NA
Bravo robot	3	Fixed	NA
Bravo robot - service contract	3	Fixed	NA
Covaris ultrasonicator	3	Fixed	NA
Covaris ultrasomicator service contract	3	Fixed	NA
Quantstudio (qPCR for quantifying library preparation before sequencing)	3	Fixed	NA
Quantstudio service contract	3	Fixed	NA
NovaSeq 6000	3	Fixed	NA
NovaSeq 6000 service contract	3	Fixed	NA
Small equipment			
Metal racks	3	Fixed	NA
Shipping & Ordering (Units: counts per trio)			
Samples	3	Fixed	NA
Informatics (Units: uses per trio)			
Software and storage	3	Fixed	NA

Abbreviations: cap., Capture; DNA, Deoxyribonucleic acid; ES, exome sequencing; GS, genome sequencing; min, minute; Hybe., Hybridization; NA, not applicable; PCR, polymerase chain reaction; qPCR, quantitative polymerase chain reaction.

Table 3. Unit price point estimates, distributions and ranges for GS and ES cost items

Items	Estimate	Job Class	Distribution	Range
Labour(Unit: CAD/min)				
Sample preparation				
DNA Extraction method 1				
QIAsymphony platform	Conf.	Laboratory technician	Trun. Norm. ~ [μ, σ = Conf.]	[0.52, 0.78]
DNA Extraction method 2				
Puregene	Conf.	Laboratory technician	Trun. Norm. ~ [μ, σ = Conf.]	[0.52, 0.78]
QIAamp DNA Mini kit	Conf.	Laboratory technician	Trun. Norm. ~ [μ, σ = Conf.]	[0.52, 0.78]
Specimen receipt	Conf.	Laboratory technician	Trun. Norm. ~ [μ, σ = Conf.]	[0.52, 0.78]
Specimen accessioning	Conf.	Laboratory technician	Trun. Norm. ~ [μ, σ = Conf.]	[0.52, 0.78]
Test approval and ordering	Conf.	Coordinator/G.C.	Trun. Norm. ~ [μ, σ = Conf.]	[0.76, 1.14]
Library preparation - GS				
Covaris	Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
TapeStation	Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
Build worksheets/pull samples	Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
library preparation	Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
Master mix preparation	Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
Kapa preparation	Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
HSD1000	Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
Library Preparation - ES				
Covaris	Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
Build worksheets/pull samples	Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
Library prep Bravo	Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
Pre-cap PCR	Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
Pre-cap ampure	Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
D1000 TapeStation	Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
Hybe. -cap. + wash	Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
Post-cap PCR	Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
Final ampure	Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
HSD1000	Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]

Kapa prep	Conf.		Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
Sequencing					
Calculations		Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
Pooling		Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
Dilution/PCR set up		Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
Extension PCR		Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
Analysis		Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
Technical supervision		Conf.	Resource lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.73, 1.09]
Informatics					
Senior bioinformatician		Conf.	Senior bioinformatician	Trun. Norm. ~ [μ, σ = Conf.]	[0.80, 1.19]
Confirmatory testing					
Deletion					
PCR setup		Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
Analysis		Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
Duplication					
DNA plate preparation to PCR steps		Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
Purification		Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
Sanger sequencing					
PCR setup; electrophoresis; sequencing set-up and Sequencing clean up		Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
Analysis (first and second check)		Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
Primer design		Conf.	Medical lab. technologist	Trun. Norm. ~ [μ, σ = Conf.]	[0.68, 1.02]
Interpretation & report					
Genome analysis		Conf.	Laboratory specialist	Trun. Norm. ~ [μ, σ = Conf.]	[0.81, 1.22]
Medical/Scientific review		Conf.	Laboratory director	Trun. Norm. ~ [μ, σ = Conf.]	[1.19, 1.79]
Customer service		Conf.	Customer service rep	Trun. Norm. ~ [μ, σ = Conf.]	[0.43, 0.64]
Administrative support		Conf.	Admin assistant	Trun. Norm. ~ [μ, σ = Conf.]	[0.80, 1.19]
Case conference					
Genome analysis		Conf.	Laboratory specialist	Trun. Norm. ~ [μ, σ = Conf.]	[0.81, 1.22]
Clinical geneticist		Conf.	Clinical geneticist	Fixed	
Laboratory director		Conf.	Laboratory director	Trun. Norm. ~ [μ, σ = Conf.]	[1.19, 1.79]

Supplies (Unit: CAD/trio) - GS			
Reagents			
Illumina NovaSeq S1/S4 reagents kit	2418.59	Trun. Norm. ~ [2418.59,72.56]	[2200.92, 2636.26]
Confirmatory testing - deleting	329.73	Trun. Norm. ~ [329.73, 8.24]	[305.00,354.46]
Confirmatory testing - duplication	427.29	Trun. Norm. ~ [427.29, 10.68]	[395.24, 459.34]
Confirmatory testing - Sanger sequencing	10.21	Trun. Norm. ~ [10.21, 0.26]	[9.44, 10.98]
Other reagents	243.73	Trun. Norm. ~ [243.73,6.09]	[225.45, 262.01]
Consumables			
Confirmatory testing - deleting	20.98	Trun. Norm. ~ [20.98, 0.52]	[19.41, 22.55]
Confirmatory testing - duplication			
Confirmatory testing - Sanger sequencing	7.05	Trun. Norm. ~ [7.05, 0.18]	[6.52, 7.58]
Other consumables	118.90	Trun. Norm. ~ [118.90, 3.96]	[107., 130.79]
Supplies (Unit: CAD/trio) - ES			
Reagents			
Illumina NovaSeq S1/S4 reagents kit	794.94	Trun. Norm. ~ [794.94, 23.85]	[723.40, 866.48]
Confirmatory testing - deleting	329.73	Trun. Norm. ~ [329.73, 8.24]	[305.00,354.46]
Confirmatory testing - duplication	427.29	Trun. Norm. ~ [427.29, 10.68]	[395.24, 459.34]
Confirmatory testing - Sanger sequencing	10.21	Trun. Norm. ~ [10.21, 0.26]	[9.44, 10.98]
Other reagents	171.24	Trun. Norm. ~ [159.29, 4.28]	[158.40, 184.08]
Consumables			
Confirmatory testing - deleting	20.98	Trun. Norm. ~ [20.98, 0.52]	[19.41, 22.55]
Confirmatory testing - duplication			
Confirmatory testing - Sanger sequencing	7.05	Trun. Norm. ~ [7.05, 0.18]	[6.52, 7.58]
Other consumables	490.63	Trun. Norm. ~ [441.57, 16.35]	[441.57, 539.69]
Equipment (Unit: CAD per machine)			
Large equipment			
BioAnalyzer/TapeStation	60000.00	Trun. Norm. ~ [60000.00, 2000.00]	[54000.00, 66000.00]
BioAnalyzer/TapeStation service contract	4828.80	Trun. Norm. ~ [4828.80, 160.96]	[4345.92,5311.68]
Bravo robot	307228.00	Trun. Norm. ~ [307228.00, 10240.93]	[276505.20, 337950.80]
Bravo robot - service contract	48556.48	Trun. Norm. ~ [48556.48, 1618.55]	[43700.83, 53412.13]
Covaris ultrasonicator	90815.00	Trun. Norm. ~ [90815.00, 3027.17]	[81733.50, 99896.50]

Covarisultrasonicator service contract	1104.00	Trun. Norm. ~ [1104.00, 110.4]	[993.60, 1214.40]
Quantstudio	19,600.00	Trun. Norm. ~ [19,600.008, 653.33]	[17640.00, 21560.00]
Quantstudio service contract	7250.00	Trun. Norm. ~ [7250.00, 241.67]	[6525.00, 7975.00]
NovaSeq 6000	1429848.00	Trun. Norm. ~ [1429848.00, 47661.6]	[1286863.20, 1572832.80]
NovaSeq 6000 service contract	140125.52	Trun. Norm. ~ [140125.52, 4670.85]	[126112.97, 154138.07]
Small equipment			
Metal racks	0.11	Trun. Norm. ~ [0.11,0.01]	[0.1, 0.12]
Shipping & ordering (Unit: CAD per trio)			
Shipping & order	67.45	Trun. Norm. ~ [67.45, 2.25]	[60.71, 74.20]
Software (CAD/trio)			
GS			
Software cost for informatics	635.67	Trun. Norm. ~ [635.67, 16.95]	[584.82, 686.52]
ES			
Software cost for informatics	271.49	Trun. Norm. ~ [271.49, 7.24]	[249.77, 293.21]

Abbreviations: CAD, Canadian dollar; cap., Capture; Conf., confidential; DNA, Deoxyribonucleic acid; ES, exome sequencing; G.C., genetic counsellor; GS, genome sequencing; Hybe., Hybridization; lab., laboratory; min, minute; PCR, polymerase chain reaction; Trun. Norm., truncated normal distribution.

2.5.2.1 Genome sequencing

This section presents the calculations for valuing inputs related to labour, reagents, consumables, equipment, and software. The resource use and unit price data for each input were provided either by laboratory staff, manufacturers or from publicly available sources. Costs were calculated by multiplying the quantity used by the unit price. Labour cost was calculated by multiplying the minutes spent per task by the labour cost per minute. All prices were collected in 2022 dollars, except equipment, which was adjusted to 2023 dollars over the multi-year shelf-life as explained below. For equipment, a straight-line calculation method was used whereby depreciation and opportunity cost were applied to the purchase price of each item in the year of purchase, ensuring that the price was evenly distributed over the asset's shelf life. Then based on the chosen year of calculation, the present value was calculated by applying the present value formula at a 3% discount rate from the year of purchase.

2.5.2.1.1 Labour

Hospital and laboratory professionals involved in GS testing included registered nurse, laboratory technician, clinical geneticist, lead laboratory technician, medical laboratory technologist, resource laboratory technician, genetic counsellor, laboratory specialist, senior bioinformatician, clinical laboratory director, laboratory quality specialist and administrative staff. The type of professionals and their labour time (measured in minutes) were recorded for each task within the workflow as per Table 2. Where applicable, labour workload units followed the Management Information Systems (MIS) Standards (38) which were applied to laboratory components. The MIS standards provide national standards for gathering and processing data and for reporting financial and statistical data on day-to-day operations of health service organizations (38). Approximations for labour time for tasks not covered by MIS Standards were provided by laboratory staff or GSO research staff (i.e., research coordinator). Labour categories consisted of sample preparation, library preparation, sequencing, confirmatory testing, informatics and clinical interpretation and reporting.

Thanks to technological advancement and automation, samples were processed in batches for several workflow steps. The batch size for each workflow step varied based on the type of platform being used. For DNA extraction, the QIA Symphony platform extracted DNA in batches of 24 samples. Pre-sequencing preparation procedures such as shearing, library preparation, and quantification were also conducted in batches of 24 samples. The large equipment sequencing machine, NovaSeq 6000,

processed 24 samples per run, while sample identification verification (SIV) was handled in batches of 46 samples. The labour volume (in minutes) for each sample preparation step was initially calculated per batch. This figure was then divided by the batch size to determine the labour volume per sample.

For the sample preparation step, two methods were used to extract and purify DNA. Approximately 95% of all samples sequenced underwent QIAsymphony platform for DNA extraction. For the remaining 5% of samples, QIAamp DNA Mini Kit and the Puregene DNA Kit were used for DNA extraction (personal communication, laboratory director). The latter method was used to extract DNA from limited blood specimens such as samples collected from infants. Labour time for each DNA extraction method was determined by laboratory staff following MIS standards (38). Because two methods were used for DNA extraction, a proportion was applied to each method and aggregated to determine the total cost of the sample preparation step.

