INVESTIGATIONAL DEVICE EXEMPTION (IDE) APPLICATION

IDE Title:

All of Us Research Program: Return of Genetic Results (AoURP gRoR)

Submission Number:

G200165

June 8, 2020

As amended November 18, 2020 to include changes requested by FDA during review.

Additions to original submission are highlighted.

Table of Contents

Table of Contents (continued)

List of Figures

List of Tables

List of Tables (continued)

List of Abbreviations and Definitions

Abbreviation	Definition
ACD	Advisory Committee to the Director (of NIH)
ACMG	American College of Medical Genetics and Genomics
AoU	*All of Us*
AoUHDR	*All of Us* Hereditary Disease Risk
AoUPGx	*All of Us* Pharmacogenetic(s)
AoURP	*All of Us* Research Program
AoUWGS	*All of Us* Whole Genome Sequencing
B	Benign
BCBS	Blue Cross Blue Shield
BCM	Baylor College of Medicine
BCM-HGSC	Baylor College of Medicine-Human Genome Sequencing Center
BD	Becton Dickinson
BI	Broad Institute
BRCA1	Breast Cancer Gene Type 1
BRCA2	Breast Cancer Gene Type 2
CAP	College of American Pathologists
CDC	Centers for Disease Control and Prevention
Chr	Cell cycle genes homology region
CI	Concordance Index
CL	Color Genomics
CLIA	Clinical Laboratory Improvement Amendments
CLSI	Clinical and Laboratory Standards Institute
Co-Is	Co-responsible investigators
CPIC	Clinical Pharmacogenetics Implementation Consortium
CSER	Clinical Sequencing Evidence-Generating Research
CVL	Clinical Validation Laboratory
DRAGEN	Illumina Dynamic Read Analysis for GENomics
DRC	Data and Research Center
DV	Direct Volunteer
eMERGE	Electronic MEdical Records and GEnomics
EMSI	Examination Management Services, Inc.
FFPE	Formalin-fixed paraffin-embedded
FISMA	Federal Information Security Management Act
FN	False negative
FP	False positive
FQHCs	Federally Qualified Health Centers

Abbreviation	Definition
G6PD	Glucose-6-phosphate dehydrogenase
GA4GH	Genome Alliance for Genomic Health
GC(s)	Genome Center(s)
GCR	Genetic Counseling Resource
GeT-RM	Genetic Testing Reference Materials Coordination Program
GIV	Global Imbalance Value
GLP	Good laboratory practices
GP	Genotype probability
gRoR	Return of Genetic Results
GT	genotype
HDR	Hereditary Disease Risk
HDRR	Hereditary Disease Risk Report
Het	heterozygous
HGSC-CL	Human Genome Sequencing Center Clinical Laboratory
HIPAA	Health Insurance Portability and Accountability Act
Hom	homozygous
HPO	Healthcare Provider Organization
Indel	Insertion and Deletion Mutations
IRB	Institutional Review Board
ISO	International Organization for Standardization
IVD	In vitro diagnostic
LB	Likely Benign
LDLR	Familial hypercholesterolemia
LDT	Laboratory Developed Test
LIMS	Laboratory Information Management System
LoF	Loss of function
LP	Likely Pathogenic
HDR gROR	Hereditary Disease Risk Return of Genetic Results
N/A	Not applicable
Neg	negative
ng	nanogram
NGS	Next generation sequencing
NHGRI	National Human Genome Research Institute
NHLBI	National Heart, Lung, and Blood Institute
NIH	National Institutes of Health
NIST	National Institute of Standards and Technology
NPA	Negative percent agreement

Abbreviation	Definition
NPV	Negative predictive agreement
NWGC	Northwestern Genome Center
OHRP	Office of Human Research Protections
OSHA	Occupational Safety and Health Administration
P	Pathogenic
P/LP	Pathogenic/likely pathogenic
PCR	Polymerase Chain Reaction
PF	Pass-fail
PGx	Pharmacogenomics or Pharmacogenetics
PHI	Protected Health Information
PI	Principal Investigator
PMA	Premarket approval
PMID	PubMed Identifier
Pos	positive
PPA	Positive percent agreement
PPV	Positive predictive value
QC	Quality Control
QMS	Quality Management System
qPCR	Quantitative polymerase chain reaction
RDR	Raw Data Repository
Recall	Calculation of the true positive rate; a measure of sensitivity
RMCs	Regional Medical Centers
ROC	Receiver operator characteristic
RoR	Return of Results
RTA	Real time analysis
SBS	Sequencing-By-Synthesis
SNP	Single-nucleotide polymorphisms
SNVs	Single-nucleotide variants
SOPs	Standard Operating Procedure
Std dev	Standard deviation
TN	True Negative
TOPMED	Trans-Omics for Precision Medicine
TP	True Positive
TPC	The Participant Center
TPMT	Thiopurine s-methyltransferase
UC	University of California
UW	University of Washington

Abbreviation	Definition
VA	Veteran Affairs
VAMC	Veteran Affairs Medical Center
VCF	Variant call format/files
VUS	Variants of Uncertain/Unknown Significance
WES	Whole Exome Sequencing
WGS	Whole Genome Sequencing

1. Name and Address of the Sponsor

Name	National Institutes of Health *All of Us* Research Program c/o Stephanie Devaney, PhD, Chief Operating Officer
Address	6001 Executive Blvd, Suite 8112 Rockville, MD 20895
Phone number	(301) 451-5996
Fax	(301) 451-6084
Email address	stephanie.devaney@nih.gov

Name and contact information of alternate contact:

Bradley Ozenberger, PhD, Genomics Team Lead
All of Us Research Program
National Institutes of Health
bradley.ozenberger@nih.gov

2. Report of Prior Investigations

2.1 Bibliography

The nascent field of precision medicine has developed substantial evidence for the potential value of personal genomic results to improve health and wellness. Several consortia efforts have published research on patient outcomes, behaviors, healthcare usage, etc., after receiving health-related genetic results derived from genome sequencing data. These consortia include: eMERGE, CSER, MedSeq Project, PeopleSeq, Geisinger MyCode, HealthSeq, PGen.

Key findings:

- Patients showed good comprehension of most key facts about the study, including its purpose, procedures, potential risks, and policies regarding the return of study results. (Roberts 2018)
- Healthy individuals who underwent predispositional sequencing were enthusiastic about their experience and not distressed by their results. While reporting value in their health-related results, few participants reported making medical or lifestyle changes. (Zoltick 2019)
- Disclosure of incidental genetic results shows little to no adverse impact on participants and adds only modestly to near term healthcare costs (Hart 2019)
- WGS neither worsened nor improved self-rated health, anxiety, or depression scores. (Vassy 2017)

Bibliography References:

Carere DA, VanderWeele TJ, Vassy JL, et al. Prescription Medication Changes Following Direct-To-Consumer Personal Genomic Testing: Findings From the Impact of Personal Genomics (PGen) Study. Genet Med. 2017 May;19(5):537-545. doi: 10.1038/gim.2016.141.

Carey DJ, Fetterolf SN, Davis FD, et al. The Geisinger MyCode community health initiative: an electronic health record-linked biobank for precision medicine research. Genet Med. 2016;18(9):906–13.

Gordon AG, Zouk H, Venner E et al. Frequency of genomic incidental findings among 21,915 eMERGE network participants. Genet Med. In press.

Hart MR, Biesecker BB, Blout CL, et al. Secondary findings from clinical genomic sequencing: Prevalence, patient perspectives, family history assessment, and healthcare costs from a multi-site study. Genet Med. 2019 May;21(5):1100–1110. doi:10.1038/s41436-018-0308-x.

Roberts J, Robinson J, Diamond P, et al. Patient understanding of, satisfaction with, and perceived utility of whole-genome sequencing: findings from the MedSeq Project. Genet Med. 2018 September; 20(9): 1069–1076. doi:10.1038/gim.2017.223.

Sanderson SC, Linderman MD, Suckiel SA, et al. Psychological and behavioural impact of returning personal results from whole-genome sequencing: the HealthSeq project. Eur J Hum Genet. 2017;25(3):280–92.

Vassy JL, Christensen KD, Schonman EF, et al. The impact of whole-genome sequencing on the primary care and outcomes of health adult patients: a pilot randomized trial. Ann Intern Medicine. 2017;167(3):159-169. doi: 10.7326/M17-0188.

