The impact of clinical genome sequencing in a global population of patients with suspected rare genetic disease

Ryan Taft (rtaft@illumina.com)
Illumina

Erin Thorpe
Illumina, Inc.

Taylor Williams
Genetic and Genomic Services PBC

Chad Shaw
Genetic and Genomic Services PBC; Department of Molecular and Human Genetics, Baylor College of Medicine; Department of Statistics, Rice University

Evgenii Chekalin
Illumina

Julia Ortega
C2N Diagnostics

Keisha Robinson
Illumina

Jason Button
Illumina

Marilyn Jones
Rady Children's Hospital; University of California San Diego

Miguel del Campo
Rady Children's Hospital; University of California San Diego

Donald Basel
Children's Wisconsin, Medical College of Wisconsin

Julie McCarrier
Children's Wisconsin, Medical College of Wisconsin

Laura Davis Keppen
Sanford USD Medical Center

Erin Royer
Sanford Children's Specialty Clinics at Sanford Health; USD Sanford School of Medicine

Romina Foster-Bonds
Rare Genomics Institute

Milagros Duenas-Roque

Instituto de Evaluación de Tecnologías en Salud e Investigación - IETSI, EsSalud

Nora Urraca

University of Tennessee Health Science Center; Le Bonheur Children’s Hospital

Kerri Bosfield

University of Tennessee Health Science Center; Le Bonheur Children’s Hospital

Chester Brown

University of Tennessee Health Science Center; Le Bonheur Children’s Hospital

Holly Lydigsen

University of Tennessee Health Science Center; Le Bonheur Children’s Hospital

Henry Mroczkowski

University of Tennessee Health Science Center; Le Bonheur Children’s Hospital

Jewell Ward

University of Tennessee Health Science Center; Le Bonheur Children’s Hospital

Fabio Sirchia

Department of Molecular Medicine, University of Pavia; Medical Genetics Unit, IRCCS San Matteo Foundation

Elisa Giorgio

Department of Molecular Medicine, University of Pavia, Pavia, Italy Medical Genetics Unit, IRCCS Mondino Foundation

Keith Vaux

Point Loma Pediatrics

Hildegard Peña Salguero

Padrino Children's Foundation

Aimé Lumaka Zola

Centre de Genetique Humaine, Universite de Kinshasa; Center for Human Genetics, Centre Hospitalier Universitaire, Liège, Belgium

Gerrye Mubungu

Centre de Genetique Humaine, Universite de Kinshasa

Prince Makay

Centre de Genetique Humaine, Universite de Kinshasa

Mamy Ngole

Centre de Genetique Humaine, Universite de Kinshasa https://orcid.org/0000-0002-1041-9257

Prosper Lukusa

Centre de Genetique Humaine, Universite de Kinshasa

Adeline Vanderver

Children's Hospital of Philadelphia

Kayla Muirhead

Division of Neurology, Department of Pediatrics, Children's Hospital of Philadelphia

Omar Sherbini

Division of Neurology, Department of Pediatrics, Children's Hospital of Philadelphia

Melissa Lah
Indiana University School of Medicine
Katelynn Anderson
Indiana University School of Medicine
Jeny Bazalar Montoya
Instituto Nacional de Salud del Niño-San Borja
Richard Rodriguez
Instituto Nacional de Salud del Niño-San Borja
Mario Cornejo Olivas
Neurogenetics Research Center, Instituto Nacional de Ciencias Neurológicas; Carrera de Medicina, Universidad Científica del Sur
Karina Milla-Neyra
Neurogenetics Research Center, Instituto Nacional de Ciencias Neurológicas
Marwan Shinawi
Department of Pediatrics, Division of Genetics and Genomic Medicine, Washington University School of Medicine, St. Louis, MO https://orcid.org/0000-0003-1329-4100
Pilar Magoulas
Texas Children's Hospital
Duncan Henry
UCSF Benioff Children’s Hospitals
Kate Gibson
Canterbury District Health Board
Samuel Wiafe
Rare Disease Ghana Initiative
Parul Jayakar
Nicklaus Children's Hospital
Daria Salyakina
Nicklaus Children's Hospital
Diane Masser-Frye
Rady Children’s Hospital
Arturo Serize
South Miami Hospital
Jorge Perez
South Miami Hospital
Alan Taylor
Al Jalila Genomics Center of Excellence, Al Jalila Children’s Specialty Hospital
Shruti Shenbagam
Al Jalila Genomics Center of Excellence, Al Jalila Children’s Specialty Hospital
Ahmad Abou Tayoun
Al Jalila Children's Specialty Hospital https://orcid.org/0000-0002-9134-1673

Alka Malhotra
Illumina
Maren Bennett
Illumina
Vani Rajan
Illumina
James Avecilla
Illumina
Andrew Warren
Illumina
Max Arseneault
Illumina
Tasha Kalista
Illumina
Ali Crawford
Illumina
Subramanian Ajay
Illumina
Denise Perry
Illumina
John Belmont
Genetic and Genomic Services PBC

Article

Keywords:

Posted Date: August 4th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3117254/v1

Abstract

Clinical genome sequencing (cGS) holds promise as a unified diagnostic testing platform in patients with a suspected rare genetic disease (RGD), however its performance and impact on clinical management in a diverse global population has yet to be investigated. The iHope program established a network of 24 clinical sites in eight countries to provide cGS to individuals with signs or symptoms of a RGD and constrained access to molecular testing. A retrospective, observational analysis of 1,004 individuals who received cGS testing from June 2016 through September 2021 was performed. The cGS diagnostic yield in this diverse cohort (51.8% non-majority European) was 41.4% (416/1004), with patients from sites in low- and middle-income countries (LMIC) 2.6-times more likely to receive a positive test result compared to sites in high-income countries (HIC) (95% CI 1.9–3.4, $p < 0.0001$). Changes in diagnostic evaluation and management were reported in 76.9% and 69.2% of cases, respectively. Comparison of LMIC and HIC patients with positive test results demonstrated that LMIC patients were equally likely to experience a change in DE (OR 6.1, 95% CI 1.1- , $p = 0.05$) and COM (OR 0.9, 95% CI 0.5–1.3, $p = 0.49$), indicating that increased access to cGS may support diagnostic equity and the reduction of global health care disparities.