To estimate labour resource use per trio (consisting of samples from the proband and two biological parents), the labour volume per sample was multiplied by three to calculate volume per trio. This formula was applied to several workflow steps, but not all of them. Some workflow steps did not depend on the batch size, such as confirmatory testing. The labour time for analytical, informatics and confirmatory testing was assumed to include sample logistics management, including tasks such as starting computing jobs, tracking samples, transferring data, as well as data processing. The latter included periodic updates to the annotation pipeline (personal communication, laboratory director).

Labour time used in interpretation and reporting included time devoted to primary variant analysis for all cases, secondary variants analysis when applicable, and case conferences for complex cases. As analysts were involved in both interpretation and report writing, blinding to sequencing strategy was not possible. As most variants were exonic, interpretation was agnostic to the use of ES or GS. Analysis time invested by the genome analyst and by the laboratory director were collected prospectively for analysis of every primary variant and for secondary variants when applicable. Only trios with complex findings (e.g., potentially diagnostic) underwent case conferences, which took place at both institutions (SickKids & CHEO). The occurrence of a case conference was noted in the REDCap database. A case conference typically included one genetic analyst, one clinical geneticist and one laboratory director with an average of 10 minutes per case assumed.

Labour price inputs consisting of hourly wage rates for each professional were collected by the laboratory staff in December 2022. Employee benefits were added at 18% for laboratory director and 25% for laboratory staff based on SickKids policy. The fee code for clinical geneticist was adopted using the 2022 SOB (37). Because of the confidential nature of salary information, reporting of wages for labour items has been suppressed. For most of the salary price inputs, ranges were applied at 20%, except for the clinical geneticist, which was unvaried because it reflects the Ontario fee-for-service schedule. Let 'l' denote labour, the labour cost per sample for each task was calculated using the following formula:

$$c_{j,k}^{l} = \frac{t_{j,k}^{l} * r_k}{b_j} \tag{1}$$

Where

$t_{j,k}^{l}$: $Labour\ time\ per\ batch\ spent\ on\ task\ k\ in\ each\ workflow\ step\ j$

r_k: $Labour\ rates\ on\ task\ k$

b_j: $Batch\ size\ per\ run\ in\ workflow\ step\ j$

The costs of labour per sample for all tasks K in each workflow step j were then aggregated using the following equation:

$$C_{j,K}^{l} = \sum_{k=1}^{K} c_{j,k}^{l} \tag{2}$$

Where

$C_{j,k}^{l}$: $Total\ cost\ per\ singleton\ sample\ for\ all\ tasks\ K\ used\ in\ workflow\ step\ j$

Hence, to calculate total labour cost per trio, the following equation was adopted:

$$C_{trio}^{l} = \sum_{i}^{I} \sum_{j=1}^{J} 3 * C_{i,j,K}^{l} \tag{3}$$

Where

C_{trio}^{l}: $Total\ labour\ cost\ per\ trio\ for\ all\ workflow\ stages$

i: $Index\ of\ workflow\ stage$

j: $Index\ of\ workflow\ step$

$C_{i,j,K}^{l}$: $Total\ cost\ per\ singleton\ sample\ for\ all\ task\ K\ in\ workflow\ step\ j$

2.5.2.1.2 Supplies

The cost of supplies was determined from the costs of the reagents and consumables used in various workflow steps, including sample preparation, library preparation, sequencing and confirmatory testing. While some workflow steps had similar calculations across different platforms, others were distinct depending on the type of platform. In general, supply cost was initially calculated on a per-sample basis and then multiplied by three to determine the cost per trio. To reduce redundancy in calculation notation, the letter 's' was used to denote supply, representing either reagents or consumables. A supply cost per sample for all items in each workflow step was computed using the following formula:

$$c_{j,k}^{s} = \frac{v_{j,k}^{s} * p_k}{b_j} \tag{4}$$

$$C_{j,K}^{s} = \sum_{K=1}^{K} c_{j,k}^{s} \tag{5}$$

Where

$v_{j,k}^{s}$: $Volume\ use\ per\ batch\ for\ supply\ k\ in\ workflow\ step\ j$

p_k: $Unit\ price\ of\ supply\ k$

b_j: $Batch\ size\ per\ run\ in\ workflow\ step\ j$

$C_{j,K}^{s}$: $Total\ cost\ per\ singleton\ sample\ for\ all\ supplies\ K\ in\ workflow\ step\ j$

A total supply cost per trio in each workflow stage was determined by multiplying the cost per sample by three and summing costs across all workflow steps as follows:

$$C_{trio}^{s} = \sum_{i}^{I} \sum_{j=1}^{J} 3 * C_{i,j,K}^{s} \tag{6}$$

Where

C^{s}_{trio}: $Total\ supply\ cost\ per\ trio\ for\ all\ workflow\ stages$

i: $Index\ of\ workflow\ stage\ in\ GS$

j: $Index\ of\ workflow\ steps\ in\ GS$

$C^{s}_{i,j,K}$: $Total\ cost\ per\ singleton\ sample\ for\ all\ supplies\ K\ in\ workflow\ step\ j$

Applying formulae (4), (5) and (6), the total cost per trio for all supplies was calculated and aggregated for all workflow stages. However, some workflow steps operated differently due to the platform it used. The above formulae were modified to accommodate different workflow steps i (i.e., sample preparation and confirmatory testing). Confirmatory testing consisted of DNA deletion and duplication detection and Sanger sequencing. Due to the nature of the sequencing platform, the total supply cost per trio was not tripled for deletion/duplication detection as the workflow steps were undertaken regardless of the number of samples per run.

2.5.2.1.2.1 Sample preparation supplies

The sample preparation workflow step was slightly different than other workflow steps. In sample preparation, the cost of supplies varied based on the DNA extraction method. Each method represented a distinct workflow step. There were two methods for extracting DNA from blood specimen: (1) the QIAsymphony platform or (2) a combination of a two-step extraction method using QIAamp DNA Mini Kit and the Puregene DNA Kit. The main difference between these two methods was that the latter required less blood, making it suitable for specimens drawn from newborns. Approximately 5 percent of samples underwent the QIAamp and the Puregene extraction methods (personal communication, laboratory director). Therefore, adopting formulae (4) and (5) and applying the following formula, the total supply cost per trio used in sample preparation was calculated as:

$$C^{sp}_{j,K} = p_1 * C^{s}_{j,K} + (1 - p_1) * (C^{s}_{j+1,K} + C^{s}_{j+2,K}) \quad (7)$$

Where

$C^{SP}_{j,K}$: $Total\ supply\ cost\ per\ sample\ of\ all\ sample\ preparation\ steps$

p_1: $Proportion\ of\ trios\ who\ underwent\ DNA\ extraction\ method\ 1$

$C^{s}_{j,K}$: $Cost\ per\ sample\ for\ total\ supply\ K\ used\ in\ workflow\ step\ j, defined\ in\ (5)$

2.5.2.1.2.2 Confirmatory testing supplies

Due to the complexity of sequencing, GS could produce sequencing errors, incorrect alignment and random sampling issues (39). To address this, follow-up testing methods such as qPCR and Sanger sequencing were incorporated into the GS workflow to ensure the accuracy and reliability of the sequencing results. Confirmatory testing was only performed on primary variants using two different testing strategies: deletion detection (qPCR) or duplication detection (Affymetrix Array), and Sanger sequencing. Deletions included the absence of genetic material whereas duplications signified an extra segment of DNA compared to a reference genome (40). According to laboratory staff, only 1% of the GSO samples sequenced at SickKids underwent either deletion or duplication detection during GS and those samples were split evenly between deletion and duplication detection. Due to the nature of the sequencing system, deletion or duplication detections were performed regardless of the number of samples. Thus, the supply cost per trio for deletion and for duplication detection was the same as the cost per sample, adopting the following equation:

$$C_{j,K}^{sdd} = q\left[p * C_{j,K}^{s} + (1-p) * \ C_{j+1,K}^{s}\right] \quad (8)$$

Where

$C_{j,K}^{sdd}$: $Cost\ per\ trio\ for\ all\ supplies\ K\ used\ in\ deletion\ and\ duplication$

p: $Proportion\ of\ samples\ underwent\ workflow\ step\ j$

q: $Proportion\ of\ samples\ underwent\ deletion\ or\ duplication\ in\ GS$

$C_{j,K}^{s}$: $Cost\ per\ sample\ for\ total\ supply\ K\ used\ in\ workflow\ step\ j, defined\ in\ (5)$

Sanger sequencing is the accepted gold standard to detect errors or ambiguous results produced by GS (39). It can help ensure the accuracy and reliability of suspicious primary variants. It was assumed that 5% of total samples sequenced under GSO required Sanger sequencing (personal communication, laboratory director). Adopting equation (8) above and assigning p = 1, a cost per sample for Sanger sequencing was calculated as follows:

$$C_{j,K}^{ss} = q * C_{j,K}^{s} \quad (9)$$

Where

$C_{j,K}^{ss}$: $Cost\ per\ trio\ for\ all\ supplies\ K\ used\ in\ Sanger\ sequencing$

q: $Proportion\ of\ samples\ underwent\ Sanger\ sequencing$

$C^{s}_{j,K}$: $Cost\ per\ sample\ for\ total\ supply\ K\ used\ in\ workflow\ step\ j, defined\ in$ (5)

Unlike qPCR/Affymetrix Array which were run per trio, Sanger sequencing was run per sample and the cost was tripled to determine a cost per trio.

2.5.2.1.3 Equipment

Costs were estimated for both small and large equipment. Large equipment included the NovaSeq 6000, TapeStation, Bravo Robot, Quantstudio and Covaris ultrasonicator. For large equipment, cost was calculated based on its acquisition price and service contract. Each large equipment machine required maintenance on a yearly basis covered by a service contract. Service contracts were for one year for all large equipment except the Bravo Robot, which was serviced under a three-year contract. The prices of the platforms and service contracts were provided either by the manufacturer (e.g., Illumina) or by SickKids laboratory staff. Over the course of the study, equipment prices varied from -3% to 11% depending on the machine and the lab manager estimated an overall fluctuation of 10%. Small equipment included reusable items such as metal racks for storing chemical boxes. The cost of small equipment was first calculated on a per-sample basis, then tripled to determine the cost per trio.

A 2023 expected cost for each piece of large equipment and each equipment contract were determined based on the shelf-life or term, and the year of purchase. A 3% opportunity cost and a 3% discount rate were applied over the shelf-life/term. For each piece of large equipment and corresponding contract, an institutional cost per sample was determined based on the 2023 expected cost (E_{pv}), the actual sample volume (V_A) for all indications within the institution per year, the actual sample volume for all GSO applications V_{GSO} (randomized and non-randomized), and the actual volume undergoing GS. The institutional sample volume was multiplied by the proportion of patients sequenced at SickKids that were GSO ($\frac{V_{GSO}}{V_A}$). The GSO volume was multiplied by the proportion of GSO patients that were GS ($\frac{V_{GS}}{V_{GSO}}$) for each equipment e. That fraction was subsequently multiplied by cost per sample (E_{pv}) and by three to determine a GSO-specific trio cost for GS. Thus, adopting the following formula:

$$v^{e}_{GS} = \left(\frac{1}{V_A}\right) * \left(\frac{V_{GSO}}{V_A}\right) * \left(\frac{V^{e}_{GS}}{V_{GSO}}\right)$$

(10)

$$C_{trio}^{e} = 3 * E_{PV}^{e} * v_{GS}^{e}$$

(11)

Where

v_{GS}^{e}: $Actual\ volume\ (\%) per\ sample\ for\ equipment\ e\ underwent\ GS$
V_A: $Actual\ sample\ volume\ per\ year\ for\ all\ indications\ in\ the\ institution$
V_{GSO}: $Actual\ sample\ volume\ per\ year\ for\ GSO\ cohort\ in\ the\ institution$
V_{GS}^{e}: $Actual\ sample\ volume\ per\ year\ for\ GS\ cohort\ in\ the\ institution$
C_{trio}^{e}: $Cost\ per\ trio\ for\ given\ large\ equipment\ e$
E_{PV}^{e}: $Expected\ cost\ of\ equipment\ e\ after\ applying\ discount\ and\ depreciation$

A total cost for all equipment was calculated by summing the cost per sample for each equipment i and its service contract using the following formula:

$$C_{trio}^{E} = \sum_{e}^{E} C_{trio}^{e}$$

(12)

Where

C_{trio}^{E}: $Total\ cost\ per\ trio\ for\ all\ equipment\ used\ in\ Sequencing$
C_{trio}^{e}: $Total\ cost\ per\ trio\ for\ each\ equipment\ e$

Applying formulae (10), (11) and (12), the cost of all equipment and their service contracts was calculated and aggregated for each trio GS test.