Vassy JL, Brunette CA, Majahalm N, et al. The Integrating Pharmacogenetics in Clinical Care (I-PICC) Study: Protocol for a Point-Of-Care Randomized Controlled Trial of Statin Pharmacogenetics in Primary Care. Contemp Clin Trials 2018;75:40-50. doi:10.1016/j.cct.2018.10.010.

Zoltick ES, Linderman MD, McGinniss MA, et al. Predispositional genome sequencing in healthy adults: design, participant characteristics, and early outcomes of the PeopleSeq Consortium. Genome Med. 2019;11:10. *https://doi.org/10.1186/s13073-019-0619-9*.

2.2 Summary of Non-Clinical Laboratory Studies

All studies have been conducted in compliance with the Good Laboratory Practice (GLP) regulations in 21 CFR Part 58.

Performance characteristics of the *All of Us* Research Program (AoU) Return of Genetics Report (gRoR) device were established through a series of laboratory studies that are detailed in this section. The accuracy and precision studies used DNA inputs amounts of 350 ng and 750 ng, with approximately 1/3 of the samples being at 350 ng and 2/3 at 750 ng. A summary of the studies and numbers of specimens included in each study are shown in **Table 1**.

Table 1. Summary of assessments and numbers of specimens included in each study

	Assessment							
Sample Cohort	Reportable Range	Accuracy	Precision	Interlab Concordance	Limit of Detection	Extraction Performance	Molecular Index Performance	Assay Fail Rates
Blood-derived Patient Samples		430						430
Blood-derived Healthy Donor Aliquots	5		28	20	21	15	28	28
Genome-in-a-bottle Reference Cell Lines	1	1			37		37	37
Cell lines with known HDRR variants	30	30	30	30			30	30
CDC GET-RM PGx cell lines	68	164	62	175			164	175
Historical blood-derived WGS								25,028

2.2.1 Test Elements

2.2.1.1 Specimen Type

The specimen type is genomic deoxyribonucleic acid (DNA) derived from whole blood (collected in 4.0 mL BD Vacutainer® Plus plastic whole blood tube, Becton Dickinson Cat No. 367861) or buffy coat (collected in 10.0 mL BD Vacutainer® Plus plastic whole blood tube, Becton Dickinson Cat No. 366643).

2.2.1.2 Interrogated Regions of the Genome

Whole genome sequencing will be performed on specimens collected from participants in the AoURP. The interrogated portion of the whole genome that will be examined for return of results to AoURP participants includes 223,913 bases across 66 genes. Of these, 59 genes comprise the AoURP Hereditary Disease Risk (AoUHDR) panel: *ACTA2, ACTC1, APC, APOB, ATP7B, BMPR1A, BRCA1, BRCA2, CACNA1S, COL3A1, DSC2, DSG2, DSP, FBN1, GLA, KCNH2, KCNQ1, LDLR, LMNA, MEN1, MLH1, MSH2, MSH6, MUTYH, MYBPC3, MYH11, MYH7, MYL2, MYL3, NF2, OTC, PCSK9, PKP2, PMS2, PRKAG2, PTEN, RB1, RET, RYR1, RYR2, SCN5A, SDHAF2, SDHB, SDHC, SDHD, SMAD3, SMAD4, STK11, TGFBR1, TGFBR2, TMEM43, TNNI3, TNNT2, TP53, TPM1, TSC1, TSC2, VHL,* and *WT1*. Seven genes are solely for reporting of specific AoU pharmacogenetic (AoUPGx) associated genes: *TPMT, NUDT15, DPYD, UGT1A1, SLCO1B1, CYP2C19,* and *G6PD*. More details of the specific exclusions from the gene regions listed are detailed in Section 2.2.2.5 (Reportable Range).

2.2.1.3 Performance Needs

Sponsor specifications

- Minimize risk for prospective enrollees by attaining high accuracy (>99%) when calling reportable variants in previously characterized clinical (blood-derived) patient samples.
- Achieve an accuracy >99% in determination of calling Pharmacogenetic (PGx) alleles in previously characterized clinical (blood-derived) patient samples and reference cell lines such that further confirmatory testing is not required.
- As a demonstration of risk minimization for prospective enrollees in the study, demonstrate equivalency in variant calling across the sites doing whole genome sequencing. For a set of samples processed at all sites there should be high concordance (>99%) in reportable pathogenic/likely pathogenic (P/LP) variant calls and PGx alleles.
- As a demonstration of risk minimization for prospective enrollees in the study, provide evidence of robustness and reproducibility across all sites doing whole genome sequencing (WGS). When examining variants called across a set of samples run at each laboratory, achieve >99% equivalency between variants in an inter- and intra-laboratory analysis.
- As a demonstration of risk minimization for prospective enrollees in the study, provide evidence of superior participant comprehension of health-related reports. When assessing each type of report, >90% of participants must understand the key concepts intended to be communicated in the reports.

2.2.2 Test Performance Characteristics

2.2.2.1 Accuracy

<u>Part 1 – Representation of genes</u>

A DNA sequence-based assay (WGS) underlies the AoU gRoR device and therefore the accuracy of that component of the device was determined in several ways. In Part 1, we present the orthogonally validated, true positive variants that were identified in clinical specimens across the range of reportable AoUHDR genes. We note that the representation of variants and genes tracks closely with the expected prevalence of reportable variants in the population based on an independent study of incidental findings (Gordon AG, et al., In press) in a largely overlapping set of genes (58 out of 59 AoUHDR genes were the same).

All three of the orthogonal comparator assays that contributed to the accuracy assessment are CLIA-validated panels used for return of results and all three largely overlap the AoUHDR target regions:

1. The eMERGESeq panel - A gene panel comprising a total of 109 genes and 1,551 SNV sites. The 109 genes included 56 based upon the American College of Medical Genetics and Genomics (ACMG) actionable finding list as well as additional genes for specific conditions. The gene and SNV list was used to direct construction of targeted capture platforms at two sequencing centers (SCs): The Baylor College of Medicine Human Genome Sequencing Center (BCM-HGSC), Houston TX and the Broad Institute and Partners Laboratory for Molecular Medicine, Cambridge, MA. Broad used Illumina Rapid Capture probes for this panel and the BCM-HGSC used Roche-Nimblegen methods. Each group created in-solution capture probes spanning the entire targeted regions of the eMERGEseq panel. Sequence reads were aligned against human genome reference GRCh37.p12, and variants were identified using a suite of bioinformatic tools designed to detect single nucleotide variants, small insertions and deletions. Variants were classified according to the American College of Medical Genetics and Genomics guidelines for sequence and variants were signed out by board certified clinical geneticists.

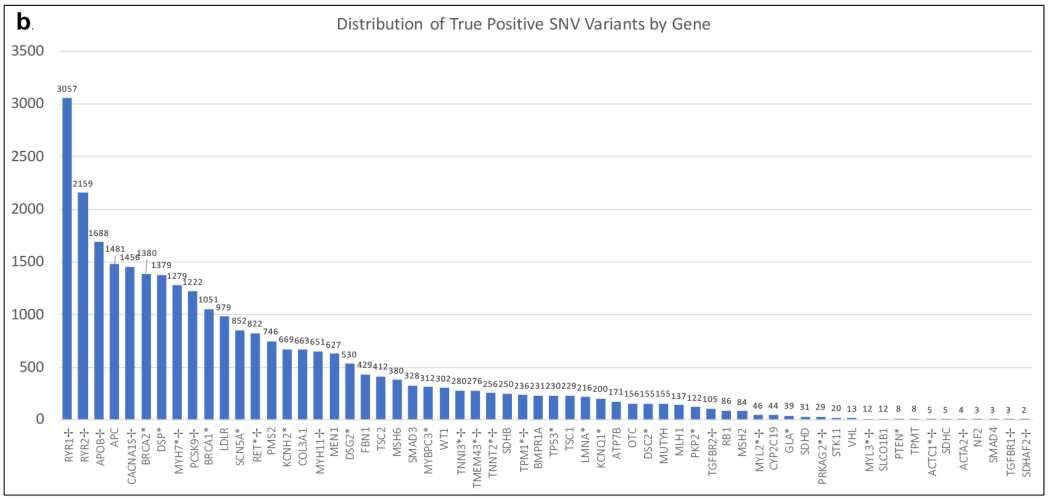
2. The UW Test: Laboratory procedures were performed by the Northwest Genomics Center under CLIA license MTS-60326571. Briefly, gDNA was subjected to a series of shotgun library construction steps and enriched for the target regions using probes synthesized by Twist Biosciences. Sequencing was accomplished using the Illumina NovaSeq 6000 instrument. Sequence reads were aligned against human genome reference GRCh37.p12, and variants were identified using a suite of bioinformatic tools designed to detect single nucleotide variants, small insertions and deletions. Variants were classified according to the American College of Medical Genetics and Genomics 2015 guidelines and variants were signed out by board certified clinical geneticists.