Main

Globally there are estimated to be at least 250 million individuals affected by rare diseases, the vast majority of which are genetic in origin and undiagnosed.[1] While there is an increasing recognition of the growing burden of noncommunicable diseases in low- and middle-income countries (LMIC),[2] less attention has been paid to patients with one of the more than 4,400 well-described rare monogenic diseases.[1] These disorders have a wide range of presentations inclusive of epilepsy, developmental or intellectual disability, neuromuscular, movement disorders, structural abnormalities, and primary immune deficiencies.[1,3] The majority (~ 70%) present in childhood,[1] and are associated with unfavorable utilization and health outcomes.[4,5] A precision diagnosis provides the basis for care and its absence or delay can result in both missed and inappropriate interventions.[4,5] Recognizing the importance of the availability of molecular testing for precision diagnosis, the World Health Organization Science Council recently stated that it is neither ethically nor scientifically justifiable for a significant delay to exist between the availability of genomic technologies in high-income countries (HIC) and LMIC.[6]

There is increasing evidence for the use of clinical genome sequencing (cGS) as a first-line test for patients with signs or symptoms of a genetic disease in the pediatric outpatient population,[7–11] but there have been few efforts to improve equity of access for LMI individuals.[12,13] The iHope program was established with an aim of providing cGS testing to patients with limited means and reduced access to molecular testing in resource-limited communities. Here we describe cohort of 1,004 patients drawn from 24 global sites with heterogeneous, suspected rare genetic disorders, and report on cGS diagnostic yield and the impact of genomic findings on diagnostic evaluation (DE) and change of management (COM).

Results

The iHope program enabled cGS testing across a network of 24 clinical sites in the United States, Mexico, Peru, the Democratic Republic of Congo, Ghana, Italy, New Zealand and the United Arab Emirates. Participation was limited to individuals with signs or symptoms indicative of a RGD, consistent with current professional guidelines and prior large RGD population investigations, and limited access to molecular testing.[7,9,14,15]

Cohort characteristics

A total of 1,004 individuals from 981 families received cGS testing through the iHope program from June 2016 through September 2021 (Fig. 1 and Supplementary Table 1), thirty-four percent (345/1004) of which were from LMIC sites including Mexico (209), Peru (89), the Democratic Republic of Congo (35), Ghana (9) and individual cases from Romania (1), Brazil (1) and India (1) (Fig. 1). Patients were from diverse ancestral backgrounds, with representation from each of seven superpopulations derived from the 1000 Genomes Project, Human Genome Diversity Project and Simons Genome Diversity Project cohorts (Fig. 2A and Supplementary Figs. 1 and 2). A non-European superpopulation was the highest ancestral contributor in 51.8% (521) of the cohort. The distribution of patient age was similar between LMIC and HIC (median age (y) LMIC 6.6 vs HIC 6.3, 95% CI -0.7-1.9, $p = 0.36$). For sex, the odds ratio male was 1.1 (95% CI 0.8–1.4, $p = 0.69$) (Fig. 2B). Trio or higher-order family structures including both biological parents were available for the majority of cases (713/1004 [71.0%]).

Patient phenotypes were complex with nervous system, skeletal system, and head or neck the most frequently identified Human Phenotype Ontology root ancestor terms in both LMIC and HIC groups (Fig. 2C). Digestive system (LMIC 24% vs HIC 41%), respiratory system (LMIC 7% vs HIC 24%), and cardiovascular system (LMIC 19% vs HIC 34%) root ancestor terms had the largest proportional difference between LMIC and HIC groups. Patients from HIC sites had an average of 2.3 more phenotypic terms submitted per case (95% CI 1.5–3.1, $p < 0.001$) but computed phenotypic information content showed no association with likelihood of a positive test result (point estimate 6.4, 95% CI 1.7–11.6, $p = 0.17$) (Supplementary Fig. 3).

Diagnostic Yield

The total diagnostic yield across the cohort was 41.4% (416/1004), with patients from LMIC sites 2.6-times more likely to receive a positive test result compared to patients from HIC sites (LMIC 195/345 [56.5%] vs HIC 221/659 [33.5%], 95% CI 1.9–3.4, $p < 0.0001$) (Fig. 3A). In an additional 26.1% of patients inconclusive test results were reported, with a higher proportion reported in patients from HIC sites (195/659, 29.5%) compared to LMIC sites (67/354, 19.4%) (Fig. 3A). Although patients from HIC sites were 6.6-times more likely than patients from LMIC sites to have had a least one prior genetic test (95% CI 4.9–8.8, $p < .0001$), there was no observed difference in the likelihood of a positive test result based on prior genetic testing within either the HIC or LMIC cohorts (Supplementary Fig. 4).

Reported variants (1,033) spanned the mutational spectrum, including nuclear genome SNVs (714), small indels (130), CNVs (165), STRs (10), mitochondrial SNVs (9), uniparental disomy (3), and spinal muscular atrophy detected by biallelic absence of the c.840C allele (2) (Methods and Supplementary Fig. 5). LMIC

patients had a greater proportion of copy number variants reported (LMIC 84/377 [22.3%] vs HIC 81/650 [12.5%]), which spanned a larger size range, compared to HIC patients (Supplementary Figs. 5 and 6).

Diagnostic Evaluation and Change of Management

A cGS impact survey was provided to the ordering clinicians for patients who pursued cGS from June 2016 through October 2020 (Fig. 1 and Supplementary Table 2). Clinical GS impact surveys were completed by a clinical provider for 694 cases (694/1004 [69.1%] of the total cohort) (Fig. 1). Survey response rates were comparable between LMIC and HIC sites (LMIC 254/297 [85.5%] vs HIC 440/521 [84.5%]) and were highest for patients with positive test results (LMIC 157/172 [91.3%] vs HIC 162/176 [92.0%]) compared to inconclusive (LMIC 44/57 [77.2%] vs HIC 122/148 [82.4%]) or negative test results (LMIC 53/68 [77.9%] vs HIC 156/197 [79.2.0%]).

To control for the elevated LMIC diagnostic yield, and the potential influence of LMIC or HIC site designation on DE and COM, which is mediated by test result category (Supplementary Fig. 7), comparative analyses of GS test impact were stratified by test result category. Clinical GS results impacted DE in 514 of 668 patients (76.9%) and was more common in patients from LMIC sites regardless of test result category (LMIC 221/253 [87.4%] vs HIC 293/415 [70.1%], OR 2.2, 95% CI 1.3–3.6, $p < 0.0001$) (Fig. 3C and Supplementary Table 3). Across both cohorts, patients with a positive test result were 64.2-times more likely to have an impact to DE (95% CI 32.5-202.8, $p < 0.0001$) (Fig. 3C). Clinical GS findings led to a change in the clinical diagnosis in 326 of 694 patients (46.9%). Positive test results corresponded to the highest rates of change in clinical diagnosis and were comparable between HIC and LMIC sites (HIC 130/163 [79.8%] vs LMIC 122/157 [77.7%], OR 0.9, CI 95% 0.5–1.5, $p = 0.50$) (Fig. 3B).