2.5.2.1.4 Informatics

Informatics refers to the use of computational tools and techniques to analyze and interpret the vast amount of output data generated by GS. The analysis involved a complex series of steps, including registration, intensity extraction, phasing correction, base calling and quality scoring (Figure 4). Items related to informatics were analysis software and storage, informatics common equipment and informatics software licensing (personal communication, senior laboratory manager) as listed in Table 1. Volumes of use for informatics software, storage and equipment were determined initially by sample, then tripled to calculate cost per trio. Unit prices for informatics for software, storage and equipment were provided by laboratory staff. Most software and storage costs were determined by multiplying unit

price by volume of consumption. This included Illumina ICA active storage (transition), data archive storage, Isilon storage (transition), VM ICA connector and VM thing3. Software licensing fees were included for Emedgene and Illumina ICA at cost per trio. Illumina Dragen analysis software which was used for storage and analysis on the cloud, was costed differently. To perform analysis on the cloud, iCredits were used to purchase data storage and analysis within the Illumina cloud software ecosystem. Each sample consumed fixed iCredit units cr_{sample} and additional cr_{trio} if a trio analysis was performed. For GS, each sample consumed 18 iCredit units and for trios, an additional 15 iCredit units were required. Hence, the cost per trio for Dragen analysis software was calculated as follows:

$$C_{trio\_Dragen} = \left[\left(3 * cr_{sample}\right) + cr_{trio}\right] * c * e \quad (13)$$

Where

cr_{sample}: $icredit\ units\ used\ per\ sample$

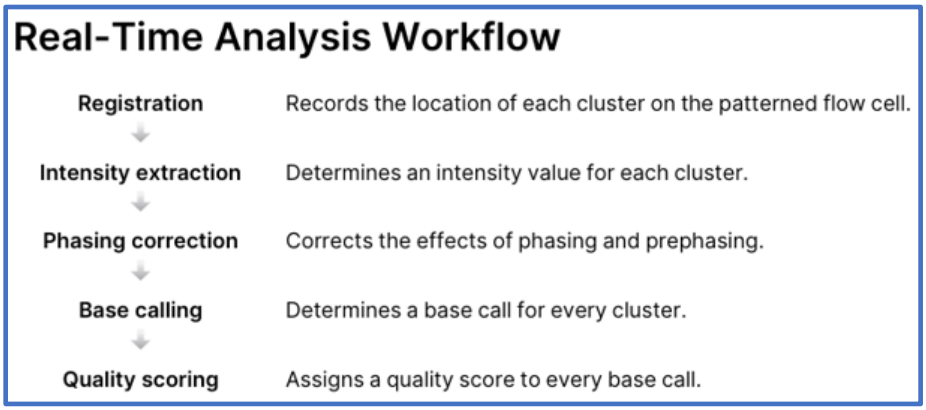
cr_{trio}: $icredit\ units\ used\ per\ trio$

c: $cost\ per\ icredits\ in\ USD$

e: $exchange\ rate\ of\ USD/CAD$

Some software and cloud storage used servers located in the US (e.g., Illumina Dragen analysis and Illumina ICA active storage). US unit prices were converted to CAD using 2022 USD/CAD rate from BoC (36) (1 USD = 1.30 CAD). The cost per trio for each item used under informatics was then aggregated to total cost per trio.

Figure 4. Informatics analysis workflow map



2.5.2.2 Exome sequencing

This section presents the calculations for valuing inputs related to labour, reagents, consumables, equipment, and software in the ES workflow. The resource use and unit price data for each input were provided either by laboratory staff, manufacturers or from publicly available sources. Costs were calculated by multiplying the quantity used by the unit price. Labour cost was calculated by multiplying the minutes spent per task by the labour cost per minute. All prices were collected in 2022 dollars except equipment, which was adjusted to 2023 dollars over the multi-year shelf-life as explained below. For equipment, a straight-line calculation method was used whereby depreciation and opportunity cost were applied to the purchase price of each item in the year of purchase, ensuring that the price was evenly distributed over the asset's shelf life. Based on the chosen year of calculation, the present value was then calculated by applying the present value formula at a 3% discount rate from the year of purchase.

In the analytical stage of ES, additional steps, including PreCapture PCR, hybridization/capture wash, PostCapture PCR, and Kapa qPCR were involved in library preparation. This is because ES focused on a very small fraction of the genome (i.e., the exonic regions) while non-coding regions are largely irrelevant to the analysis. These additional steps were designed to enrich the exonic regions and minimize the sequencing of non-coding DNA. In contrast, GS sequenced all regions of the genome without the need for specific enrichment, making these additional steps unnecessary.

2.5.2.2.1 Labour

Hospital and laboratory professionals involved in ES testing included registered nurse, laboratory technician, clinical geneticist, lead laboratory technician, medical laboratory technologist, resource laboratory technician, genetic counsellor, laboratory specialist, senior bioinformatician, clinical laboratory director, laboratory quality specialist and administrative staff. The type of professionals and their labour time (measured in minutes) were recorded for each task within the workflow as per Table 2. Where applicable, labour workload units followed the MIS Standards (38) which were applied to laboratory components. The MIS standards provide national standards for gathering and processing data and for reporting financial and statistical data on day-to-day operations of health service organizations (38). Approximations for labour time for tasks not covered by MIS Standards were provided by laboratory staff or GSO research staff (e.g., research coordinator). Labour categories

consisted of sample preparation, library preparation, sequencing, confirmatory testing, informatics and clinical interpretation and reporting.

Similar to GS, samples were processed in batches for various ES workflow steps. The batch size for each workflow step varied based on the type of platform being used. For DNA extraction, the QIA Symphony platform extracted DNA in batches of 24 samples. Pre-sequencing preparation procedures, such as shearing, library preparation, and quantification were also conducted in batches of 24 samples. The large equipment sequencing machine, NovaSeq 6000, processed 24 samples per run, while SIV was handled in batches of 46 samples. The labour volume (in minutes) for each sample preparation step was initially calculated per batch. This figure was then divided by the batch size to determine the labour volume per sample.

In the sample preparation step, the same methods were used to extract and purify DNA for ES. Approximately 95% of all samples sequenced underwent QIAsymphony platform for DNA extraction. For the remaining 5% of samples, QIAamp DNA Mini Kit and the Puregene DNA Kit were used for DNA extraction (personal communication, laboratory director). The latter method was used to extract DNA from limited blood specimens such as samples collected from infants. Labour time for each DNA extraction method was determined by laboratory staff following MIS Standards (38). Because two methods were used for DNA extraction, a proportion was applied to each method and aggregated to determine the total cost of the sample preparation step.

To estimate labour resource use per trio (consisting of samples from the proband and two biological parents), the labour volume per sample was multiplied by three to calculate volume per trios. This formula was applied to several workflow steps, but not all of them. Some workflow steps did not depend on the batch size, such as confirmatory testing. The labour time for analytical, informatics and confirmatory testing were assumed to include sample logistics management, including tasks such as starting computing jobs, tracking samples, transferring data, as well as data processing, which included periodic updates to the annotation pipeline (personal communication, laboratory director).

Labour time used in interpretation and reporting included time devoted to primary variant analysis for all cases, secondary variants analysis when applicable, and case conferences for complex cases. As

analysts were involved in both interpretation and report writing, blinding to sequencing strategy was not possible. As the majority of variants were exonic, interpretation was agnostic to the use of ES or GS. Analysis time invested by the genome analyst and the laboratory director were collected prospectively for analysis of every primary variant and for secondary variants when applicable. Only trios with complex findings (e.g., potentially diagnostic) underwent case conferences, which took place at both institutions (SickKids & CHEO). The occurrence of a case conference was noted in the REDCap database. A case conference typically included one genetic analyst, one clinical geneticist and one laboratory director with an average of 10 minutes per case assumed.

Labour price inputs consisting of hourly wage rates for each professional were collected by the laboratory staff in December 2022. Employee benefits were added at 18% for laboratory director and 25% for laboratory staff based on SickKids policy. The fee code for clinical geneticist was adopted using the 2022 SOB (37). Because of the confidential nature of salary information, reporting of wages for labour items has been suppressed. For most of the salary price inputs, ranges were applied at 20%, except for the clinical geneticist, which was unvaried because it reflects the ON fee-for service schedule. In ES workflow, Labour cost per sample for each task was calculated using fomulae (1), (2) and (3) defined in section 2.5.2.1.1.

2.5.2.2.2 Supplies

Cost of supplies used in ES was determined from the costs of the reagents and consumables used in various workflow steps, including sample preparation, library preparation, sequencing and confirmatory testing. While some workflow steps had similar calculations across different platforms, others were distinct depending on the type of platform. In general, supply cost was initially calculated on a per-sample basis and then multiplied by three to determine the cost per trio. Applying formulae (4), (5) and (6) defined previously in section 2.5.2.1.2, the total cost per trio for all supplies was calculated and aggregated for all workflow stages. Similar to GS, some workflow steps operated differently due to the platform used. The above formulae (4), (5) and (6) were modified to (7), (8) and (9) accordingly to accommodate cost of supplies in different ES workflow steps.

2.5.2.2.2.1 Sample preparation supplies

The sample preparation workflow step was slightly different than other workflow steps. In sample preparation, the cost of supplies varied based on the DNA extraction method. Each method represented

a distinct workflow step. There were two methods for extracting DNA from blood specimen: (1) the QIAsymphony platform or (2) a combination of a two-step extraction method using QIAamp DNA Mini Kit and the Puregene DNA Kit. The main difference between these two methods was that the latter required less blood, making it suitable for specimens drawn from newborns. Approximately 5 percent of samples underwent the QIAamp and the Puregene extraction methods (personal communication, laboratory director). Applying equation (7), the total supply cost per trio was calculated for sample preparation.

2.5.2.2.2.2 Confirmatory testing supplies

Due to the complexity of sequencing, ES could also produce sequencing errors, incorrect alignment and random sampling issues (39). To address this, follow-up testing methods such as qPCR and Sanger sequencing were incorporated into the ES workflow to ensure the accuracy and reliability of the sequencing results. Confirmatory testing was only performed on primary variants using two different testing strategies: deletion detection (qPCR) or duplication detection (Affymetrix Array), and Sanger sequencing. Deletions included the absence of genetic material whereas duplications signified an extra segment of DNA compared to a reference genome (40). According to laboratory staff, around 2% of the GSO samples sequenced at SickKids underwent either deletion or duplication detection during ES, evenly split between deletion and duplication detection. Due to the nature of the sequencing system, deletion or duplication detections were performed regardless of the number of samples. Thus, the supply cost per trio for deletion and for duplication detection was the same as the cost per sample. For Sanger sequencing, it was assumed that 5% of total samples sequenced under GSO required Sanger sequencing (personal communication, laboratory director). Supply cost per trio in confirmatory testing was computed using previously defined formulae (8) and (9).