3. The Color test: Laboratory procedures were performed at the Color laboratory under CLIA (#05D2081492) and CAP (#8975161) compliance. Briefly, DNA was extracted, enriched for select regions using SureSelect XT probes, and then sequenced using NextSeq 500/550 or NovaSeq 6000 instruments. Sequence reads were aligned against human genome reference GRCh37.p12, and variants are identified using a suite of bioinformatic tools designed to detect single nucleotide variants, small insertions and deletions, and large structural variants. Variants were classified according to the American College of Medical Genetics and Genomics 2015 guidelines for sequence variant interpretation, and all variant classifications were signed out by a board-certified medical geneticist or pathologist.

Blood-derived genomic DNA from 271 unique patients who had undergone prior targeted sequencing with clinically validated gene panels were subjected to WGS using the AoU process and pipelines. Variants within the AoUHDR regions were assessed for accuracy and assigned a status of true positive (TP), false positive (FP), or false negative (FN). The remaining sites within the interval were designated true negative (TN). The number and distribution of TP variants across all samples are shown in Figure 1.

a.

Gene	Number of TP SNVs	Number of TP Insertions and Deletions
RYR1✢	3057	3
RYR2✢	2159	8
APOB✢	1688	54
APC	1481	2
CACNA1S✢	1456	1
BRCA2*	1380	46
DSP*	1379	62
MYH7*✢	1279	12
PCSK9✢	1222	58
BRCA1*	1051	32
LDLR	979	5
SCN5A*	852	1
RET*✢	822	1
PMS2	746	51
KCNH2*	669	6
COL3A1	663	41
MYH11✢	651	24
MEN1	627	0
DSG2*	530	4
FBN1	429	5
TSC2	412	6
MSH6	380	4
SMAD3	328	0
MYBPC3*	312	138
WT1	302	0
TNNI3*✢	280	0
TMEM43*✢	276	1
TNNT2*✢	256	122
SDHB	250	6
TPM1*✢	236	0
BMPR1A	231	0
TP53*	230	1
TSC1	229	33
LMNA*	216	0
KCNQ1*	200	1
ATP7B	171	0
OTC	156	0
DSC2*	155	3
MUTYH	155	0
MLH1	137	1
PKP2*	122	4
TGFBR2✢	105	0
RB1	86	3
MSH2	84	42
MYL2*✢	46	0
CYP2C19	44	0
GLA*	39	1
SDHD	31	0
PRKAG2*✢	29	0
STK11	20	0
VHL	13	1
MYL3*✢	12	0
SLCO1B1	12	0
PTEN*	8	2
TPMT	8	0
ACTC1*✢	5	0
SDHC	5	1
ACTA2✢	4	0
NF2	3	0
SMAD4	3	2
TGFBR1✢	3	13
SDHAF2✢	2	0



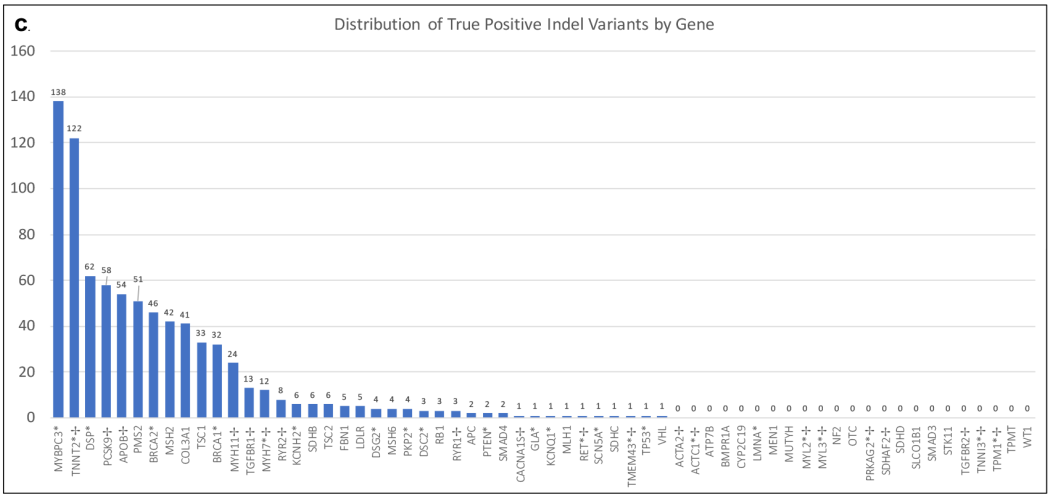


Figure 1. Number of TP variants (Single-nucleotide variants [SNVs], Insertion and Deletion Mutations [indels]) observed in the AoUHDR genes. *indicates genes with potential surgical interventions indicated for a pathogenic variant. ✢indicates genes in which Loss of Function (LoF) is not a known mechanism of disease pathogenesis, meaning insertions or deletions are typically not reportable.

In this determination of accuracy, at least one TP variant was observed in each of the AoUHDR genes. We note that three additional genes (*SLCO1B1, TMPT,* and *CYP2C19*) were included in this analysis as they were on the orthogonal gene panels used for comparison. These genes are not part of the AoUHDR panel but instead are part of the AoUPGx gene list. The accuracy and representation of AoUPGx alleles is considered separately below.

The number of different variants that were represented across the 59 AoUHDR genes varies widely. This distribution is expected and mimics the expected prevalence of reportable variants in the population based on a comparison to the eMERGE III study. (Gordon AG, et al., In press) Gordon et al. sequenced 21,915 individuals with gene panels containing 58 of the 59 AoUHDR genes.

The frequency of incidental findings that were deemed reportable by the eMERGE clinical sites is shown in **Figure 2** which is reproduced with the permission of the authors and without alteration. We note that the genes in the most frequent and least frequent bins of the distribution shows substantial overlap between the AoURP validation set and those found in the eMERGE study. The ATP7B gene is not included in the eMERGE study. However, this gene is on the other two comparator assay panels that were used to provide truth data for the accuracy assessments.

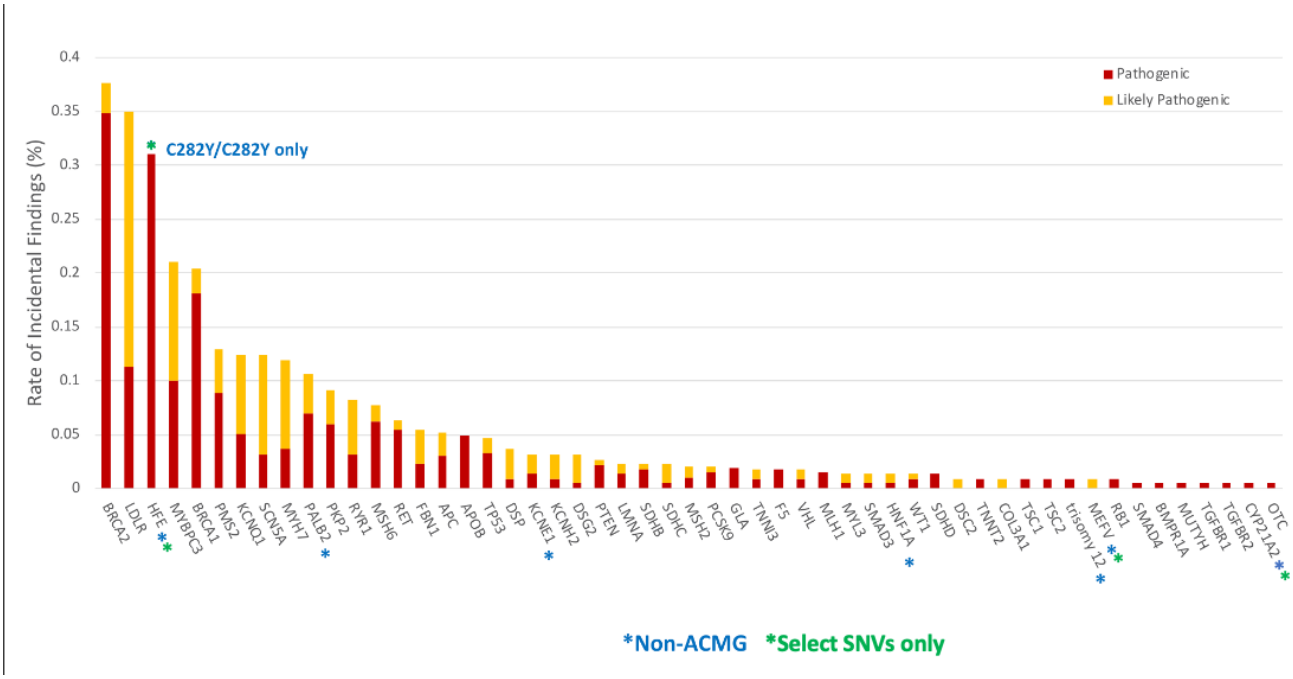


Figure 2. Excerpt from Gordon et al. paper. Observed frequency of reportable findings across 21,915 people in a set of genes highly overlapping with AoUHDR.