Overall, clinical GS resulted in COM in 285 of 694 patients (41.1%), inclusive of specialty referrals, imaging and testing, therapeutic interventions and palliative care (Fig. 4). When comparing across GNI site designations, patients with a positive test result from LMIC sites were equally likely to experience a COM compared to patients with a positive test result from HIC sites (OR 0.9, 95% CI 0.5–1.4, $p = 0.49$), and no statistically significant differences were observed for inconclusive or negative test results (Supplementary Table 3). When genetic counseling and avoidance of additional testing were also considered, a total of 480 patients (69.2%) experienced COM (LMIC 198/254 [77.9%] vs HIC 282/440 [64.1%]) (Fig. 4). Overall, genetic counseling was the most frequently endorsed change in management category (377/602 [62.6%]), followed by referrals, imaging and testing (265/617 [42.9%]), avoidance of additional testing (198/668 [29.6%]), therapeutics (72/575 [12.3%]) and palliative care (23/694 [3.3%]) (Fig. 4, Supplementary Fig. 8). Therapeutic COM was endorsed in 72 patients and was comparable between HIC and LMIC sites (HIC 42/353 [11.9%] vs LMIC 30/222 [13.5%], OR 0.9, 95% CI 0.5–1.5, $p = 0.62$) (Fig. 4A and C).

Amongst LMIC cases, six notable cGS diagnoses with changes of management were described (Table 1). For example, in a 22-year-old Peruvian female with early-onset spastic paraparesis and lower limb hyperreflexia, and classic spastic gait, cGS identified a heterozygous, paternally inherited inframe insertion in the *GCH1* gene, classified as likely pathogenic for autosomal dominant GTP cyclohydrolase-1-deficiency (OMIM # 128230). This finding resulted in a change of the patient's clinical diagnosis from early-onset

hereditary spastic paraparesis to dopa responsive dystonia (DRD), and both she and her affected father, previously diagnosed with sporadic early-onset parkinsonism, responded to low doses of levodopa with almost complete control of symptoms.

Table 1

Notable LMIC Patient Management Changes. yo: years old; m: male sex; f: female sex; (G.XX.X): ICD code provided by clinician; kb: kilobase; PO: proband-only analysis; AD: autosomal dominant inheritance of the preceding disorder; XL: X-Linked inheritance of preceding disorder; LP- likely pathogenic variant classification; P- pathogenic variant classification; ID: intellectual disability; DD: developmental delay; NDD: neurodevelopmental delay; Het: heterozygous

Study ID, Location, age (sex)	Phenotype	Clinical Diagnosis Prior to cGS	cGS Finding & Disorder	Change in Management
00638 Peru 9 yo (m)	Early-onset generalized dystonia, hyperreflexia, knee pain, cramping, bilateral renal hydronephrosis	Early-onset dystonia, (G24.9)	Het. 71 kb deletion including exon 1 of *GCH1*; unknown inheritance (PO); P; Dopa-responsive dystonia (AD)	Low doses of levodopa/ decarboxylase inhibitor
00800 Peru 22 yo (f)	Early onset spastic paraparesis with lower limb hyperreflexia, spastic gait, mild scoliosis, pes equinovarus	Early-onset spastic paraplegia (G11.4)	Het. paternally inherited in-frame insertion in *GCH1*; LP; GTP cyclohydrolase – 1-deficiency (AD)	Low doses of levodopa/decarboxylase inhibitor
00316 Mexico 2 yo (f)	Muscle hypotonia, hyporeflexia, foot malposition, myopathic facies, normal early motor development but unable to walk independently	Spinal muscular atrophy	Absence of the *SMN1* c.840C allele, which is consistent with an absence of wild-type SMN1 & predicted to result in spinal muscular atrophy	Nusinersen
00818 Peru 14 yo (m)	Moderate ID, DD, short stature, epilepsy, obesity, aggressive behavior, poor attention & dysmorphic facial features	Undiagnosed ID with obesity & short stature	Het. *de novo* missense variant in *SLC2A1*; P; glucose transporter type I deficiency syndrome (AD)	Ketogenic diet with carnitine supplementation
00346 Mexico 10 yo (m)	Progressive hypotonia, weakness & muscle atrophy especially pronounced in lower extremities & proximal limbs, elevated CK	Undiagnosed muscular dystrophy	Hemizygous, maternally inherited intronic variant in *DMD*; LP; Dystrophinopathies (XL)	Cardiology referral, deflazacort recommended

Study ID, Location, age (sex)	Phenotype	Clinical Diagnosis Prior to cGS	cGS Finding & Disorder	Change in Management
00338 Mexico 3 yo (m)	Epilepsy, heterotopia, NDD, allergic rhinitis, cerebral palsy, generalized weakness, uncoordinated walk, hypertrichosis, minor dysmorphic features, focal cortical dysplasia.	Undiagnosed developmental disorder	Het. *de novo* missense variant in *GFAP*, P; Alexander disease (AD)	Recommended to follow neuroprotective measures such as avoiding long fasting periods, flickering lights & continuing care with a neuro-pediatrician.

Discussion

There is increasing evidence supporting the value of cGS as a first-line test in pediatric patients with a suspected rare genetic disease, particularly in large cohorts drawn from single high-income geographies.[9,15] There has been little investigation, however, of the impact of cGS in patients with reduced access to medical care. This is likely driven by three factors: (1) challenges deploying cGS to patients with limited resources, the majority of whom fall outside academic medical center referral areas; (2) difficulties maintaining patient and clinician engagement to enable assessment of COM after cGS; and (3) an untested assumption that the complexities of pediatric genetic disease, in combination with limited local resources, are unlikely to yield meaningful clinical benefit to patients in resource-encumbered settings. The iHope program has directly addressed the first two challenges, and enabled assessment of the third.

The analyses performed in this geographically diverse cohort focused on three elements of potential cGS impact – diagnostic yield, diagnostic evaluation (DE) and change of management (COM) – investigating the cohort as a whole and the differences between HIC and LMIC populations while controlling for possible confounders. The diagnostic yield of cGS across the cohort was comparable to similar studies in genetically and phenotypically diverse cohorts but was increased in LMIC compared to HIC patients.[7,8,14] Some of this difference may be explained by differential test utilization as patients from HIC sites were 6.6-times more likely than patients from LMIC sites to have had at least one prior genetic test, suggesting that cGS in the LMIC population was more often utilized as a first-line test, potentially in patients with more severe phenotypic presentations. This hypothesis is supported by cGS detection of an increased number and size of copy number variants in the LMIC population, which may have been identified by chromosomal microarray or karyotype if otherwise available.