2.5.2.2.3 Equipment

Costs were estimated for both small and large equipment. Large equipment included the NovaSeq 6000, TapeStation, Bravo Robot, Quantstudio and Covaris ultrasonicator. For large equipment, cost was calculated based on its acquisition price and service contract. Each large equipment machine required maintenance on a yearly basis covered by a service contract. Service contracts were for one year for all large equipment except the Bravo Robot, which was serviced under a three-year contract. The prices of the platforms and service contracts were provided either by the manufacturer (e.g., Illumina) or by SickKids laboratory staff. Over the course of the study, equipment prices varied from -3% to 11%

depending on the machine and the lab manager estimated an overall fluctuation of 10%. Small equipment included reusable items such as metal racks for storing chemical boxes. The cost of small equipment was first calculated on a per-sample basis, then tripled to determine the cost per trio.

A 2023 expected cost for each piece of large equipment and each equipment contract were determined based on the shelf-life or term, and the year of purchase. A 3% opportunity cost and a 3% discount rate were applied over the shelf-life/term. For each piece of large equipment and corresponding contract, an institutional cost per sample was determined based on the 2023 expected cost (E_{pv}), the actual sample volume (V_A) for all indications within the institution per year, actual sample volume for all GSO applications V_{GSO} (randomized and non-randomized), and the actual volume undergoing ES. The institutional sample volume was multiplied by the proportion of patients sequenced at SickKids that were GSO ($\frac{V_{GSO}}{V_A}$). The GSO volume was multiplied by the proportion of GSO patients that were ES ($\frac{V_{ES}}{V_{GSO}}$) for each equipment e. That fraction was subsequently multiplied by cost per sample (E_{pv}) and by three to determine a GSO-specific trio cost for ES. Thus, adopting the following formula:

$$v_{ES}^{e} = \left(\frac{1}{V_A}\right) * \left(\frac{V_{GSO}}{V_A}\right) * (\frac{V_{ES}^{e}}{V_{GSO}}) \tag{14}$$

$$C_{trio}^{e} = 3 * E_{PV}^{e} * \ v_{ES}^{e} \tag{15}$$

Where

v_{ES}^{e}: $Actual\ volume\ (\%)\ per\ sample\ for\ equipment\ e\ underwent\ ES$
V_A: $Actual\ sample\ volume\ per\ year\ for\ all\ indications\ in\ the\ institution$
V_{GSO}: $Actual\ sample\ volume\ per\ year\ for\ GSO\ cohort\ in\ the\ institution$
V_{ES}^{e}: $Actual\ sample\ volume\ per\ year\ for\ GS\ cohort\ in\ the\ institution$
C_{trio}^{e}: $Cost\ per\ trio\ for\ given\ large\ equipment\ e$
E_{PV}^{e}: $Expected\ cost\ of\ equipment\ e\ after\ applying\ discount\ and\ depreciation$

Applying formulae (12), (14) and (15), the total cost of all equipment and their service contracts was calculated for each trio ES test.

2.5.2.2.4 Informatics

Informatics refers to the use of computational tools and techniques to analyze and interpret the vast amount of output data generated by GS. The analysis involved a complex series of steps, including registration, intensity extraction, phasing correction, base calling and quality scoring (Figure 4). Items related to informatics were analysis software and storage, informatics common equipment and informatics software licensing (personal communication, senior laboratory manager) as listed in Table 1. Volumes of use for informatics software, storage and equipment were determined initially by sample, then tripled to calculate cost per trio. Unit prices for informatics for software, storage and equipment were provided by laboratory staff. Most software and storage costs were determined by multiplying unit price by volume of consumption. This included Illumina ICA active storage (transition), data archive storage, Isilon storage (transition), VM ICA connector and VM thing3. Software licensing fees were included for Emedgene and Illumina ICA at cost per trio. Illumina Dragen analysis software which was used for storage and analysis on the cloud, was costed differently. To perform analysis on the cloud, iCredits were used to purchase data storage and analysis within the Illumina cloud software ecosystem. For ES, each sample consumed 9 iCredit units and for trios, an additional 3 iCredit units were required. Hence, the cost per trio for Dragen analysis software was calculated using formula (13) defined above.

2.6 Assumptions

Due to uncertainty in cost item volumes and/or unit prices, some assumptions were required. The assumptions of the micro-costing model are detailed in Table 4. All professional wages as described in section 2.5.2.1.1 for GS and 2.5.2.2.1 for ES were reported based on 2022 rates and it was assumed that there were no significant increases between 2022-2023. Case conferences were arranged when needed to discuss the sequencing results of the trio (proband and their biological parents). The trio's primary and secondary variants, if applicable, were discussed during the meeting. Key personnel present included a genome analyst, a laboratory director, and a clinical geneticist with an average of 10 minutes per conference assumed for all staff (personal communication, GSO project coordinator). To capture uncertainty, time spent by genome analyst and laboratory director were assumed to vary between 5 to 10 minutes. A total of 93/329 (28.27%) and 81/324 (25.00%) of trios underwent case conferences for GS and ES, respectively.

Examining manufacturer invoices and price lists over the course of the study revealed that prices for reagents and consumables increased by 2-9%, depending on the item. At the same time, vendor

discounts varied between 4-16% on 19 consumable products (mean 9.6%) and currency exchange rates fluctuated between 1-4%. Given the large number of items affected, a pragmatic 10% variation was assigned to prices.

Within the sample preparation steps, two distinct DNA extraction methods were employed to obtain DNA from blood. It was assumed that the QIA Symphony Platform accounted for 95% of samples and the Puregene Kit plus QIAamp Mini Kit was used for 5% of samples labour (personal communication, laboratory director).

In confirmatory testing step, qPCR and Affymetrix array were used for deletion and duplication detection and were assumed to be employed at a rate of 1% in GS and 2% in ES (personal communication, laboratory director). Furthermore, Sanger sequencing was assumed to be conducted at a rate of 5% for both GS and ES (personal communication with laboratory director).

A rate of 3% for opportunity cost and 3% for depreciation were applied to the cost of large equipment and associated service 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 the equipment at each time point. Furthermore, the useful lifetime of small equipment was assumed to be 5 years, whereas the useful lifetime of large equipment varied from 8 to 10 years. A 3% discount rate was therefore applied to all equipment-related items, including their service contracts, over their lifespan. Estimated total sample volume run for each piece of large equipment (excluding NovaSeq 6000) across all indications in the institution was 5000 samples per year, whereas the NovaSeq 6000 core sequencing systems were estimated to run 4600 samples annually across both machines (personal communication, laboratory director).

Overhead costs encompassing administrative and infrastructure expenses were incorporated into labour, equipment (both small and large), and informatics costs. A previous query to the Ontario MoH Case Costing Initiative database provided an estimate of overhead costs for major interventions across different hospitals in Ontario. The average overhead cost in Ontario for the year 2016/2017 was 22.3%, with a range of 15.8% to 35.1%. An internal request to SickKids' case costing and decision support unit yielded a SickKids-specific overhead estimate for patient wards of 31.6%. Considering these sources, the reference overhead cost case was assumed to be 22.3%, with a range of 10% variation for 2023.

Table 4. Assumptions used in micro-costing analysis

GS/ES	Description
Labour	• No significant changes in labour price between 2019-2023 • Labour time spent on tasks was assumed to include sample logistics management, including starting computing jobs, tracking samples, transferring data, as well as data processing, which included periodic updates to the annotation pipeline • Each case conference was assumed to be hosted by a genome analyst, a laboratory director and a clinical geneticist, with an average of 10 minutes per case conference. Both genome analyst and laboratory director time were assumed to vary between 5 to 10 minutes
DNA Extraction	• Two DNA extraction methods were used: (i) QIAsymphony platform, assumed for 95% of the samples (ii) Puregene Kit + QIAamp Mini Kit, assumed for 5% of the samples
Confirmatory testing	• Confirmatory testing was used for primary variants, with qPCR and Affymetrix array used to detect deletions/duplications (1% of GS and 2% for ES) • Sanger sequencing was applied to 5% of GS and ES samples
Equipment	• Small equipment useful lifetime was 5 years • Large equipment useful lifetime varied from 8 to 10 years • 3% opportunity cost and 3% depreciation applied annually to large equipment and their service contracts (if terms > 1 year) • The total sample volume run on each piece of large equipment (excluding NovaSeq 6000) for all indications in the institution was estimated at 5000 samples per year • Both NovaSeq 6000 systems were estimated to run a total of 4600 samples annually for all indications in the institution
General	• Overhead cost of 22.3% applied to labour, supplies, small and large equipment • Variation of 10%

Abbreviations: DNA, deoxyribonucleic acid; ES, exome sequencing; GS, genome sequencing; qPCR, Real-time polymerase chain reaction

2.7 Micro-costing analysis

Micro-costing of GS and ES involved identifying, measuring and valuing the individual cost items for each sequencing workflow stage. By applying the formulae defined in method section 2.3, the total cost for GS and ES per trio was calculated deterministically by aggregating the cost of each component by category for all workflow stages. The model was built in Microsoft Excel (version 2023).

2.8 Probabilistic analysis

A PA was carried out to incorporate uncertainty in model parameters such as volume and price inputs. Each uncertain parameter was characterized by an estimate, range and distribution. The resource use volumes for inputs were either proportions or point estimates. Distributions were defined and assigned for resource use and unit price of inputs that were potentially changing over the study period (Tables 3 and 4). For uncertain inputs, a lower and an upper bound were assigned based on empirical data or personal communication with laboratory staff (laboratory director or senior research manager) as previously described.

Most price and resource use parameters were assigned to truncated normal distributions where a point estimate corresponded to the mean of the normal distribution and lower and upper bounds corresponded to a 99.7 confidence interval (CI) (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 volume or unit price for each input, assumed to follow a normal distribution (N) bounded at zero, $0 < X < \infty$, μ corresponds to the point estimation of X; $\sigma = \frac{u-l}{6}$, u represents the upper bound, and l represents the lower bound of a range. The 99.7% CI was chosen to determine a level of confidence for the upper and lower bound for an input. The normal distribution was truncated at zero since resource use and prices cannot be negative. This method was applied to supplies (reagents and consumables), equipment (small and large equipment, service contracts) and informatics. The resources used for confirmatory testing, including qPCR, Affymetrix array and Sanger sequencing, was quantified as the proportion of cases that underwent testing as described earlier. At the individual case

level, confirmatory testing was described by a binomial distribution. To address uncertainty in the proportion of cases that had confirmatory testing, a beta distribution was chosen and applied as:

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

Where X is a resource use parameter for confirmatory testing; α is the number of confirmatory tests (deletion detection, duplication detection or Sanger sequencing); β is the total number of tests less the number of confirmatory tests. Since the proportions were provided by the laboratory director, that proportion was applied to the total number of tests to obtain the number of confirmatory tests.

The labour volumes used in sample accessioning, informatics analysis and clinical interpretation and case conferences were recorded in minutes. These labour volumes were spent by genome analysts, the laboratory director, senior bioinformaticians and administrative coordinators as described earlier. For clinical interpretation and reporting labour time, data were collected per trio. Primary variants were classified by a genome analyst and a laboratory director as: i) no/non-diagnostic, ii) diagnostic or partially diagnostic or iii) potentially diagnostic (variant of uncertain significance [VUS]). Secondary variants were classified as: i) no findings, ii) known pathogenic or iii) expected (likely) pathogenic.
On average, it took a genome analyst 201 minutes (SD 105) for GS analysis and 193 minutes (SD 102) for ES analysis per trio. For the laboratory director, the average analysis times were 88 minutes (SD 48) for GS analysis and 81 minutes (SD 40) for ES analysis per trio. Based on empirical data, the analysis time for genome analyst varied between 50 to 780 minutes for GS and 50 to 720 minutes for ES. The laboratory director spent relatively less time, ranging from 29 to 360 minutes for GS and 30 to 330 minutes for ES.