Specific variants in the population with evidence of increased disease risk are represented in this dataset, including founder alleles in *BRCA1* and *BRCA2*. Table 2 below summarizes these variants, called in independent samples in all three AoU Genome Centers.

Table 2. Founder mutations in BRCA1 and BRCA2 called with 100% accuracy in independent samples (BCM, Baylor College of Medicine; BI, Broad Institute; and UW, University of Washington)

Sample ID	Gene	DNA Change	Amino Acid Change	GRCh37 Coordinates	Variant Type	Zygosity	Site	Notes
1000070014	BRCA1	c.68_69delAG	p.Glu23Valfs*17	17:41276045_41276046del	Deletion	Heterozygous	BCM	Founder Mutation
1000073979	BRCA1	c.5266dupC	p.Gln1756Profs*74	17:41209080dupC	Insertion	Heterozygous	BCM	Founder Mutation
SM-IN9H2	BRCA1	c.68_69delAG	p.Glu23Valfs*17	17:41276045_41276046delCT	Deletion	Heterozygous	BI	Founder Mutation
SM-IN8QQ	BRCA1	c.5266dupC	p.Gln1756Profs*74	17:41209079_41209080insG	Insertion	Heterozygous	BI	Founder Mutation
SM-JPV8J	BRCA1	c.5266dupC	p.Gln1756Profs*74	17:41209079T>TG	Insertion	Heterozygous	BI	Founder Mutation
SM-JPV8Y	BRCA1	c.5266dupC	p.Gln1756Profs*74	17:41209079T>TG	Deletion	Heterozygous	BI	Founder Mutation
SM-JPVA4	BRCA1	c.5266dupC	p.Gln1756Profs*74	17:41209079T>TG	Insertion	Heterozygous	BI	Founder Mutation
SM-JPVAE	BRCA1	c.5266dupC	p.Gln1756Profs*74	17:41209079T>TG	Insertion	Heterozygous	BI	Founder Mutation
1000074456	BRCA2	c.5946delT	p.Ser1982Argfs*22	13:32914438del	Deletion	Heterozygous	BCM	Founder Mutation
SM-GZQK6	BRCA2	c.5946delT	p.Ser1982Argfs*22	13:32914438delT	Deletion	Heterozygous	BI	Founder Mutation
SM-JPTV9	BRCA2	c.5946delT	p.Ser1982Argfs*22	13:32914437GT>G	Deletion	Heterozygous	BI	Founder Mutation
SM-JPV89	BRCA2	c.5946delT	p.Ser1982Argfs*22	13:32914437GT>G	Deletion	Heterozygous	BI	Founder Mutation
SM-JPVA9	BRCA2	c.5946delT	p.Ser1982Argfs*22	13:32914437GT>G	Deletion	Heterozygous	BI	Founder Mutation
323249	BRCA2	c.5946delT	p.Ser1982Argfs*22	13:32914438delT	Deletion	Heterozygous	UW	Founder Mutation

Conclusion: We have demonstrated the ability of the AoURP device to accurately call variants in each of the genes that comprise the AoUHDR. These include clinically significant variants e.g., the BRCA founder mutations. The representation of variants that were tested in our validation study closely resembles the expected prevalence of reportable variants in these genes based on an independent study by eMERGE.

Part 2 – Accuracy as a function of variant type and context

To assess variant calling accuracy across a range of genomic contexts, variant sub-type, and zygosity, we examined data from a series of patient samples. The genomic contexts were defined using bed files from the Genome Alliance for Genomics and Health (GA4GH, *www.ga4gh.org*) benchmarking-tools repository (*https://github.com/ga4gh/benchmarking-tools/tree/d88448a68a79ed322837bc8eb4d5a096a710993d/resources/stratification-bed-files*).

Specifically, the genomic contexts examined were:

1. Segmental duplications: sequences with homologous replications throughout the genome.
2. Low mappability regions: regions where a 100 bp single ended read with no base errors could be mapped to a different location with at most two mismatches and one insertion or deletion.
3. Low complexity regions: regions of repetitive repeating sequence, such as homopolymers or short tandem repeats.
4. Low GC regions: regions where fewer than 25% of bases are guanine/cytosine (GC) pairs.
5. High GC regions: regions where greater than 85% of bases are GC.

In addition to the genomic contexts above, genes with regions of high homology to known pseudogenes within the AoUHDR were examined. These genes are BMPR1A, PMS2, PTEN, SDHC, and SDHD. Within these regions, we observed a Positive Percent Agreement (PPA) of 99.05% (95% Confidence Interval [CI]: 98.30% – 99.81%) and a Negative Percent Agreement (NPA) of 100% (95% CI: 100% – 100%).

In the analysis of the accuracy that is summarized in Table 3, the device False Negative Rate (FNR) is 0.26% for all variant types combined based on analysis of the accuracy cohort of clinical specimens

that are represented in Section 2, Table 3. In that analysis we categorize the variants identified by the orthogonal panel testing as True Positive (29,475) and any of those that are not called by the device as False Negatives (76). We calculate FNR as FN/FN+TP (76/76+475) and express it as a percentage. The three genes that are not in the HDR (SLCO1B1, TMPT, and CYP2C19) do not contribute any False Negative calls to the overall number of False Negatives used to calculate the false negative rate in our response. They contribute 64 True Positive calls. To calculate PPA and NPA performance metrics for the methods used for the AoU gRoR, we compared our data to data generated from the clinical gene panels used in the eMERGE study (The eMERGE Consortium, 2019) and at UW (Pritchard CC, et al., 2012) (**Table 3**).

Table 3. Accuracy of variant calling across genomic contexts and variant types in patient samples. PPA, NPA. Here + or – infer whether a variant was present in either the panel or the corresponding WGS sample.

	Panel+/ WGS-	Panel-/ WGS+	Panel+/ WGS+	Panel-/ WGS-	PPA [95% CI]	NPA [95% CI]
All compiled	76	201	29,475	45847148	99.74% [99.68%-99.81%]	100% [100%-100%]
P/LP Variants only	0	0	145	28474503	100% [100%-100%]	100% [100%-100%]
SNP only	23	6	28,710	45848108	99.92% [99.87%-99.97%]	100% [100%-100%]
Insertions only	9	8	225	29887204	96.20% [94.30%-98.00%]	100% [100%-100%]
Deletions only	33	12	540	42170158	94.24% [92.65%-95.83%]	100% [100%-100%]
Segmental Duplications	12	2	2,119	38466267	99.44% [99.10%-99.78%]	100% [100%-100%]
Low Mappability Regions	18	12	1,069	36912107	98.34% [97.84%-98.85%]	100% [100%-100%]
Low Complexity Regions	24	69	956	38072735	97.55% [96.35%-98.75%]	100% [100%-100%]
Low GC Regions	6	12	197	14792575	97.04% [95.87%-98.22%]	100% [100%-100%]
High GC Regions	0	0	23	3891553	100% [100%-100%]	100% [100%-100%]
Heterozygous Variants only	75	147	18,295	45858382	99.59% [98.48%-99.70%]	100% [100%-100%]
Homozygous Variants only	1	83	11,180	45865561	99.99% [99.97%-100%]	100% [100%-100%]

Note: There were five frequently observed variants that failed consistently across data sets and, therefore, were not included in the analysis. These five variants are benign or of uncertain significance and would not be reportable. These variants are shown in **Table 4**.