As reflected in Fig. 2C, patient selection was consistent with genetic testing guidelines and prior investigations of the impact of exome and genome sequencing in patients with suspected rare genetic disorders, inclusive of patients with neurodevelopmental disorders, multiple congenital anomalies and epilepsy.[9] However, to control for the elevated LMIC diagnostic yield, and the influence of site-specific patient selection variation on DE and COM, analyses were stratified by test result category (Supplementary

Fig. 7 and Supplementary Table 3). The impact of cGS on DE revealed a high proportion of individuals across the cohort with results that informed clinician assessment, with more patients from LMIC sites experiencing a change in DE compared to patients from HIC sites even when stratified by test result category. This may in part reflect implicit patient heterogeneity given the diversity of rare genetic disease phenotypes, and differential prior testing and clinical specialist availability at LMIC sites. Assessment of COM revealed that up to 70% of patients had a change in clinical care. When stratified by test result category the overall likelihood of COM was comparable between the LMIC and HIC populations, suggesting that precision care in this population is mediated by access to an etiological diagnosis. Indeed, as highlighted in Table 1, unanticipated changes in management can lead to marked improvements in quality of life for patients in LMIC.

Recent studies have raised the possibility that access to cGS, or equivalent molecular testing, supports diagnostic equity.[16] The data presented here further strengthens this position and suggests that for patients with RGD, the highest likelihood of receiving precision care is tied to a precision diagnosis, irrespective of geographic location or resourcing of the local care team. Despite the high rate of COM in the LMIC population, lack of access to therapeutic interventions was reported for some patients, illustrating the need to bolster international funding for under-resourced clinical systems. For example, a patient from the Democratic Republic of the Congo with *THRA*-related congenital non-goitrous hypothyroidism was referred to an endocrinologist for specialty evaluation and medication management but was lost to follow-up after the family indicated they could not afford the recommended changes in care. Unlike infectious diseases, for which there is an extensive philanthropic network to support surveillance, diagnosis and treatment, patients with suspected RGDs remain underrecognized and underserved. Multi-national public-private partnerships focused on the delivery of diagnostics and therapeutics in LMIC are necessary to reduce the morbidity and mortality associated with monogenetic disorders.

Limitations of this investigation include its duration, use of multiple survey versions to assess DE and COM (Supplementary Table 2), and potential differences in barriers to obtaining access to the iHope program (which are likely to mimic inherit limitations in the availability and accessibility of molecular testing for patients with RGD). Future studies investigating clinically indicated cGS test utilization in the context of clinician ordering behaviors, and detailed analyses of both patient and clinical resource encumbrances, will refine our understanding of appropriate care and implementation pathways in LMIC.

The findings of this study have supported expansion of the iHope program to China and the development of iHope Genetic Health, a multi-institutional effort to reduce barriers to genomic testing worldwide, and suggest that equitable access to clinical genomic testing, in concert with additional investments in LMICs to support RGD clinical care, may reduce global healthcare inequities.

Methods

This study was reviewed by the WIRB-Copernicus Group (WCG) Institutional Review Board and determined to be exempt research as defined in 45CFR46.104(d)(4).

Clinical Genome Sequencing

Clinical GS was performed by the Illumina Clinical Services Laboratory in San Diego, California, and included interrogation of single nucleotide variants (SNVs), small insertions and deletions (indels), and copy number variants (CNVs) for all samples. Mitochondrial SNVs, biallelic absence of the *SMN1* c.840C allele and short tandem repeat analyses were added to the test during the study period (Supplementary Methods and Supplementary Table 4). Test result categories included: positive, in which a likely pathogenic or pathogenic variant(s) was reported in a disease-associated gene(s) consistent with the clinical presentation and expected disease inheritance pattern; inconclusive, in which variants of potential clinical significance were reported, inclusive of variants of unknown significance; and negative, in which no variants were reported related to the indication for testing. As recommended by the American College of Genetics and Genomics, optional reports for secondary findings were provided for all individuals tested in a family for 56 or 59 genes dependent of the time of testing, unless an individual opted-out of secondary findings (Supplementary Table 4).[17,18] Ancillary pharmacogenomics screen reports were provided for all tested individuals but were not considered in this analysis (Supplementary Table 5).

Phenotype Distribution Assessment

To systematically assess patient phenotypes, PhenoTagger was applied to clinician-provided patient phenotypes and Human Phenotype Ontology (HPO) terms were extracted.[19] HPO terms were mapped to one of 25 top-level terms by organ system (direct descendants of “Phenotypic abnormality”, HP:0000118).[20] To assess the impact of phenotype on diagnostic yield the unique number of phenotypic terms, and the phenotypic information content of these HPO terms, was assessed (Methods in the Supplementary Information).

Ancestry Assessment

Principal component analysis of the iHope cohort was performed using genotypes extracted from the cGS data overlayed on a reference set of 3,320 genomes derived from the 1000 Genomes Project, Human Genome Diversity Project and Simons Genome Diversity Project.[21–23] Admixture was assessed using ADMIXTURE 1.3.0 with a supervised background of the 3,320 individuals from the reference cohort.[24] For additional details see the Methods section in the Supplementary Information.

Assessment of cGS test results on the clinical evaluation and patient management

A cGS impact survey was provided to the ordering clinicians for patients who pursued cGS from June 2016 through October 2020 (Fig. 1 and Supplementary Table 2). The survey included multiple choice questions and free-text responses designed to assess the impact of cGS results, regardless of report status, on DE and COM. Three survey versions were used over the course of the study (Supplementary Table 2). Electronic survey responses were received 472 days, on average, after clinical report delivery. For additional details see the Methods section in the Supplementary Information.

Outcomes

Diagnostic yield was assessed as the proportion of individuals with a positive test result. A change in DE was defined as the endorsement of any of six categories of clinical assessment. Two delineations of COM were utilized for analysis: a restricted definition which included specialty referrals, imaging and testing, therapeutic interventions and palliative care; and an extended definition that also considered avoidance of additional testing and genetic counseling. To avoid inflation of estimates of diagnostic yield, secondary and incidental findings were not included in diagnostic yield assessment, but DE and COM survey responses allowed for consideration of these results. For additional details see the Methods Outcomes in the Supplementary Information and Supplementary Table 2.

Statistical Analysis

Cohort characteristics

Cohort characteristics including phenotypes, age, sex and family structure were summarized using descriptive statistics. iHope program sites were categorized based on country of site location according to the World Bank gross national income (GNI) per capita designations: low (<$1,085), lower-middle (<$4,255), upper-middle (<$13,205), and high income (≥ $13,205).[25] The categories low, lower-middle, and upper-middle were collapsed into a single category, LMIC, and compared to HIC.