The analysis time data were heavily skewed due to a minority of complex cases requiring significant analysis and interpretation time. Two methods were employed to determine the best way to capture the distribution of skewed analysis time data: (i) bootstrapping with replacement and (ii) lognormal transformation using a random distribution generator. Both parametric and non-parametric bootstrapping were tested for resampling. Since non-parametric bootstrapping does not rely on assumptions of the distribution, it resulted in higher variability in bootstrapped data compared to the actual data in this study. This can lead to discrepancies in estimates because of the increased variance. Parametric bootstrapping assuming a lognormal distribution slightly improved accuracy by reducing the

standard error; however, the standard deviation (SD) was still notably different from the sample SD (P < 5%, 95% CI).

With a smaller sample size than the Ontario RD population, bootstrapping the analysis time is likely to introduce more bias relative to the true populations. Also, the analysis time spent by the genome analyst and laboratory director has decreased over time due to improvements in efficiency (personal communication, laboratory director). As such, a log-normal distribution was applied using the sample mean and SD as follows:

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

Where X is labour time in minutes for sample accessioning, informatics analysis, clinical interpretation or case conferences; μ represents the mean of log-transformed minimum and maximum values and σ^2 is estimated by $\frac{\log(\max(t) - \mu}{F_X^{-1}(k)}$ where $F_X^{-1}(k)$ is the inverse of the cumulated distribution function, and *k* was set to 0.95 to obtain the value of *X* at the 95th percentile of the distribution; and *t* represents the labour time spent in each workflow step. A 95% confidence level was selected for $F_X^{-1}(k)$ to ensure that the mean of *X* closely approximated the actual mean for each variable.

The distributions of inputs were assumed to be independent. To propagate variation in the model, 5,000 values were drawn from each input's range and distribution using Monte Carlo simulations. Inputs without variation, i.e., inputs in certain workflow steps with known fixed resource use, were duplicated in each simulation. As indicators of precision, 95% CIs were determined for each input by analyzing the distribution of each simulated value. These intervals were calculated by taking the 2.5th and 97.5th percentiles of the simulated values measuring uncertainty around each input.

Cost results were summarized as cost per trio for both GS and ES by cost category. The cost per trio by major cost category was summarized graphically for GS and ES. Sequencing costs of GSO during the two-year pilot period were summarized.

2.9 Cost-consequence analysis

The RCT included calculation of diagnostic yield as a measure of test clinical utility, defined as the proportion of the test population in whom a diagnostic or partially diagnostic finding was detected in the index case. Determination of which variants were considered diagnostic or partially diagnostic was based on the list of pathogenic variants current at the time of sequencing. A positive diagnostic finding was defined as detecting pathogenic/likely pathogenic variant(s) that provided a complete explanation of phenotype. A partially diagnostic finding was defined as a pathogenic/likely pathogenic variant(s) that provided a partial explanation of phenotype. A potentially diagnostic finding was the presence of variant(s) of unknown significance that could provide a complete explanation of phenotype or a pathogenic/likely pathogenic variant in a recessive gene without a second variant identified. A CCA was conducted to ascertain the incremental costs for GS compared to ES and the incremental diagnostic yield for GS compared to ES. In the event that GS demonstrated greater costs with improved diagnostic yield compared to ES, an incremental cost-effectiveness ratio could be calculated. Mean total and major category trio costs with 95% CIs were calculated for GS and ES across distinct diagnostic categories: (1) positive diagnostic findings (diagnostic plus partially diagnostic), (2) potentially diagnostic (VUS) and (3) not diagnostic (no or non-diagnostic).

Chi-square tests were used to assess whether the difference in proportions of a positive diagnosis as well as potentially diagnostic findings between GS and ES were statistically significant.

2.10 Ethics

Individuals who underwent GWS as part of this study provided clinical consent for GWS and research consent through Clinical Trials Ontario [CTO-1577], the provincial body responsible for approving clinical trials and observational studies that involve two or more academic or health care institutions in Ontario.

3. Results

3.1 Demographics

Over a two-year period from April 2021 to March 2023, 324 trios were randomized to GS and 329 trios were randomized to ES. Sample characteristics are presented in Table 5.

A large majority of probands (>75%) were under one year of age in both groups at symptom onset. More than two-thirds experienced syndromic ID or DD with or without developmental or functional impairment and multi-system involvement. More than 50% of patients were of non-white ethnicity.

Table 5. GSO randomized controlled trial sample characteristics

Patient Characteristics	GS (n=324)		ES (n=329)	
Age group at onset of symptoms	n	%	n	%
Prenatal	84	25.9	69	21.0
Birth (<12mo)	166	51.2	190	57.8
Childhood (1-10y)	69	21.3	67	20.4
Other	5	1.5	3	0.9
Sex at birth				
Male	186	57.4	184	55.9
Female	136	42.6	145	44.1
Clinical features				
Syndromic ID/DD	217	67.0	217	66.0
Multiple congenital anomalies without ID/DD	33	10.2	34	10.3
Multisystem disorder without ID/DD	64	19.8	65	19.8
Single system (organ) disorder without ID/DD	3	0.9	6	1.8
(Moderate-severe) Isolated ID/DD	7	2.2	7	2.1
Ethnicity				
Non-white	171	52.8	178	54.1
White	153	47.2	151	45.9

Abbreviations: ES, exome sequencing; GS, genome sequencing; DD, developmental disability; ID, intellectual disability

3.2 Costs per trio

The mean costs of GS and ES per estimated using micro-costing models are shown in Table 6. The total estimated cost per trio was CAD 4364.02 (95% CI 3984.94, 5013.67) for GS and CAD 2888.79 (95% CI 2567.72, 3492.72) for ES. The major cost components for GS trio were reagents, software and labour, while for ES, the major cost categories were reagents, consumables and labour. The smallest cost component category for both sequencing strategies was shipping and ordering.

Labour cost per trio for both GS and ES were similar as labour tasks were performed on the same sequencing platform. The differences in cost of reagents between GS and ES is explained by the volume of genetic material undergoing sequencing. Since GS targets the entire nucleotide region of an individual's DNA, it requires more reagents compared to ES, which focuses only on the protein-coding regions. The cost of consumables was higher for ES in comparison with GS due to its targeted nature. ES focused solely on the protein-coding region (1-2% of the genome) which demanded specialized, sometimes labour-intensive procedures for capturing, amplifying and sequencing these discontinuous regions. Customized capture probes and enrichment kits added complexity and expense, which resulted in a higher cost of consumables during library preparation. The cost of equipment was higher for ES compared to GS because of how the NovaSeq 6000 sequencer costs were allocated to the two applications. These sequencers are used mainly for clinical ES in the institution, and therefore a higher proportion of the equipment cost was assigned to ES compared to GS. Shipping and ordering costs were identical for GS and ES. The cost of software per GS trio was approximately 2.34 times higher than ES due to larger data volume generated, and the associated computational and storage demands in comparison to ES. GS had lower overhead cost because of the lower equipment cost compared to ES. The incremental cost per trio for GS versus ES was CAD 1475.23 (95% CI 1417.22, 1520.95).

The proportional cost per trio by category for GS and ES is shown in Figures 5 and 6. For GS, the cost of reagents accounted for the highest proportion at 61%, compared to 40% of total cost for ES. The proportional cost per trio for consumables was higher for ES at 20% of the total cost compared to 3% for GS. Software cost accounted for a higher proportion of total cost for GS at 15%, compared to 11% for ES. The lowest cost proportion cost was for shipping and ordering, which represented 1% of the total cost for GS and 2% for ES.

3.3 Total program cost

The total program cost by group is presented in Table 7. The total enrollment of the GSO study was 1472 trios, with 653 from the randomized cohort and 819 from the non-randomized clinical cohort (Table 8). The total program sequencing costs were CAD 1.97 million in the first year (04/2021 to 03/2022) and CAD 2.76 million in the second year (04/2021 to 03/2022). The mean sequencing costs per year for the GSO RCT study was CAD 2.37 million, with a total cost of CAD 4.73 million over a two-year period.

Table 6. Mean cost per trio of GS vs ES by cost category and diagnostic yield

Cost Category	ES	GS
	Mean cost per trio CAD (95% CI)	**Mean cost per trio CAD (95% CI)**
Total	**$2,888.79 (2,567.72, 3,492.72)**	**$4,364.02 (3,984.94, 5,013.67)**
Labour	$603.13 (348.59, 1089.58)	$603.83 (332.67, 1113.54)
Reagents	$976.10 (924.17, 1036.08)	$2,676.99 (2,534.09, 2,825.32)
Consumables	$502.69 (478.29, 526.57)	$130.26 (123.93, 136.40)
Small & Large Equipment	$252.85 (242.45, 263.23)	$75.90 (72.80, 78.99)
Shipping & Ordering	$67.43 (63.06, 71.84)	$67.45 (63.06, 71.84)
Software	$271.43 (257.10, 285.46)	$635.49 (593.79, 676.19)
Overhead	$215.15 (156.16, 329.22)	$174.09 (110.56, 292.90)

Abbreviations: CAD, Canadian dollar; CI, confidence interval; ES, exome sequencing; GS, genome sequencing