Table 4. Excluded recurrent false positive and false negative variants

Gene	Variant start site
MSH2	chr2:47641559
MSH2	chr2:47641562
APOB	chr2:21266774
PMS2	chr7:6037057
TSC1	chr9:135773000
PCSK9	chr1:55505552

Conclusion: We have presented performance data as a function of variant type and genomic context. This includes 100% accuracy calling the reportable P/LP variants in the validation cohort samples.

Part 3 – Accuracy of select AoUHDR pathogenic variant calling at all sites

We assessed accuracy using concordance of variant calls in human cell lines (30) with previously characterized pathogenic variants in the AoUHDR genes. These previously characterized human cell line-derived DNA samples (purchased from Coriell Biorepository) were processed through the production workflows for WGS, capillary, or panel sequencing at all AoU Genome Centers (GCs) and Clinical Validation Laboratories (CVLs). Calls of the known pathogenic variants were assessed at each site and concordance measured. All CGs and CVLs produced 100% concordant calls for each variant examined (**Table 5**).

Table 5. Call concordance of select pathogenic variants in human cell lines. * samples marked with asterisks were run in triplicate at one site. All calls between replicates were identical

Sample	Gene	Variant	Broad (WGS) Correct calls	Broad (WGS) Incorrect Calls	Broad (WGS) No-calls	BCM (WGS) Correct calls	BCM (WGS) Incorrect Calls	BCM (WGS) No-calls	UW (WGS) Correct calls	UW (WGS) Incorrect Calls	UW (WGS) No-calls	BCM CVL (Sanger) Correct calls	BCM CVL (Sanger) Incorrect Calls	BCM CVL (Sanger) No-calls	Color CVL (Panel) Correct calls	Color CVL (Panel) Incorrect Calls	Color CVL (Panel) No-calls	UW CVL (Panel) Correct calls	UW CVL (Panel) Incorrect Calls	UW CVL (Panel) No-calls	Overall Concordance
NA00483	LDLR	CYS646TYR	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	6/6
NA01448	LDLR	GLY197DEL	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1*	0	0	1/1	0	0	6/6
NA03949	APC	GLN541TER	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	6/6
NA04391	GLA	ASN215SER	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	6/6
NA05258	ATP7B	GLU1064ALA	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	6/6
NA05761	ATP7B	THR977MET	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	6/6
NA06102	TSC2	ARG1743GLN	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	6/6
NA07414	TSC1	c.994_995insA	2/2	0	0	2/2	0	0	2/2	0	0	2/2	0	0	2/2	0	0	2/2	0	0	12/12
NA10080	PTEN	GLN261TER	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	6/6
NA11410	APC	3149delC	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	6/6
NA11630	MEN1	5-BP INS, NT 317	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1*	0	0	1/1	0	0	6/6
NA13713	BRCA1	GLU1250TER	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1*	0	0	1/1	0	0	6/6
NA14090	BRCA1	2-BP DEL, 185AG	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	6/6
NA14095	BRCA1	5256delG	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	6/6
NA14170	BRCA2	1-BP DEL, 6174T, FS	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1*	0	0	1/1	0	0	6/6
NA14624	BRCA2	5946delCT	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	6/6
NA14637	BRCA1	ARG1443TER	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	6/6
NA14805	BRCA2	TRP194TER	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	6/6
NA16658	RET	CYS620PHE	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	6/6
NA21987	FBN1	ARG2057TER	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	6/6
NA22051	COL3A1	GLY708ASP	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	6/6
NA22606	COL3A1	766delA	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	6/6
NA23079	COL3A1	1763_1769delGTGCTCC insTAAG	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1*	0	0	1/1	0	0	6/6
NA23429	OTC	ARG320TER	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	6/6
NA23650	RYR1	ARG401CYS | c.7463_7475del13	2/2	0	0	2/2	0	0	2/2	0	0	2/2	0	0	2/2	0	0	2/2	0	0	12/12
NA23780	LMNA	LEU35PRO	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	6/6
NA23850	OTC	ARG40HIS	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	6/6
NA25330	CACNA1S	THR1354SER	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	6/6
NA25516	RYR1	ARG109TRP | p.Arg2241	2/2	0	0	2/2	0	0	2/2	0	0	2/2	0	0	2/2	0	0	2/2	0	0	12/12
NG03506	LMNA	GLY608GLY	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	1/1	0	0	6/6

Conclusion: We demonstrated 100% concordance across all three AoU Genome Centers and CVLs when calling P/LP variants in the same samples at each site.

Part 4 – Accuracy of variant calling on an established reference standard sample

Accuracy, measured by TP, TN, FP, FN, PPV, NPV were computed by comparing WGS results from the National Institute of Standards and Technology (NIST) human reference cell line (NA12878) to the Genomes In a Bottle (GIAB) v3.3.2 gold standard truth set (Zook JM, et al. Epub 2019/04/03). Results are shown for different genomic contexts and variant types over the reportable region in **Table 6** and over the whole genome in **Table 7**. Note: PPV and NPV were not calculated if there were fewer than ten events in the given category.

Table 6. Accuracy of variant calls across the reportable region of NA12878

Reportable Region	TP	TN	FP	FN	PPV [95% CI]	NPV [95% CI]

All variants	131	218551	0	0	100% [97.2% - 100%]	100% [100% - 100%]
SNP only	129	218551	0	0	100% [97.2% - 100%]	100% [100% - 100%]
Insertions and Deletions	2	218551	0	0	***	100% [100% - 100%]
Segmental Duplications	12	19073	0	0	100% [73.5% - 100%]	100% [100% - 100%]
Low Mappability Regions	1	3532	0	0	***	100% [100% - 100%]
Low Complexity Regions	4	2435	0	0	***	100% [100% - 100%]
Low GC Regions	2	2722	0	0	***	100% [100% - 100%]
High GC Regions	0	7	0	0	***	100% [100% - 100%]
Heterozygous variants only	82	218551	0	0	100% [95.6% - 100%]	100% [100% - 100%]
Homozygous variants only	49	218551	0	0	100% [92.7% -100%]	100% [100% - 100%]

Table 7. Accuracy of variant calls across the whole genome of NA12878

Whole Genome	TP	TN	FP	FN	PPV [95% CI]	NPV [95% CI]
All variants	3688413	2571369697	3907	2448	99.89% [99.89% - 99.90%]	100% [100% - 100%]
SNP only	3207969	2571369697	2522	1346	99.92% [99.92% - 99.92%]	100% [100% - 100%]
Insertions and Deletions	480444	2571369697	1385	1102	99.71% [99.70% - 99.73%]	100% [100% - 100%]
Segmental Duplications	59054	43102836	375	329	99.37% [99.30% - 99.43%]	100% [100% - 100%]
Low Mappability Regions	172097	99469838	1458	1296	99.16% [99.12% - 99.20%]	100% [100% - 100%]
Low Complexity Regions	257660	62578838	879	823	99.66% [99.64% - 99.68%]	100% [100% - 100%]
Low GC Regions	179834	120814998	276	211	99.85% [99.83% - 99.86%]	100% [100% - 100%]
High GC Regions	413	377579	5	0	98.80% [97.23% - 99.61%]	100% [100% - 100%]
Heterozygous variants only	2233008	2571369697	3614	1621	99.84% [99.83% - 99.84%]	100% [100% - 100%]
Homozygous variants only	1454259	2571369697	256	790	99.98% [99.98% - 99.98%]	100% [100% - 100%]

Clinical specimens and cell lines had small numbers of insertions and deletions of different sizes. To assess the performance of variant calling as a function of insertion and deletion size we looked at variants across the entire genome in the NA12878 reference sample (**Figure 3**).