Outcomes analysis

Observational factors potentially associated with cGS test results, DE, and COM were modeled using a directed acyclic graph to facilitate selection of factors for adjustment when examining outcomes in the context of LMIC and HIC sites (Supplementary Fig. 7). A multinomial regression was performed to assess the impact of LMIC or HIC site designation on test result category. The influence of LMIC or HIC site designation on survey question endorsement was evaluated by multiple logistic regression using a logit link where the survey question determined a binary outcome. The logistic regression included test result category and LMIC or HIC site designation as explanatory variables. The multiplicative influence of LMIC or HIC site designation on the odds of endorsement of DE or COM was estimated by treating HIC as the reference category. A stratified analysis approach was employed to investigate the potential association of LMIC or HIC site designation on survey response outcomes within each test result category. The data were separated by test result and binary survey responses were binned according to LMIC or HIC site designation. An ANOVA framework was utilized to compare the number of phenotypic terms submitted per proband between patients from HIC and LMIC sites. Bootstrap resampling was used to compute CIs, and a permutation analysis was performed to determine a p-value. The same procedure was applied to the computed phenotypic information content to consider the association of phenotypic information with test result. For additional details please see the Methods section in the Supplementary Information.

Declarations

Acknowledgements

The international iHope team wishes to thank the patients and parents for their participation in the program. Illumina acknowledges Sylwia Urbaniak and Joey Flores for operational support of the iHope program; Illumina Laboratory Services inclusive of the software, bioinformatics, and interpretation and reporting teams; Krista Bluske and Amanda Buchanan for test result category coding; and the Illumina shipping and logistics team. Instituto Nacional de Ciencias Neurológicas Perú acknowledges Elison Sarapura-Castro, Andrea Rivera-Valdivia, Jorge La Serna and the team from Neurogenetics DNA Bank for clinical and logistic support. Edgardo Rebagliati Martins National Hospital acknowledges doctors Nelson Purizaca and Ana Protzel and the staff of the cytogenetics laboratory for their clinical and logistic support. The University of Kinshasha, Democratic Republic of Congo acknowledges Laetitia Mavinga, Dahlie Tshika and Nadine Nsiangani for patient care and support, and Cathy Songo and Chloé Musuamba for sample collection and shipment preparation. The University of Tennessee Health Science Center and Le Bonheur Children's Hospital acknowledges Megan Fonville, MS, CGC, Emma Bradley, MS, CGC, Roya Mostafavi, MS, CGC and Shannon Holtrop, MS, CGC for patient care and support.

Author Disclosures

Erin Thorpe, Evgenii Chekalin, Keisha Robinson, Jason Button, Alka Malhotra, Maren Bennett, James Avecilla, Andrew Warren, Max Arseneault, Tasha Kalista, Ali Crawford, Subramanian S. Ajay, and Denise. L. Perry are employees of and stockholders in Illumina, Inc. Ryan J. Taft is an employee of and stockholder in Illumina, Inc. and a compensated advisor to Creyon Bio. Julia Ortega is a stockholder in Illumina, Inc. and employee of C2N Diagnostics. John Belmont is a stockholder in Illumina, Inc. and was compensated as a research advisor through Genetics & Genomics Services Inc. Chad Shaw was compensated as a consultant through Genetics & Genomics Services Inc. for statistical analysis performed. Vani Rajan is a stockholder in Illumina, Inc.

References

1. Nguengang Wakap S, Lambert DM, Olry A, et al. Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. *Eur J Hum Genet*. 2020;28(2):165-173. doi:10.1038/s41431-019-0508-0
2. Hunter DJ, Reddy KS. Noncommunicable Diseases. *N Engl J Med*. 2013;369(14):1336-1343. doi:10.1056/NEJMra1109345
3. Online Mendelian Inheritance in Man, OMIM®. World Wide Web URL: https://omim.org/
4. Bick D, Jones M, Taylor SL, Taft RJ, Belmont J. Case for genome sequencing in infants and children with rare, undiagnosed or genetic diseases. *J Med Genet*. 2019;56(12):783-791. doi:10.1136/jmedgenet-2019-106111
5. John E. Dicken, Robert Copeland, Kristen Joan Anderson, et al. *Rare Diseases: Although Limited, Available Evidence Suggests Medical and Other Costs Can Be Substantial*. United States Government Accountability Office; 2021.
6. World Health Organization Science Council. *Accelerating Access to Genomics for Global Health: Promotion, Implementation, Collaboration, and Ethical, Legal, and Social Issues: A Report of the WHO*

Science Council. World Health Organization; 2022:46. https://www.who.int/publications/i/item/9789240052857

7. 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. *Npj Genomic Med*. 2018;3(1):16. doi:10.1038/s41525-018-0053-8
8. Bertoli-Avella AM, Beetz C, Ameziane N, et al. Successful application of genome sequencing in a diagnostic setting: 1007 index cases from a clinically heterogeneous cohort. *Eur J Hum Genet*. 2021;29(1):141-153. doi:10.1038/s41431-020-00713-9
9. Manickam K, McClain MR, Demmer LA, et al. Exome and genome sequencing for pediatric patients with congenital anomalies or intellectual disability: an evidence-based clinical guideline of the American College of Medical Genetics and Genomics (ACMG). *Genet Med*. 2021;23(11):2029-2037. doi:10.1038/s41436-021-01242-6
10. Lindstrand A, Ek M, Kvarnung M, et al. Genome sequencing is a sensitive first-line test to diagnose individuals with intellectual disability. *Genet Med*. 2022;24(11):2296-2307. doi:10.1016/j.gim.2022.07.022
11. van der Sanden BPGH, Schobers G, Corominas Galbany J, et al. The performance of genome sequencing as a first-tier test for neurodevelopmental disorders. *Eur J Hum Genet*. Published online September 16, 2022. doi:10.1038/s41431-022-01185-9
12. El Naofal M, Ramaswamy S, Alsarhan A, et al. The genomic landscape of rare disorders in the Middle East. *Genome Med*. 2023;15(1):5. doi:10.1186/s13073-023-01157-8
13. Scocchia A, Wigby KM, Masser-Frye D, et al. Clinical whole genome sequencing as a first-tier test at a resource-limited dysmorphology clinic in Mexico. *Npj Genomic Med*. 2019;4(1):5. doi:10.1038/s41525-018-0076-1
14. 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. *Genet Med*. Published online May 2023:100896. doi:10.1016/j.gim.2023.100896
15. The 100,000 Genomes Project Pilot Investigators, Smedley D, Smith KR, et al. 100,000 Genomes Pilot on Rare-Disease Diagnosis in Health Care — Preliminary Report. *N Engl J Med*. 2021;385(20):1868-1880. doi:10.1056/NEJMoa2035790
16. The NICUSeq Study Group, Krantz ID, Medne L, et al. Effect of Whole-Genome Sequencing on the Clinical Management of Acutely Ill Infants With Suspected Genetic Disease: A Randomized Clinical Trial. *JAMA Pediatr*. 2021;175(12):1218. doi:10.1001/jamapediatrics.2021.3496
17. Green RC, Berg JS, Grody WW, et al. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet Med*. 2013;15(7):565-574. doi:10.1038/gim.2013.73
18. Kalia SS, Adelman K, Bale SJ, et al. Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics. *Genet Med*. 2017;19(2):249-255. doi:10.1038/gim.2016.190