Figure 5. Proportion of total cost by category for genome sequencing

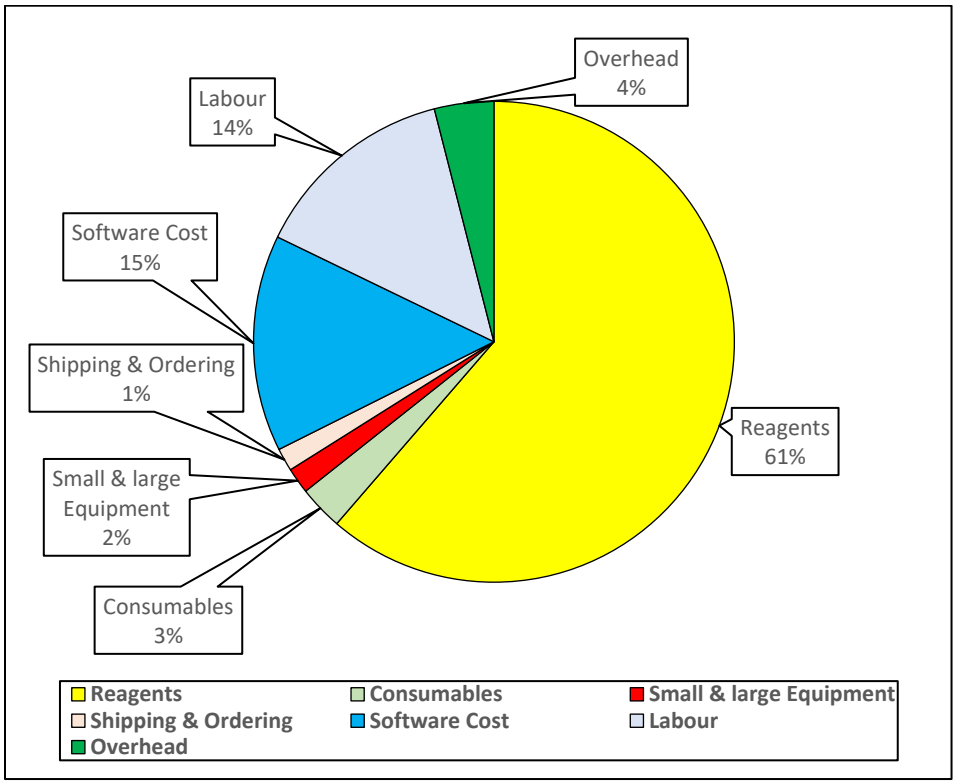


Figure 6. Proportion of total cost by category for exome sequencing

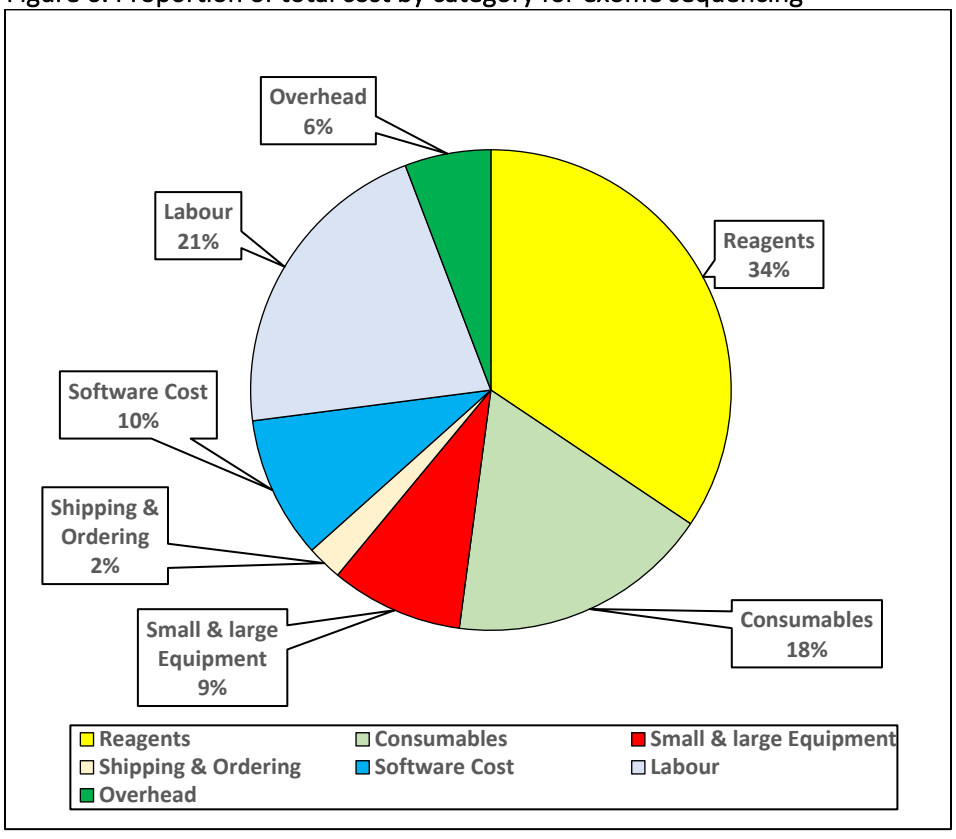


Table 7. Genome-wide sequencing program enrollment and costs by study period

Cohort	Number of Trios Enrolled			Program Cost (CAD)			
	04/2021 to 03/2022	04/2022 to 03/2023	Total	04/2021 to-03/2022	04/2022 to 03/2023	Mean/year	Total
Clinical ES	469	350	819	1,354,843	1,011,077	1,182,960	2,365,919
Randomized GS	84	240	324	366,578	1,047,365	706,971	1,413,942
Randomized ES	86	243	329	248,436	701,976	475,206	950,412
Total	**639**	**833**	**1472**	**1,969,856**	**2,760,417**	**2,365,137**	**4,730,273**

Abbreviations: CAD, Canadian dollar; ES, exome sequencing; GS, genome sequencing

3.4 Diagnostic yield

The diagnostic yield, defined as one or more diagnostic or partially diagnostic findings in a trio, was 32.7% (n=106) for GS compared to 35.9% for ES (n=118). The proportion difference was 0.032 (95% CI: -0.041, 0.104, p-value 0.397).

Cost per trio varied across diagnostic categories for both GS and ES. Diagnostic categories were divided into (1) diagnostic and partially diagnostic, (2) potentially diagnostic (including VUS), and (3) no/non-diagnostic. Table 4 and Table 5 exhibit the cost per trio by each diagnostic category. For both GS and ES, the highest cost per trio was observed in the potentially diagnostic patient group, while the lowest cost per trio was found in the no/non-diagnostic group. The higher cost in the potentially diagnostic/VUS group was due to a greater time spent by the genome analyst for these cases.

While diagnostic yields were similar for GS and ES, the rate of potentially diagnostic/VUS findings was higher for GS (25.3%) compared to ES (20.0%). The Chi-squared tests revealed that there were no statistically significant differences between GS and ES in diagnostic yield or rates of potentially diagnostic findings for this patient population.

Table 8. Cost per genome trio by cost and diagnostic category

Cost category	Mean Cost per Trio (95% CI)		
	Diagnostic + partially diagnostic (n=106)	**Potentially diagnostic (n=82)**	**No + non-diagnostic (n=136)**
Total	**$4,350.79 (4015.73, 4863.44)**	**$4,382.64 (4028.28, 4940.1)**	**$4,325.78 (3966.14, 4953.66)**
Labour	$593.11 (357.66, 983.11)	$618.93 (365.28, 1045.07)	$572.81 (319.13, 1062.08)
Reagents	$2,676.99 (2534.09, 2825.32)	$2,676.99 (2534.09, 2825.32)	$2,676.99 (2534.09, 2825.32)
Consumables	$130.26 (123.93, 136.4)	$130.26 (123.93, 136.4)	$130.26 (123.93, 136.4)
Equipment & Service contract	$75.90 (72.8, 78.99)	$75.90 (72.8, 78.99)	$75.90 (72.8, 78.99)
Shipping & Ordering	$67.45 (63.06, 71.84)	$67.45 (63.06, 71.84)	$67.45 (63.06, 71.84)
Software	$635.49 (593.79, 676.19)	$635.49 (593.79, 676.19)	$635.49 (593.79, 676.19)
Overhead	$171.59 (116.68, 262.61)	$177.61 (118.48, 276.67)	$166.87 (107.54, 280.56)
Diagnostic Yield	106/324 32.7%	82/324 25.3%	136/324 42.0%

Total and category mean costs derived from 5,000 simulations. All costs in 2022/23 CAD.
CAD, Canadian dollar; CI, confidence interval.

Table 9. Cost per exome trio by cost and diagnostic category

Cost Category	Mean Cost per Trio (95% CI)		
	Diagnostic + partially diagnostic (n=118)	**Potentially diagnostic (n=66)**	**No + non-diagnostic (n=145)**
Total	**$2,913.45** (2588.45, 3507.18)	**$2,899.18** (2623.31, 3343.41)	**$2,899.53** (2546.18, 3414.77)
Labour	$623.13 (364.97, 1099.43)	$611.56 (393.97, 964.87)	$563.18 (333.83, 1020.4)
Reagents	$976.10 (924.17, 1036.08)	$976.10 (924.17, 1036.08)	$976.10 (924.17, 1036.08)
Consumables	$502.69 (478.29, 526.57)	$502.69 (478.29, 526.57)	$502.69 (478.29, 526.57)
Equipment & Service contract	$252.45 (242.45, 263.23)	$252.45 (242.45, 263.23)	$252.45 (242.45, 263.23)
Shipping & Ordering	$67.43 (63.06, 71.84)	$67.43 (63.06, 71.84)	$67.43 (63.06, 71.84)
Software	$271.43 (257.10, 285.46)	$271.43 (257.10, 285.46)	$271.43 (257.10, 285.46)
Overhead	$219.81 (159.78, 332.06)	$217.12 (166.56, 299.69)	$205.85 (152.32, 312.6)
Diagnostic Yield	118/329 35.9%	66/329 20.1%	145/329 44.1%

Total and category mean costs derived from 5,000 simulations. All costs in 2022/23 CAD.
CAD, Canadian dollar; CI, confidence interval.

4. Discussion

This paper precisely estimated the cost per trio of GS and ES for children with RDs. The bottom-up micro-costing approach used quantified individual resources and activities in the sequencing workflow, which ensured high accuracy and comprehensiveness in cost estimation. The analysis also promoted transparency and reproducibility, facilitating easier comparison of costs across different studies and healthcare settings.

This analysis found the cost of GS per trio to be 1.5 times higher than that of ES per trio. Reagents were the primary cost component, accounting for 61% of total expenditures for GS and 34% for ES. Software and labour components were identified as the second and third highest categories for GS, at 15% and 14%, respectively. In contrast, labour and consumables ranked as the second and third most substantial cost components for ES at 21% and 18%, respectively. The cost of software, including analysis, storage and data management accounted for 15% of total costs for GS and 10% of total costs for ES. The lowest contributor to costs for both sequencing strategies was the cost of shipping and ordering, with 1% of the total cost for GS and 2% for ES.

The proportion of total cost for small and large equipment was higher for ES compared to GS. The higher value for ES compared to GS was due to the greater use of sequencing equipment for clinical ES compared to GS across all indications in the institution, including high volume panels on exome backbones and other exomes at the time of the study. This resulted in a larger fraction of equipment sample cost allocated to ES compared to GS. As GS grows as a proportion of all sequencing samples, the equipment costs for ES and GS will converge, i.e. ES equipment costs will decrease while GS equipment costs will increase. However, since equipment cost remained less than 9% of total costs, this convergence is expected to have a minimal impact on the incremental cost of GS.

The present analysis resulted in a lower cost per trio in GS for RD diagnosis at CAD 4364.02 (95% CI 3984.94, 5013.67) on the NovaSeq 6000 compared to our previous report of CAD 6556.00 (95% CI 6277.50, 6832.00) using the HiSeq X^{TM} platform for ASD and DD/ID populations (16-18). The lower present cost was mainly due to reduced costs for reagents and consumables supply (CAD 2807.25 vs CAD 4099.90) and in software (CAD 635.49 vs CAD 1258.3), reflecting a greater present use of automation. Our previous analysis revealed a per sample cost for ES for ASD and DD/ID populations of CAD 1960 (95% CI 1899, 2020) on HiSeq® 2500 platform and CAD 1981 (85% CI 1909, 2054) on NextSeq®

550 platform (16, 18). Since our present analysis focused on cost per trio, it cannot be directly compared to previous results.

In the diagnosis of RD, a previously published systematic review compared findings from studies that used the micro-costing approach to calculate the cost of GS (33). Estimates from 2016 to 2018 indicated the cost per GS trio ranged from USD 5272 (CAD 6854) to USD 9706 (CAD 12,618) (33). ES was predominantly used in singletons where the cost per patient ranged from USD 993 (CAD 1291) to USD 3792 (CAD 4930) (33). For patients with neurodevelopmental disorders of unknown genetic etiology, the ES cost varied from USD 1678 (CAD 2181) per singleton to USD 3388 (CAD 4404) per patient in clinical settings (41). The highest cost for a GS trio was reported by Schwarze *et al*. [2020] at USD 9706 (CAD 12,618) for RD, including developmental, neurologic, immunologic, cardiovascular and musculoskeletal patients on the Illumina HiSeq 4000 platform (42). GS was also used for singleton diagnosis in RD, with the lowest cost estimated by van Nimwegen *et al*. [2016] at USD 2094 (CAD 2722) using Illumina HiSeq X Five sequencing for germline variants (43). Our results showed the lowest cost per trio for GS compared to previous studies on RDs.