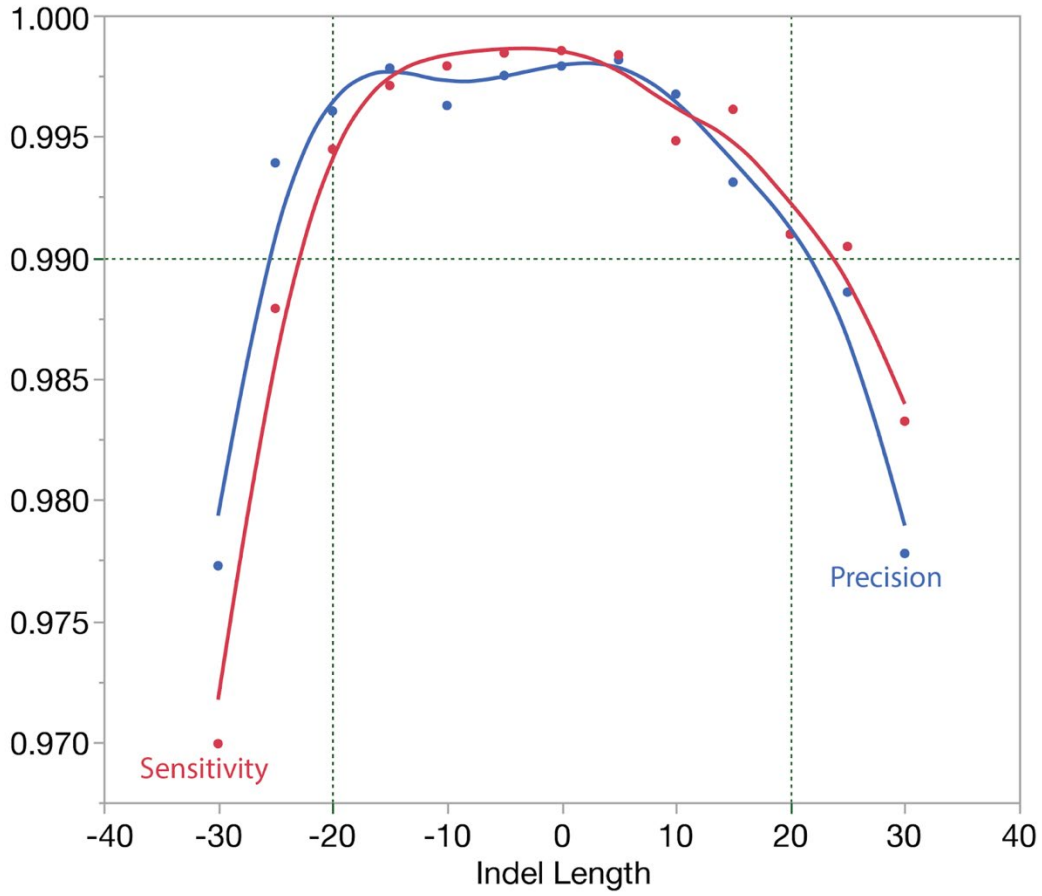


Figure 3. Performance (sensitivity red line & precision blue line) of variant calling from WGS as a function of insertion size (positive numbers) and deletion size (negative numbers). We observe accuracy of >99% (horizontal dotted line) in events up to 20 bases in length (indicated by vertical dotted lines) and ≥97% in events out to 30 bases in length.

Conclusion: We observed high accuracy using the AoUWGS device when compared to a well-established truth sample (NA12878), including highly accurate insertion and deletion calling up to at least 20 bases in length.

Part 5 – Accuracy of AoUPGx calling

To assess the accuracy of the device with respect to variant calls in PGx genes we studied 159 patient samples previously tested using orthogonal assays (Sanger sequencing, gene panel sequencing, or genotyping assays) and found to have reportable alleles (32 alleles across the seven AoUPGx genes) (**Table 8**). All patient samples used in the Accuracy analysis were derived from anticoagulated whole blood. As noted above for the AoUHDR panel, the selected AoUPGx clinical samples provide a representation of genes and alleles that align with the expected prevalence of reportable alleles in the general population as determined by the Clinical Pharmacogenetics Implementation Consortium (CPIC®) (*www.cpicpgx.org*). Observed concordance was 595/595 calls (100%).

Table 8. Call concordance of AoUPGx alleles using WGS and orthogonal methods in real patient samples

Gene and star allele	Number of Clinical sample(s)	Broad (WGS) Correct Calls	Broad (WGS) Incorrect Calls	BCM (WGS) Correct Calls	BCM (WGS) Incorrect Calls	UW (WGS) Correct Calls	UW (WGS) Incorrect Calls	Color CVL (WES) Correct Calls	Color CVL (WES) Incorrect Calls	BCM CVL (Sanger) Correct Calls	BCM CVL (Sanger) Incorrect Calls	UW CVL (Panel) Correct Calls	UW CVL (Panel) Incorrect Calls	UW CVL (Array) Correct Calls	UW CVL (Array) Incorrect Calls	Overall Concordance
CYP2C19*10	1	1/1	0	-	-	-	-	1/1	0	-	-	-	-	-	-	2/2
CYP2C19*17	40	25/25	0	15/15	0	-	-	25/25	0	15/15	0	-	-	-	-	80/80
CYP2C19*17/*17	1	1/1	-	-	-	-	-	1/1	-	-	-	-	-	-	-	2/2
CYP2C19*2	48	33/33	0	-	-	15/15	0	33/33	-	-	-	-	-	15/15	0	96/96
CYP2C19*2/*2	3	3/3	0	-	-	-	-	3/3	-	-	-	-	-	-	-	3/3
CYP2C19*22	1	-	-	1/1	0	-	-	-	-	1/1	0	-	-	-	-	2/2
CYP2C19*3	3	3/3	0	-	-	-	-	3/3	-	-	-	-	-	-	-	3/3
CYP2C19*35	1	1/1	0	-	-	-	-	1/1	-	-	-	-	-	-	-	1/1
CYP2C19*4	6	3/3	0	3/3	0	-	-	3/3	0	3/3	0	-	-	-	-	12/12
CYP2C19*8	1	-	-	-	-	1/1	0	-	-	-	-	-	-	1/1	0	2/2
CYP2C19*9	2	-	-	2/2	0	-	-	-	-	2/2	0	-	-	-	-	4/4
DPYD c.1129-5923C>G	2	2/2	0					2/2	-	-	-	-	-	-	-	4/4
DPYD c.1679T>G (*13)	3	3/3	0					3/3	-	-	-	-	-	-	-	6/6
DPYD c.2846A>T	2	-	-	2/2	0	-	-	-	-	2/2	0	-	-	-	-	4/4
DPYD*2 (c.1905+1G>A)	2	-	-	2/2	0	-	-	-	-	2/2	0	-	-	-	-	4/4
G6PD A-202A_376G	3	3/3	0	-	-	-	-	3/3	0	-	-	-	-	-	-	6/6
G6PD Canton, Taiwan-Hakka, Gifu-like, Agrigento-like	1	1/1	0	-	-	-	-	1/1	0	-	-	-	-	-	-	2/2
G6PD Ilesha	2	2/2	0	-	-	-	-	2/2	0	-	-	-	-	-	-	4/4
G6PD Mediterranean, Dallas, Panama, Sassari, Cagliari, Birmingham	5	5/5	0	-	-	-	-	5/5	0	-	-	-	-	-	-	10/10
G6PD Seattle, Lodi, Modena, Ferrara II, Athens-like	1	1/1	0	-	-	-	-	1/1	0	-	-	-	-	-	-	2/2
G6PD Union, Maewo, Chinese-2, Kalo	2	2/2	0	-	-	-	-	2/2	0	-	-	-	-	-	-	4/4
NUDT15*2	2	2/2	0	-	-	-	-	2/2	0	-	-	-	-	-	-	4/4
NUDT15*3	9	8/8	0	-	-	1/1	0	8/8	0	-	-	-	-	1/1	0	18/18
SLCO1B1*15	28	22/22	0	-	-	6/6	0	22/22	0	-	-	-	-	6/6	0	56/56
SLCO1B1*15/*15	4	4/4	0	-	-			4/4	0	-	-	-	-	-	-	8/8
SLCO1B1*17	3	-	-	-	-	3/3	0	-	-	-	-	-	-	3/3	0	6/6
SLCO1B1*5	8	7/7	0	-	-	1/1	0	7/7	0	-	-	-	-	1/1	0	16/16
TPMT*2	6	5/5	0	1/1	0	-	-	5/5	0	1/1	0	-	-	-	-	12/12
TPMT*3A	9	9/9	0	-	-	-	-	9/9	0	-	-	-	-	-	-	18/18
TPMT*3C	5	5/5	0	-	-	-	-	5/5	0	-	-	-	-	-	-	10/10
UGT1A1*27	3	3/3	0	-	-	-	-	3/3	0	-	-	-	-	-	-	6/6
UGT1A1*28	70	37/37	0	33/33	0	-	-	37/37	0	33/33	0	-	-	-	-	140/140
UGT1A1*36	12	4/4	0	8/8	0	-	-	4/4	0	8/8	0	-	-	-	-	24/24
UGT1A1*36/*36	1	1/1	0	-	-	-	-	1/1	0	-	-	-	-	-	-	2/2
UGT1A1*37	3	2/2	0	-	-	1/1	0	2/2	0	-	-	1/1	0	-	-	6/6
UGT1A1*6	8	3/3	0	5/5	0	-	-	3/3	0	5/5	0	-	-	-	-	16/16

The G6PD Asahi allele is somewhat common in only a specific ethnic group (African American) but is rare in European and Latino-descent populations (0.0003 to 0.01, respectively). A survey of available clinical samples at each site did not uncover any additional clinical samples for this allele. The defining site for Asahi is well covered in our sequencing data (32.4x) and we will be able to accurately identify the presence or absence of this allele in our samples. As stated above, coverage values are derived from a dataset of 104 samples, equally spread across the three genome centers, all selected to have a full-genome mean coverage of 30-35x (depths determined by reads with minimum mapping quality of 20 and a base quality 20). The mean coverage across these sites are greater than 30x for these callable regions, thus we will be able to accurately identify the presence or absence of these alleles in our samples. (The wider

standard deviations seen for the alleles on G6PD are due to the sex-dependent coverage of genes on chrX.)