19. Luo L, Yan S, Lai PT, et al. PhenoTagger: A Hybrid Method for Phenotype Concept Recognition using Human Phenotype Ontology. *Bioinforma Oxf Engl*. Published online January 20, 2021:btab019. doi:10.1093/bioinformatics/btab019
20. Köhler S, Gargano M, Matentzoglu N, et al. The Human Phenotype Ontology in 2021. *Nucleic Acids Res*. 2021;49(D1):D1207-D1217. doi:10.1093/nar/gkaa1043
21. The 1000 Genomes Project Consortium, Corresponding authors, Auton A, et al. A global reference for human genetic variation. *Nature*. 2015;526(7571):68-74. doi:10.1038/nature15393
22. M'Charek A. *The Human Genome Diversity Project: An Ethnography of Scientific Practice*. 1st ed. Cambridge University Press; 2005. doi:10.1017/CBO9780511489167
23. Mallick S, Li H, Lipson M, et al. The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature*. 2016;538(7624):201-206. doi:10.1038/nature18964
24. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res*. 2009;19(9):1655-1664. doi:10.1101/gr.094052.109
25. World Bank. World Development Indicators. https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

Figures

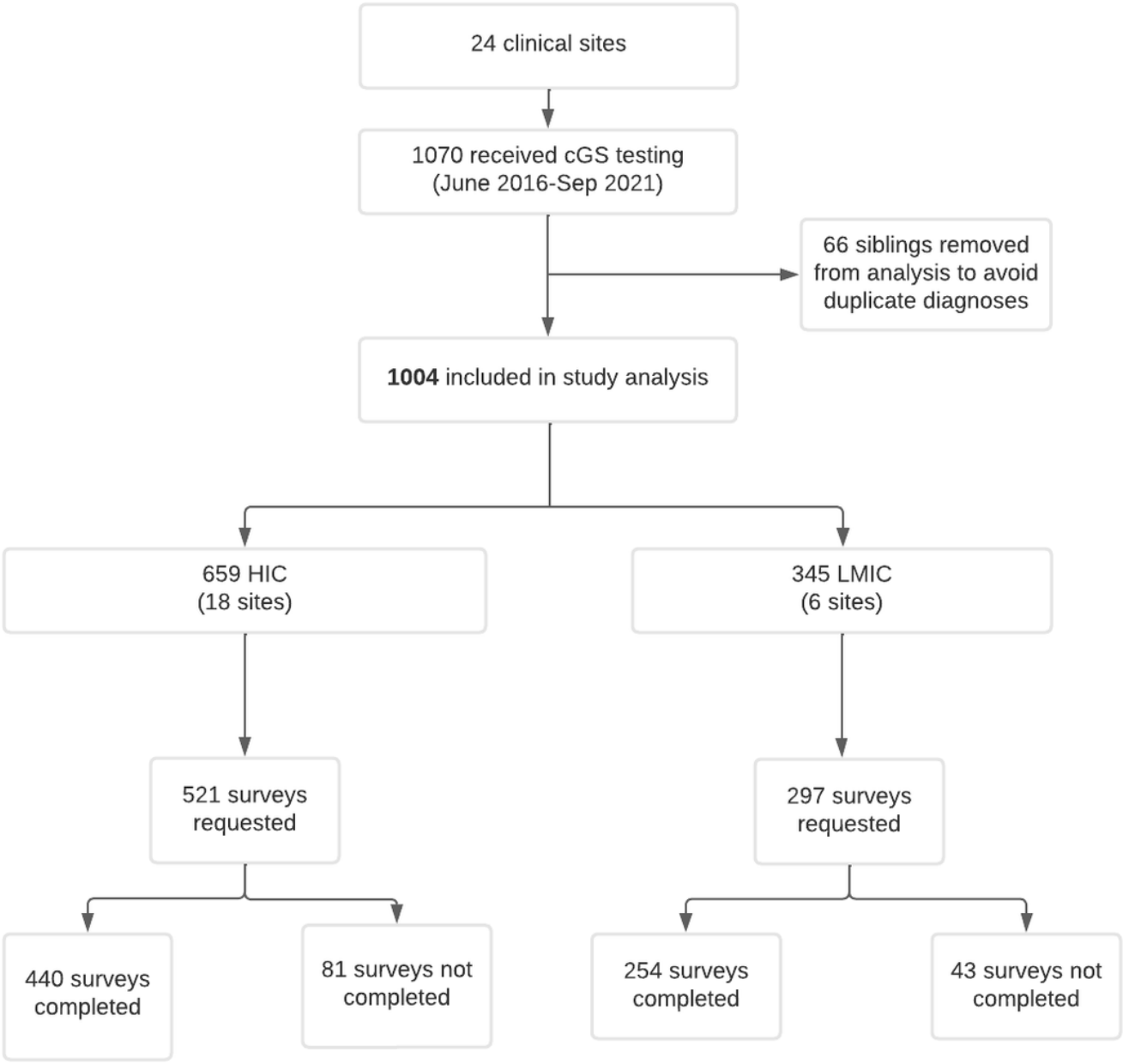

Figure 1

STROBE diagram. iHope program observational cohort patient ascertainment and cGS impact survey responses depicted using a STROBE diagram. Abbreviations: cGS, clinical genome sequencing; HIC, high-income country; LMIC, low- and middle-income country.