The cost of sequencing varies depending on the indication and patient population. Researchers have examined the cost of GS using micro-costing approaches in tumor sequencing for cancer diagnosis. Singleton testing was the primary application for both GS and ES in cancer patients (33). The cost per patient for GS ranged from USD 2236 (CAD 2907) to USD 9418 (CAD 12243) (33). The lowest cost per patient for GS was reported by Gordon *et al*. [2018] for mesothelioma using the BGISEQ-500 sequencing machine (44), compared to the highest cost reported by Schwarze *et al*. [2018] for breast, colorectal, prostate, and endometrial cancer diagnosis using Illumina HiSeq 4000 platform (42, 45). The cost per patient for ES in cancer ranged between USD 716 (CAD 931) and USD 4817 (CAD 6262) (33). For ES, the lowest cost was reported by Gordon *et al*. [2018] for diagnosing melanoma using Illumina NovaSeq 5000 (44) and the highest cost was examined at USD 4817 (CAD 6262) per patient in solid tumor diagnosis by Bayle *et al.* [2015] using Illumina HiSeq 2000 series (46). In cancer diagnosis, while the cost per patient for GS was higher than ES according to previous studies, the cost of GS appears to have decreased as technology has advanced (33).

Previous studies that used a micro-costing approach also stratified cost by categories. Consumables accounted for 2% (44) to 78% (47) of the total sequencing cost per singleton using GS in cancer diagnosis

(33). In RD diagnosis, previous studies showed that informatics and equipment costs together constituted a substantial proportion of the total cost. For GS trios, these combined costs accounted for 22% to 27% of the total (19), while for ES in singleton patients, they represented 20% to 53% of the total cost (41). Likely due to technological advancement, the separate proportional costs of informatics and equipment were lower in the present study, accounting for 15% and 2% for GS, respectively, and 10% and 9% for ES, respectively. Moreover, equipment actual use rates were reported to vary from 50% (41) to 90% (47) of total capacity in previous studies which complicates comparisons of equipment costs. Due to the variety of study settings, overhead cost was not measured consistently. Studies either precisely measured overhead costs as a proportion of the institution's annual sequencing volume (16, 18, 20), or considered sequencing platform utilization rates, along with fluctuations in the unit costs of supplies and the useful life of equipment (46).

These wide variations in estimates highlight the range of approaches used in microcosting and cost allocation. Comparisons with previous estimates are complicated by use of different platforms and workflows. Older platforms are quickly being replaced by newer more costly equipment and greater automation is contributing to improved efficiency and reduced labour time. Microcosting underscores the challenges in general predictions about the future cost of sequencing. Whereas equipment costs and wages will continue to rise, increased competition may exert a downward pressure on the costs of reagents and consumables which have consistently contributed to a large proportion of total cost. Further, centralizing sequencing and maximizing volumes to reach capacity will improve efficiency and contribute to lower overall costs. Additional steps such as re-analysis for select cases may increase costs but better selection of specific genetic tests based on phenotypic presentation may reduce overall costs in a patient's diagnostic pathway.

Whether or not to prioritize GS in the diagnostic pathway for RD and other pediatric conditions is not yet clear. The present study observed diagnostic yields of 35.9% for trio ES compared to 32.7% for trio GS in individuals with a prior negative or inconclusive CMA or single gene/gene panel test. These results are consistent with previous findings for GS or ES in children with rare diseases.10 In a meta-analysis of sequencing studies performed in children with a range of conditions, the pooled diagnostic yield was 0.38 (95% CI 0.35, 0.42) for ES and was 0.37 (95% CI 0.32, 0.42) for GS, based on 41 and 14 studies for ES and GS, respectively (48). Focusing on children with rare and undiagnosed disease, Pandey *et al*. reported pooled estimates of diagnostic yield from three within-cohort comparisons of 0.232 (95% CI

0.185, 0.287) for ES and 0.306 (05% CI 0.186, 0.459) for GS (49). An unbiased comparison of diagnostic yield between trio ES and trio GS can only be achieved in independent samples with randomization. Recognizing the high genetic and phenotypic heterogeneity of the study sample, the randomization was stratified within each site (SickKids and CHEO) by additional clinical criteria expected to potentially confound the comparison of diagnostic yield. These included phenotype and prior genetic testing. The GS yield may have been lower than expected for several reasons. First, the study design required a prior genetic test which would remove from the sample those cases with more easily detectable variants. Second, the genome analysis was focused on the coding region due to the interpretation challenges in analyzing non-coding regions. While non-coding regions can contain important regulatory elements, interpreting variants in these regions is challenging and requires sophisticated informatics tools. Third, validated bioinformatic tools for calling new variant types (repeat expansions and mitochondrial variants) were not available at the time of the study and these are expected to have a positive impact on GS diagnostic yield as will the growing understanding of how these variants contribute to disease. Finally, the marginal impact on yield of re-analysis also requires evaluation. Further exploration of clinical utility in the present sample, including measuring diagnostic yield differences by sub-group, changes over time and results reporting turn-around times was conducted in a companion implementation effectiveness trial (23).

The study possessed several strengths. This was a randomized controlled trial comparing GS trio to ES trio for diagnosing heterogenous patients with rare diseases of suspected genetic origins. Large sample sizes were achieved. Microcosting was employed in all stages in the workflows of ES trio and GS trio to approximate actual opportunity costs rather than using charges. Parameters with uncertain estimates were captured in the probabilistic analysis. By using actual consumption rather than maximum capacity, the present analysis allows for the opportunity cost of underutilized equipment to be integrated into the estimate. The introduction of the NovaSeq6000 allows ES and GS to be run on the same platform. While this provides some standardization in equipment costing, it creates a potential new source of bias if the sample cost does not take into consideration disparate use of the machine for one application vs. another. The present analysis avoided this by using standard allocation methods to account for much greater use of equipment for clinical ES vs. GS. While the estimates in this study pertained to a predominantly pediatric rare disease patient population, the costing model can easily accommodate variations in resource use items and volumes of testing for other patient populations or be adapted by other institutions. Many jurisdictions around the world are facing decisions regarding funding of

sequencing technologies. While the unit prices applied pertain to Ontario, the disaggregated presentation of cost categories, resource use volumes and cost items enhances transparency and reproducibility, facilitating comparison of resource use and costs across different studies as well as adaptation to different regions.

There were several study limitations. First, many of the price parameters in this study were assumed to vary within a 10% range based on invoices, price lists and laboratory personnel expert opinion to reflect unknown potential fluctuations in price and currency. Actual variation may have exceeded this range given pressures on supply chains that vary by product origin. Second, a five-year or eight-year time horizon was chosen for the large equipment used in this study based on the projected shelf-life. In reality, the life cycle for sequencers may be shorter or longer depending on the technological improvement in sequencing hardware and software combined with the frequency of usage and institutional budgets. Reducing the life cycle duration would lead to an increase in cost as a result of a shorter amortization period. Third, this study targeted trios, which included the proband and both biological parents. In clinical practice, singletons, duo, quads and pents may also be sequenced. Fourth, estimates for unit prices and volumes of use in the sequencing workflow are volatile and subject to supply chain fluctuations which can greatly affect cost. Fifth, all parameter estimates reflected utilization volumes and prices in effect at the time of the study and the analysis was conducted amidst rapid technical and analytic advancements and an expanding understanding of clinically relevant variant types. These factors complicate comparisons with past studies and necessitate frequent updates of microcost estimates. Finally, while the study enrolled sufficient samples sizes to produce stable cost estimates, it was not sufficiently powered to detect significance in an incremental diagnostic yield <10%.

While this study focused only on the costs and consequences of sequencing, in reality, patients with RDs typically undergo multiple serial tests to achieve a diagnosis. Economic evaluations that compare entire diagnostic pathways in terms of costs and health outcomes to determine the optimal positioning of ES and GS as first, second or third-tier strategies in conjunction with other tests for specific patient populations are needed. To date, most economic evaluations have examined costs of various sequencing approaches as a function of diagnostic yield or other proximal outcomes such as clinical management changes (22, 45, 50, 51). Studies are needed that consider measures of effectiveness that include impacts on patients' health and quality-of-life over the long-term. Until such data are available, economic evaluations comparing sequencing approaches have used decision analysis to model

hypothetical patient cohorts where parameter uncertainty required assumptions about long-term costs and consequences (52, 53). Moreover, methods to measure and quantify patient preferences and personal utility for sequencing technologies that return a result that may not be medically actionable are essential components of HTA used to inform funding decisions (54, 55).

5. Conclusion

GS and ES have been increasingly used as promising diagnostic strategies for RD patient populations (30). In this study, the average total cost per trio was CAD 4364.02 (95% CI 3984.94, 5013.67) for GS and CAD 2888.79 (95% CI 2567.72, 3492.72) for ES using the Illumina NovaSeq 6000 platform. Reagents was the largest cost component for both GS and ES. GS was CAD 1475.23 (95% CI 1417.22, 1520.95) more costly than ES.

In conclusion, GS and ES are promising strategies for rare disease diagnostics. As GS was more costly than ES and similarly effective, its application should be targeted to populations where it can achieve greatest value, including reducing the time to diagnosis. This may be particularly useful in neonatal care where a rapid diagnosis and early treatment may result in improved patient management and outcomes (32). Understanding where GS may be most advantageous based on patient phenotypic profile and incremental cost-effectiveness is an area of active investigation, including the use of machine learning to analyse phenotype data to prioritize sequencing in populations where a diagnosis is mostly likely (56, 57) . As the technology continues to evolve, microcosting and assessment of diagnostic yield must be repeated to furnish current estimates. Such estimates can be used in comparative cost-effectiveness analyses that compare not only specific tests, but entire diagnostic pathways to inform recommendations for funding as well as optimal positioning of ES and GS within assessment pathways in rare disease and other patient populations.

References

1. Richter T, Nestler-Parr S, Babela R, *et al*. Rare Disease Terminology and Definitions—A Systematic Global Review: Report of the ISPOR Rare Disease Special Interest Group. Value in Health. 2015; 18: 906-14. 10.1016/j.jval.2015.05.008

2. Boycott KM, Vanstone MR, Bulman DE, *et al*. Rare-disease genetics in the era of next-generation sequencing: Discovery to translation. Nature Reviews Genetics. 2013; 14: 681-91. 10.1038/nrg3555

3. Posey JE. Genome sequencing and implications for rare disorders. Orphanet Journal of Rare Diseases. 2019; 14: 153-63. 10.1186/s13023-019-1127-0

4. Soden SE, Saunders CJ, Willig LK, *et al*. Effectiveness of exome and genome sequencing guided by acuity of illness for diagnosis of neurodevelopmental disorders. Science Translational Medicine. 2014; 6: 265ra168-265ra168. 10.1126/scitranslmed.3010076

5. Antonarakis SE, Krawczak M, Cooper DN. The Nature and Mechanisms of Human Gene Mutation. In: Valle DL, Antonarakis S, Ballabio A, *et al*., eds., The Online Metabolic and Molecular Bases of Inherited Disease. New York, NY: McGraw-Hill Education, 2019.

6. Gross AM, Ajay SS, Rajan V, *et al*. Copy-number variants in clinical genome sequencing: deployment and interpretation for rare and undiagnosed disease. Genetics in Medicine. 2019; 21: 1121-30. 10.1038/s41436-018-0295-y

7. Hauser NS, Solomon BD, Vilboux T, *et al*. Experience with genomic sequencing in pediatric patients with congenital cardiac defects in a large community hospital. Molecular Genetics & Genomic Medicine. 2018; 6: 200-12. 10.1002/mgg3.357

8. Yavarna T, Al-Dewik N, Al-Mureikhi M, *et al*. High diagnostic yield of clinical exome sequencing in Middle Eastern patients with Mendelian disorders. Human Genetics. 2015; 134: 967-80. 10.1007/s00439-015-1575-0

9. Green RC, Berg JS, Grody WW, *et al*. ACMG Recommendations for Reporting of Incidental Findings in Clinical Exome and Genome Sequencing. Genetics in Medicine. 2013; 15: 565-74. 10.1038/gim.2013.73

10. Ahmed S, Zhou Z, Zhou J, *et al*. Pharmacogenomics of Drug Metabolizing Enzymes and Transporters: Relevance to Precision Medicine. Genomics Proteomics Bioinformatics. 2016; 14: 298-313. 10.1016/j.gpb.2016.03.008

11. Brian JA, Zwaigenbaum L, Ip A. Standards of diagnostic assessment for autism spectrum disorder. Paediatrics & Child Health. 2019; 24: 444-51. 10.1093/pch/pxz117

12. Ontario Health Study. [Available from: https://www.ontariohealthstudy.ca/for-researchers/research-approved-to-access-ohs-data/#:~:text=Project%20Summary%3A%20The%20Ontario%20Health,diabetes%2C%20and%20other%20chronic%20diseases.]