In genes where reported allele frequencies are extremely rare in the population and no clinical samples were accessible, we substituted cell line samples to provide a more comprehensive list of reportable alleles in the validation (**Table 9** and **Table 10**).

The Genetic Testing Reference Materials Coordination Program (GeT-RM) provides extensively characterized PGx cell lines for quality control and proficiency testing across labs in the genetic testing community (Pratt VM, et al., 2016). The GeT-RM program has performed multiple pharmacogenetic assays on these PGx cell lines to define a truth set based on concordance among these orthogonal assays for use by clinical labs. WGS sequencing on 135 GeT-RM cell line samples covering six AoUPGx genes and 20 alleles was performed at each AoURP Center. Overall concordance was assessed and is shown in **Table 9**. The GCs produced 99.8% concordant calls for all the AoUPGx alleles interrogated in these samples.

The consensus analysis from one cell line (UGT1A1*28/*28, NA20509) was discordant with the expected result as provided in the GeT-RM documentation. Further review of the underlying GeT-RM data revealed an inconsistency in the consensus calls in that analysis. That fact, considered with the fact that all three GCs accurately called UGT1A1*28/*28 in 10 other samples, suggest that the original GeT-RM classification of the NA20509 may have been incorrect.

Table 9. Call concordance of AoUPGx alleles using WGS and orthogonal methods in GeT-RM cell lines

Gene and Star Allele	Coriell PGx Sample	BCM Correct	BCM Incorrect Call	UW Correct	UW Incorrect Call	Color Correct	Color Incorrect Call	Overall Concordance
CYP2C19*10	NA07439	1/1	0	1/1	0	1/1	0	3/3
CYP2C19*17	NA10846 NA10851 NA12236 NA12873 NA23297 NA07019 NA07048 NA07055 NA17102 NA17288 NA19207 NA19176 NA23878 NA11881 NA19444 NA19908 NA23348 NA19147 NA23313 NA07348 NA12145 NA12375 NA18518 NA17679 NA17641 NA07357 NA23246 NA07029 NA10865 NA07000 NA24008 NA17658 NA17074 NA18519 NA19819 NA12813 NA10831 NA19239	38/38	0	38/38	0	38/38	0	114/114
CYP2C19*17/*17	NA12336 NA19035 NA19109	3/3	0	3/3	0	3/3	0	9/9
CYP2C19*2	NA18564 NA07439 NA17288 NA19207 NA19176 NA12815 NA17290 NA18617 NA18873 NA18945 NA19226 NA19917 NA23405 NA24027 NA10856 NA17012 NA19003 NA18524 NA18855 NA07348 NA12145 NA12375 NA18518 NA17679 NA12753 NA19174 NA11832 NA18484 NA23877 NA17641 NA07357 NA18868 NA23874 HG01190 NA24009 NA12878 NA21781 NA18544 NA18540 NA17673 NA15245 NA19122	42/42	0	42/42	0	42/42	0	126/126
CYP2C19*2/*2	NA23090 NA23093 NA17061 NA12717 NA18509 NA20509	6/6	0	6/6	0	6/6	0	18/18
CYP2C19*3	NA18564 NA23246	2/2	0	2/2	0	2/2	0	6/6
CYP2C19*4	NA23878 NA23881 NA18552	3/3	0	3/3	0	3/3	0	9/9
CYP2C19*6	NA23874 NA19178	2/2	0	2/2	0	2/2	0	6/6
CYP2C19*8	NA23348 NA23872 NA07029 NA10865 NA23873	5/5	0	5/5	0	5/5	0	15/15
CYP2C19*9	NA24009 NA24008	2/2	0	2/2	0	2/2	0	6/6
DPYD c.1905+1G>A(*2)	NA23873	1/1	0	1/1	0	1/1	0	3/3
G6PD Ube Konan	NA23246	1/1	0	1/1	0	1/1	0	3/3
SLCO1B1*15	NA11881 NA07357 NA18552 NA10859 NA18526 NA06993 NA12892 NA07000 NA24008 NA17658 NA19109 NA12878 NA21781 NA20509 NA12003 NA17642 NA24217 NA10847 NA15245 HG00276	20/20	0	20/20	0	20/20	0	60/60
SLCO1B1*15/*15	NA06991	1/1	0	1/1	0	1/1	0	3/3
SLCO1B1*17	NA23246 NA18563 NA18992 NA18544 NA18540 NA11993	6/6	0	6/6	0	6/6	0	18/18
SLCO1B1*5	NA21781 NA10847	2/2	0	2/2	0	2/2	0	6/6
TPMT*3A	NA12753 NA17641 NA17673 NA15245 NA17702	5/5	0	5/5	0	5/5	0	15/15
TPMT*3C	NA18855 NA17061 HG00589 NA20296 NA18966 NA19920	6/6	0	6/6	0	6/6	0	18/18
TPMT*3C/*3C	NA19035	1/1	0	1/1	0	1/1	0	3/3
UGT1A1*27	HG00436	1/1	0	1/1	0	1/1	0	3/3
UGT1A1*28	NA07439 NA19207 NA19176 NA19444 NA19908 NA23348 NA17679 NA19174 NA11832 NA18484 NA23877 NA17641 NA07357 NA12717 NA18509 NA23246 NA24009 NA24008 NA17658 NA12878 NA12003 NA17642 NA24217 NA10847 NA18544 NA18540 NA11993 NA17702 HG00436 NA07056 NA12006 NA12156 NA17235 NA17454 NA18942 NA19143 NA17074 NA18519 NA19819 NA19122 NA19178 HG00276 NA19095 NA23275 NA19239	45/45	0	45/45	0	45/45	0	135/135
UGT1A1*28/*28	NA19147 NA23313 NA18855 NA18868 NA23874 NA20509 NA17227 NA17448 NA17657 NA12813 NA10831	10/11	1	10/11	1	10/11	1	30/33
UGT1A1*36	NA19207 NA19178 NA19095 NA19213	4/4	0	4/4	0	4/4	0	12/12
UGT1A1*37	HG01190 NA23275 NA19239 NA19920	4/4	0	4/4	0	4/4	0	12/12
UGT1A1*6	NA23246 NA18966 NA19007 NA18973	4/4	0	4/4	0	4/4	0	12/12
UGT1A1*6/*6	NA18980	1/1	0	1/1	0	1/1	0	3/3

For reportable AoUPGx alleles that were not present in the GeT-RM collection, we procured and processed 29 cell lines that were part of the 1000 Genomes Project and had PGx calling performed as part of that effort (Lee SB, et al., 2019). The genome centers produced **100% concordant** calls for all the AoUPGx alleles that were interrogated in these samples.

The false negative rate specifically for the 7 PGx genes was determined from Tables 8 and 9, where we compared WGS PGx calls on 159 patient samples with clinically validated calls and WGS PGx calls on 135 cell lines with Get-RM consensus calls, respectively. One PGx allele in

UGT1A1 was discordant across 812 alleles from the seven PGx genes that we tested, therefore the FNR was determined to be 1/(811+1) = 0.12%.