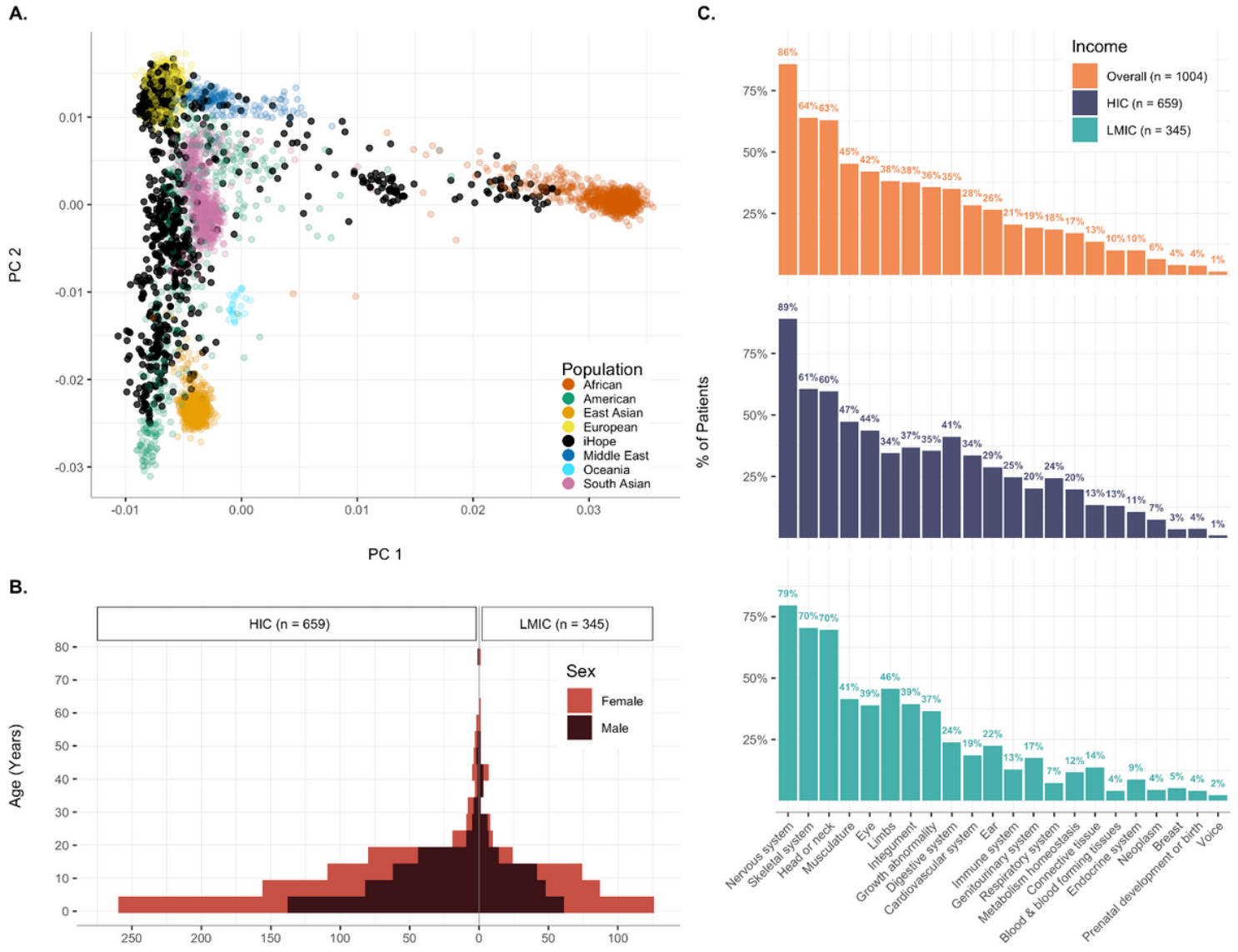

Figure 2

iHope patient demographics and phenotypic presentation. (A) Principal component analysis of the iHope cohort probands (black) overlayed on seven human superpopulations derived from the 1000 Genomes, Human Genome Diversity and Simons Genome Diversity datasets. (B) Age and sex distributions stratified by high-income country (HIC) and low- and middle-income country (LMIC) sites. HIC age (y): mean 8.9, median 6.3, range 0 days-77.1 years; LMIC age (y): mean 9.6, median 6.6, range 26 days-77.9 years with two males of unknown age assigned the mean LMIC patient age. HIC sex: male 350/659 (53.1%), female 309/659 (46.8%); LMIC sex: male 187/345 (54.2%), female 158/345 (45.7%). There are no statistically significant differences in age and sex distributions between the HIC and LMIC populations (p=0.36 and p=0.69). (C) Summary distribution of top-level Human Phenotype Ontology terms nested beneath “Phenotypic abnormality” (HP:0000118) across the iHope cohort and stratified by HIC and LMIC sites.