13. Ontario Institue for Cancer Research. [Available from: https://oicr.on.ca/news/all/]

14. Advancing Genomic Medicine through Research and Education - McLaughlin Centre. [Available from: https://utstat.utoronto.ca/sun/McLaughlin-Centre-Biennial-Report-2011-2013.pdf]

15. Technology Assessment at SickKids (TASK). [Available from: https://lab.research.sickkids.ca/task/]

16. Tsiplova K, Zur RM, Ungar WJ, *et al*. A Microcosting and Cost-consequence Analysis of Genomic Testing Strategies in Autism Spectrum Disorder. 2016. [Available from: http://lab.research.sickkids.ca/task/reports-theses/]

17. Tsiplova K, Zur RM, Marshall CR, *et al*. A microcosting and cost–consequence analysis of clinical genomic testing strategies in autism spectrum disorder. Genetics in Medicine. 2017; 19: 1268-75. 10.1038/gim.2017.47

18. Jegathisawaran J, Tsiplova K, Marshall CR, *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 https://lab.research.sickkids.ca/task/wp-content/uploads/sites/66/2021/01/Microcosting-REPORT-NO-2018-01_FINAL.pdf]

19. Jegathisawaran J, Tsiplova K, Hayeems RZ, *et al*. Trio genome sequencing for developmental delay and pediatric heart conditions: A comparative microcost analysis. Genetics in Medicine. 2022; 24: 1027-36. 10.1016/j.gim.2022.01.020

20. Jegathisawaran J, Tsiplova K, Hayeems RZ, *et al*. Microcosting of whole genome sequencing (WGS) of trios in a heterogeneous pediatric cardiac population. Technology at Sick Kids (TASK): Technology at Sick Kids (TASK), 2020. https://lab.research.sickkids.ca/task/wp-content/uploads/sites/66/2021/01/2020-01-CGC-Full-Report.pdf

21. OntarioHealth. Genome-Wide Sequencing for Unexplained Developmental Disabilities or Multiple Congenital Anomalies: A Health Technology Assessment. Ontario Health Technology Assessment Series. 2020; 20: 1-178. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7080457/

22. Li C, Vandersluis S, Holubowich C, *et al*. Cost-effectiveness of genome-wide sequencing for unexplained developmental disabilities and multiple congenital anomalies. Genetics in Medicine. 2021; 23: 451-60. 10.1038/s41436-020-01012-w

23. Hayeems RZ, Marshall CR, Gillespie MK, *et al*. Comparing genome sequencing technologies to improve rare disease diagnostics: a protocol for the evaluation of a pilot project, Genome-wide Sequencing Ontario. CMAJ Open. 2022; 10: E460. 10.9778/cmajo.20210272

24. Genome-wide Sequencing Ontario. [Available from: https://gsontario.ca/]

25. Wright CF, FitzPatrick DR, Firth HV. Paediatric genomics: diagnosing rare disease in children. Nature Reviews Genetics. 2018; 19: 253-68. 10.1038/nrg.2017.116

26. Beaulieu Chandree L, Majewski J, Schwartzentruber J, *et al*. FORGE Canada Consortium: Outcomes of a 2-Year National Rare-Disease Gene-Discovery Project. The American Journal of Human Genetics. 2014; 94: 809-17. 10.1016/j.ajhg.2014.05.003

27. Canadian Organization for Rare Disorders [Available from: https://www.raredisorders.ca/about-cord/#:~:text=One%20in%2012%20Canadians%20has,that%20severely%20impacts%20their%20lives]

28. Black N, Martineau F, Manacorda T. Diagnostic odyssey for rare diseases: exploration of potential indicators. 2015. [Available from https://www.researchgate.net/publication/291413094_Diagnostic_odyssey_for_rare_diseases_exploration_of_potential_indicators]

29. Sawyer SL, Hartley T, Dyment DA, *et al*. Utility of whole-exome sequencing for those near the end of the diagnostic odyssey: time to address gaps in care. Clinical Genetics. 2016; 89: 275-84. 10.1111/cge.12654

30. Clark MM, Stark Z, Farnaes L, *et al*. Meta-analysis of the diagnostic and clinical utility of genome and exome sequencing and chromosomal microarray in children with suspected genetic diseases. Nature Partner Journals Genomic Medicine. 2018; 3: 16. 10.1038/s41525-018-0053-8

31. Oei K, Hayeems RZ, Ungar WJ, *et al*. Genetic testing among children in a complex care program. Children. 2017; 4: 1-11. 10.3390/children4050042

32. Kingsmore SF, Cakici JA, Clark MM, *et al*. A Randomized, Controlled Trial of the Analytic and Diagnostic Performance of Singleton and Trio, Rapid Genome and Exome Sequencing in Ill Infants. The American Journal of Human Genetics. 2019; 105: 719-33. 10.1016/j.ajhg.2019.08.009

33. Santos Gonzalez F, Mordaunt D, Stark Z, *et al*. Microcosting diagnostic genomic sequencing: A systematic review. Genetics in Medicine. 2023; 25: 100829. 10.1016/j.gim.2023.100829

34. Yuen T, Carter MT, Szatmari P, *et al*. Cost-effectiveness of genome and exome sequencing in children diagnosed with autism spectrum disorder. Applied Health Economics and Health Policy. 2018; 16: 481-93. 10.1007/s40258-018-0390-x

35. Boycott K, Hartley T, Adam S, *et al*. The clinical application of genome-wide sequencing for monogenic diseases in Canada: Position Statement of the Canadian College of Medical Geneticists. Journal of Medical Genetics. 2015; 52: 431-7. 10.1136/jmedgenet-2015-103144

36. Bank of Canada exchange rates. [Available from https://www.bankofcanada.ca/rates/exchange/daily-exchange-rates/]

37. Ontario Ministry of Health. Ontario Schedule of Benefits, Physician Services Under the Health Insurance Act 2022. [Available from https://www.ontario.ca/page/ohip-schedule-benefits-and-fees]

38. Canadian Institute for Health Information Management Information System Standards 2022. [Available from https://www.cihi.ca/en/management-information-system-standards]

39. Strom SP, Lee H, Das K, *et al*. Assessing the necessity of confirmatory testing for exome-sequencing results in a clinical molecular diagnostic laboratory. Genetics in Medicine. 2014; 16: 510-15. 10.1038/gim.2013.183

40. Armour JAL, Barton DE, Cockburn DJ, *et al*. The detection of large deletions or duplications in genomic DNA. Human Mutation. 2002; 20: 325-37. 10.1002/humu.10133

41. Sabatini LM, Mathews C, Ptak D, *et al*. Genomic Sequencing Procedure Microcosting Analysis and Health Economic Cost-Impact Analysis: A Report of the Association for Molecular Pathology. The Journal of Molecular Diagnostics. 2016; 18: 319-28. 10.1016/j.jmoldx.2015.11.010

42. Schwarze K, Buchanan J, Fermont JM, *et al*. The complete costs of genome sequencing: a microcosting study in cancer and rare diseases from a single center in the United Kingdom. Genetics in Medicine. 2020; 22: 85-94. 10.1038/s41436-019-0618-7

43. van Nimwegen KJM, van Soest RA, Veltman JA, *et al*. Is the $1000 Genome as Near as We Think? A Cost Analysis of Next-Generation Sequencing. Clinical Chemistry. 2016; 62: 1458-64. 10.1373/clinchem.2016.258632

44. Gordon LG, White NM, Elliott TM, *et al*. Estimating the costs of genomic sequencing in cancer control. BMC Health Services Research. 2020; 20: 492. 10.1186/s12913-020-05318-y

45. Schwarze K, Buchanan J, Taylor JC, *et al*. Are whole-exome and whole-genome sequencing approaches cost-effective? A systematic review of the literature. Genetics in Medicine. 2018; 20: 1122-30. 10.1038/gim.2017.247

46. Bayle A, Droin N, Besse B, *et al*. Whole exome sequencing in molecular diagnostics of cancer decreases over time: evidence from a cost analysis in the French setting. The European Journal of Health Economics. 2021; 22: 855-64. 10.1007/s10198-021-01293-1

47. Pasmans CTB, Tops BBJ, Steeghs EMP, *et al*. Micro-costing diagnostics in oncology: from single-gene testing to whole- genome sequencing. Expert Review of Pharmacoeconomics & Outcomes Research. 2021; 21: 413-14. 10.1080/14737167.2021.1917385

48. Chung CCY, Hue SPY, Ng NYT, *et al*. Meta-analysis of the diagnostic and clinical utility of exome and genome sequencing in pediatric and adult patients with rare diseases across diverse populations. Genetics in Medicine. 2023; 25: 100896. 10.1016/j.gim.2023.100896

49. Pandey R, Brennan NF, Trachana K, *et al*. A meta-analysis of diagnostic yield and clinical utility of genome and exome sequencing in pediatric rare and undiagnosed genetic diseases. Genetics in Medicine. 2025; 27: 101398. 10.1016/j.gim.2025.101398

50. Smith HS, Swint JM, Lalani SR, *et al*. Clinical Application of Genome and Exome Sequencing as a Diagnostic Tool for Pediatric Patients: a Scoping Review of the Literature. Genetics in Medicine. 2019; 21: 3-16. 10.1038/s41436-018-0024-6

51. Tan TY, Dillon OJ, Stark Z, *et al*. Diagnostic Impact and Cost-effectiveness of Whole-Exome Sequencing for Ambulant Children With Suspected Monogenic Conditions. JAMA Pediatrics. 2017; 171: 855-62. 10.1001/jamapediatrics.2017.1755

52. Lavelle TA, Feng X, Keisler M, *et al*. Cost-effectiveness of exome and genome sequencing for children with rare and undiagnosed conditions. Genetics in Medicine. 2022; 24: 1349-61. 10.1016/j.gim.2022.03.005

53. Stark Z, Schofield D, Martyn M, *et al*. Does genomic sequencing early in the diagnostic trajectory make a difference? A follow-up study of clinical outcomes and cost-effectiveness. Genetics in Medicine. 2019; 21: 173-80. 10.1038/s41436-018-0006-8

54. Weymann D, Dragojlovic N, Pollard S, *et al*. Allocating healthcare resources to genomic testing in Canada: Latest evidence and current challenges. Journal of Community Genetics. 2022; 13: 467-76. 10.1007/s12687-019-00428-5

55. Payne K, Eden M, Davison N, *et al*. Toward health technology assessment of whole-genome sequencing diagnostic tests: Challenges and solutions. Personalized Medicine. 2017; 14: 235-47. 10.2217/pme-2016-0089

56. Dingemans AJM, Hinne M, Jansen S, *et al*. Phenotype based prediction of exome sequencing outcome using machine learning for neurodevelopmental disorders. Genetics in Medicine. 2022; 24: 645-53. 10.1016/j.gim.2021.10.019

57. Peterson B, Hernandez EJ, Hobbs C, *et al*. Automated prioritization of sick newborns for whole genome sequencing using clinical natural language processing and machine learning. Genome Medicine. 2023; 15: 18. 10.1186/s13073-023-01166-7