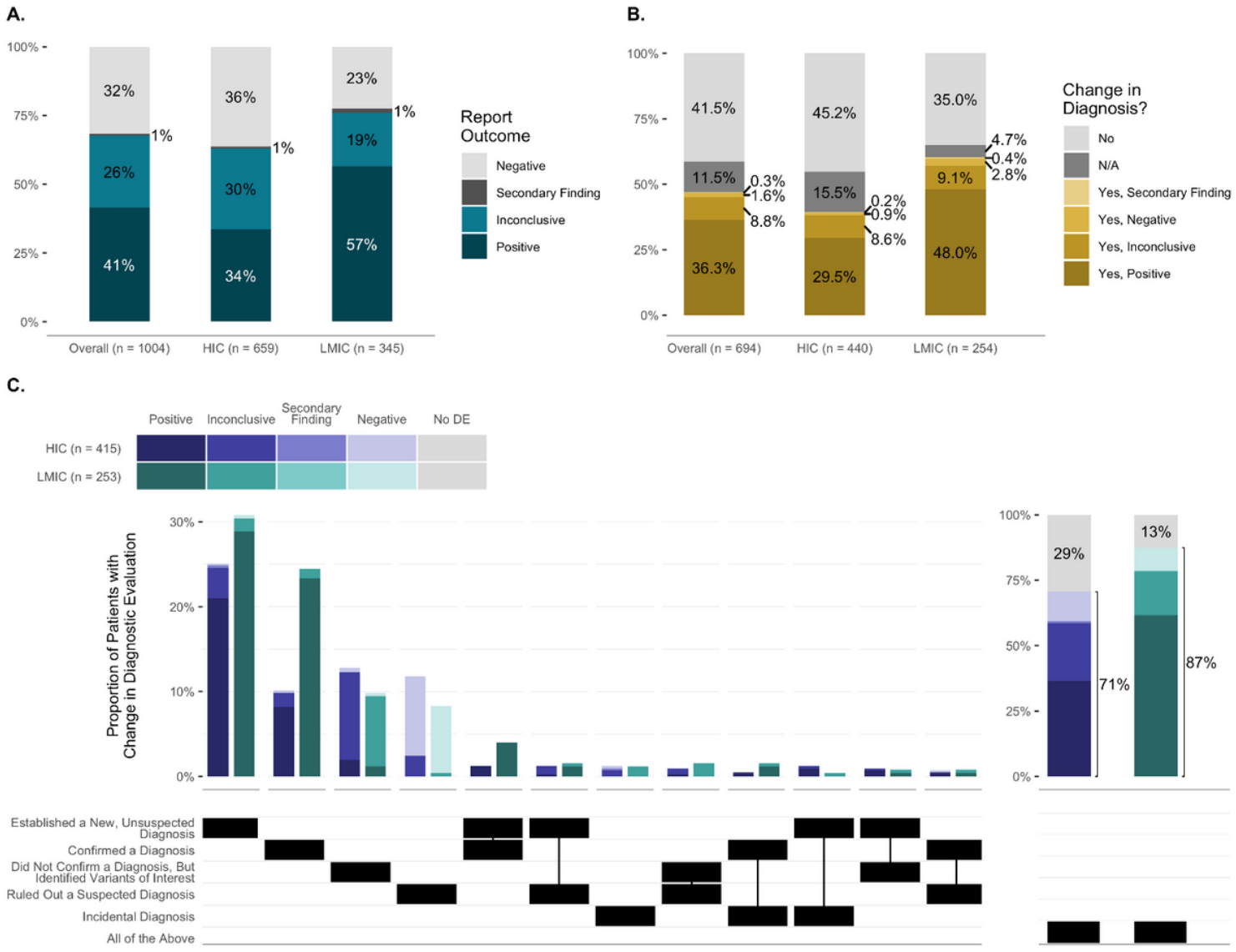

Figure 3

Diagnostic yield of cGS and its impact on clinical diagnosis and diagnostic evaluation. (A) Overall diagnostic yield of cGS stratified by test result category and by HIC and LMIC sites. (B) Change in clinical diagnosis due to cGS with stratification by test result category and HIC and LMIC sites. (C) The impact of cGS results on diagnostic evaluation. Response options are reflected in the text in the lower left, with black rectangles representing endorsement by the responding clinician. Multiple response options could be endorsed. The vertical black lines connecting black rectangles reflect a response combination. Bar plots above each set of responses indicate the proportion of patients with the indicated DE response combination stratified by test result category. Stacked bars to the far right reflect the combined total responses supportive of an impact on diagnostic evaluation.

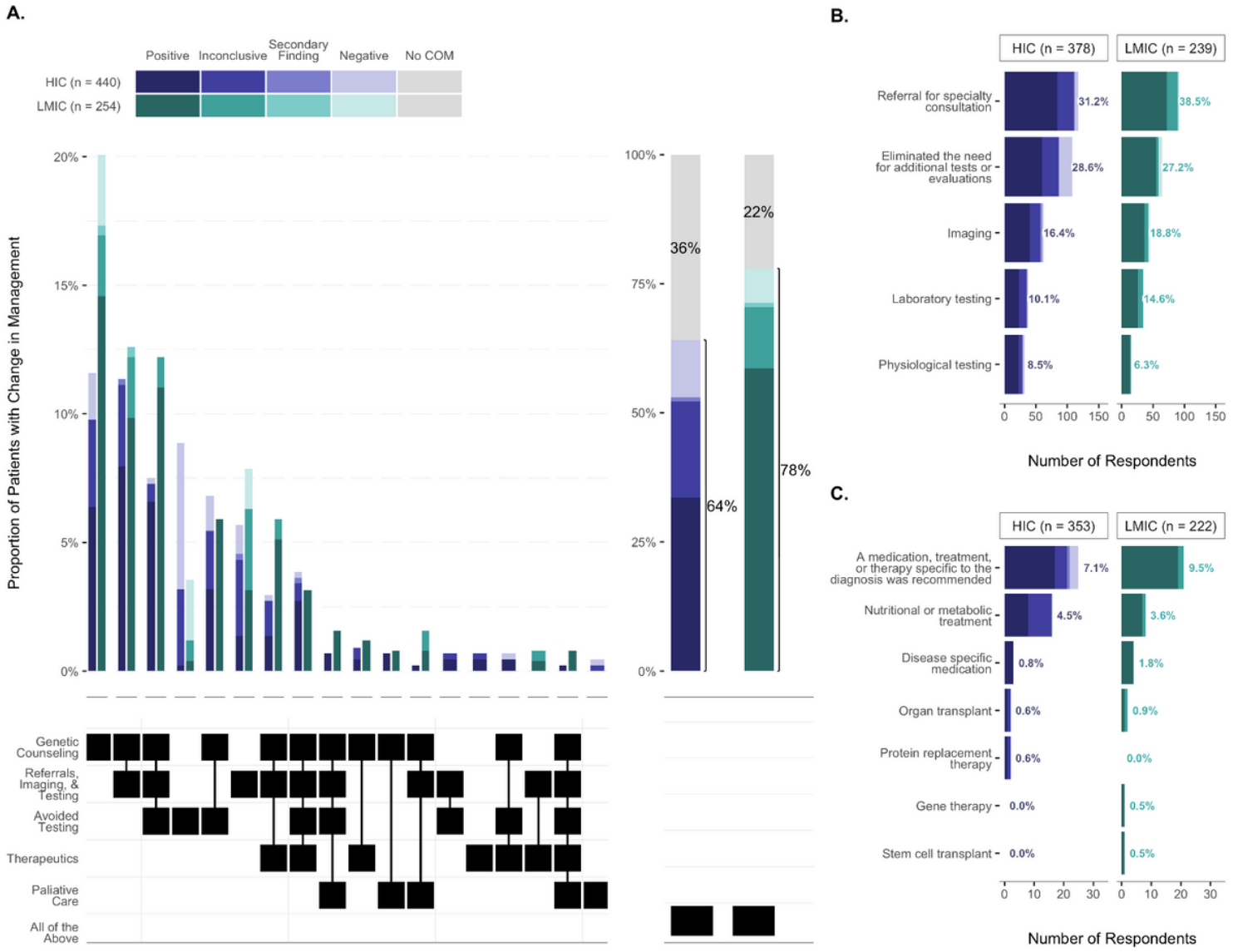

Figure 4

Changes of management associated with cGS testing results. (A) Change of management (COM) across the cohort and stratified by HIC and LMIC sites and test result category. Response options are reflected in the text in the lower left, with black squares representing endorsement of a COM category. Multiple response options could be endorsed. The vertical black lines connecting black rectangles reflect a particular combination of response options that were endorsed. Bar plots above each set of responses indicate the proportion of patients with the indicated COM response combination stratified by test result category. (B) Distribution of COM categorized to the referrals, imaging and testing COM category, stratified by HIC and LMIC sites and test result category. (C) Distributions of COM categorized to the therapeutics COM category, stratified by HIC and LMIC sites and test result category.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ThorpeSupplementaryInformationiHope1K.docx
- ThorpeSupplementaryTable1CohortsummaryinclusiveofdiagnosesandcGSImpactsurveyresponses.xlsx
- ThorpeSupplementaryTable2cGSImpactSurveyquestionsversionsandresponses.xlsx