Table 10. AoUPGx calling across additional cell line samples from the 1000 Genomes Project

Gene and Star Allele	1000 Genome PGx Sample	Number of Samples	Correct Calls	Incorrect Calls	Overall Concordance
CYP2C19*16	NA19452	1	1/1	0	1/1
CYP2C19*2	HG02087 HG02511 NA18599 NA18617 NA18964 NA18978 NA18998 NA19000 NA19020 NA21110	10	10/10	0	10/10
CYP2C19*2/*2	HG02186 HG02188	2	2/2	0	2/2
CYP2C19*22	HG02318	1	1/1	0	1/1
CYP2C19*24	NA20356	1	1/1	0	1/1
CYP2C19*3	HG02141 HG02318	2	2/2	0	2/2
G6PD Asahi	HG02511	1	1/1	0	1/1
G6PD Aures	HG02165	1	1/1	0	1/1
G6PD Canton,Taiwan-Hakka,Gifu-like,Agrigento-like	HG02087 NA18617	2	2/2	0	2/2
G6PD Chinese-5	HG02186	1	1/1	0	1/1
G6PD Ilesha	NA19020	1	1/1	0	1/1
G6PD Kalyan-Kerala,Jamnaga,Rohini	NA21110	1	1/1	0	1/1
G6PD Seattle,Lodi,Modena,Ferrara-II,Athens-like	HG01620	1	1/1	0	1/1
G6PD Sibari	NA19323	1	1/1	0	1/1
G6PD Ube Konan	NA18998	1	1/1	0	1/1
G6PD Viangchan,Jammu	HG02141 HG02188	2	2/2	0	2/2
NUDT15*2	NA18599 NA18621 NA18622 NA18626 NA18633 NA18740 NA18964 NA18978 NA19000 NA19077	10	10/10	0	10/10
NUDT15*3	NA18599 NA18998	2	2/2	0	2/2
SLCO1B1*15	NA19000	1	1/1	0	1/1
SLCO1B1*17	HG02186 HG02188 NA18599 NA19000 NA19077	5	5/5	0	5/5
TPMT*2	HG01605	1	1/1	0	1/1
TPMT*3C	NA18908 NA19114 NA19323 NA19452	4	4/4	0	4/4
UGT1A1*28	NA18510 NA18908 NA19077 NA19323	4	4/4	0	4/4
UGT1A1*6	HG02141 HG02188 NA18617 NA18621 NA18626 NA18633 NA18740	7	7/7	0	7/7

Conclusion: We have presented highly accurate calling (>99% concordance) of reportable AoUPGx alleles in both clinical specimens and reference cell lines. All reportable alleles have been observed at least once (except for G6PD Kambos and TPMT*3B which are very rare in the general population and no cell lines are available).

2.2.2.2 Precision

Part 1 – Inter- and intra- center equivalency

To assess equivalence of processing and variant calling across the three AoU GCs, we assessed concordance of variant calls from five donor-derived blood specimens collected and processed according to the AoURP Protocol. DNA was extracted with two methods (Autogen [Holliston, MA] and Chemagen [PerkinElmer Baesweiler, Germany]), using both whole blood and white-blood cell fractions (buffy coat) from these five individuals. Replicate samples were run at each lab and the equivalence between replicates was determined to demonstrate that the variability in sample processing and variant calling between labs is no greater than the variability within labs (**Table 11**). Overall equivalence was calculated using the Jaccard similarity coefficient between each pair of labs over all variants (the size of the intersection of the calls divided by the size of the union of the calls).

Table 11. Overall equivalence of called variants in donor blood samples across AoU GCs

Comparison (Donors)	Pairwise Group	SNP Concordant	SNP Discordant	InDel Concordant	InDel Discordant	Overall Equivalence [95%CI]
Interlab	BCM-UW	3830/3833	3/3833	189/195	6/195	99.77% [99.55%-99.99%]
	BI-BCM	3830/3834	4/3834	191/193	2/193	99.86% [99.71%-100%]
	UW-BI	3831/3834	3/3834	188/196	8/196	99.71% [99.46%-99.96%]
Intralab	BI-BI	1708/1709	1/1709	87/89	2/89	99.82% [99.57%-100%]
	BCM-BCM	1246/1248	2/1248	62/62	0/62	99.76% [99.32%-100%]
	UW-UW	1577/1578	1/1578	76/83	7/83	99.49% [99.02%-99.46%]

We further evaluated equivalence across and between GCs using the WGS data from 175 human cell line derived genomic DNA samples that were part of the PGx accuracy and NIST accuracy studies (**Table 12**).

Table 12. Overall equivalence of called variants in cell lines across AoU GCs

Comparison (Cell Lines)	Pairwise Group	SNP Concordant	SNP Discordant	InDel Concordant	InDel Discordant	Overall Equivalence [95%CI]
Interlab	BCM-UW	24289/24304	15/24304	1142/1178	36/1178	99.80% [99.71%-99.89%]
	BI-BCM	24288/24304	16/24304	1151/1171	20/1171	99.86% [99.79%-99.93%]
	UW-BI	24287/24302	15/24302	1145/1177	32/1177	99.81% [99.73%-99.90%]
Intralab	BI-BI	1175/1175	0/1175	56/65	9/65	99.27% [99.68%-99.86%]
	BCM-BCM	806/806	0/806	38/40	2/40	99.76% [99.32%-100%]
	UW-UW	918/918	0/918	50/50	0/50	100% [100%-100%]

Conclusion: The Genome Centers exhibited a high equivalence (>99%) based on variants called from a common set of blood and cell line samples. In addition, the variability within replicates at the same GC is generally the same or greater than the variability across the three

GCs. Any differences that exist between the GCs are not a significant source of variability for variant calling and, therefore, a participant will not experience a difference in analytical risk based on which GC processed the sample.

Part 2 – Inter- and intra-center precision - discordance by context

Study 1 - Twenty replicate samples from five individuals were examined. Clinical Panel testing of these samples was used to define the 'truth'. The panel is adapted from the multi-gene NGS panel Color test described in Neben, Zimmer, Stedden, et al., 2019 (*https://pubmed.ncbi.nlm.nih.gov/31201024/*) and Berger, Williams, Barrett, Zimmer, et al., 2020 (*https://www.biorxiv.org/content/10.1101/2020.01.15.907212v1*). Laboratory procedures, bioinformatics analysis, and variant interpretation are performed at Color under Clinical Laboratory Improvement Amendments (CLIA) (#05D2081492) and College of American Pathologists (CAP) (#8975161) compliance. Analysis, variant calling, and reporting focus on the complete coding sequence and adjacent intronic sequence of the primary transcript(s), unless otherwise indicated. Variant calls from WGS on all 20 samples at each GC were compared to the panel variants to determine concordance across sites by genomic context (**Table 13**). Across all samples we identified 7892 SNPs (4874 Heterozygous [Het], 3018 Homozygous [Hom]), 119 Insertions (90 Het, 29 Hom, size range 1-3, mean 1.7), and 129 Deletions (45 Het, 84 Hom, size range 1-10, mean 2.9).

Table 13. Concordance of variant calling by genomic context in donor blood samples

Category	Genome Center	Panel+/ WGS-	Panel-/ WGS+	Panel+/ WGS+	Panel-/ WGS-	PPA [95% CI]	NPA [95% CI]
SNPs	BCM	7	13	2609	4187271	99.74% [99.6%-99.9%]	100% [100%-100%]
	BI	8	14	2610	4187268	99.71% [99.5%-99.9%]	100% [100%-100%]
	UW	8	13	2610	4187269	99.71% [99.5%-99.9%]	100% [100%-100%]
Insertions	BCM	1	20	20	4189859	97.22% [91.8%-100%]	100% [100%-100%]
	BI	0	20	20	4189860	100% [100%-100%]	100% [100%-100%]
	UW	2	20	20	4189858	94.44% [87.1%-100%]	100% [100%-100%]
Deletions	BCM	1	0	42	4189857	98.33% [95.1%-100%]	100% [100%-100%]

Category	Genome Center	Panel+/ WGS-	Panel-/ WGS+	Panel+/ WGS+	Panel-/ WGS-	PPA [95% CI]	NPA [95% CI]
	BI	1	0	42	4189857	98.33% [95.1%-100%]	100% [100%-100%]
	UW	1	0	42	4189857	98.33% [95.1%-100%]	100% [100%-100%]
SegDup	BCM	1	0	240	4189659	99.29% [97.9%-100%]	100% [100%-100%]
	BI	1	0	240	4189659	99.29% [97.9%-100%]	100% [100%-100%]
	UW	1	0	240	4189659	99.29% [97.9%-100%]	100% [100%-100%]
LowMap	BCM	1	0	24	3770885	97.22% [91.8%-100%]	100% [100%-100%]
	BI	1	0	24	3770885	97.22% [91.8%-100%]	100% [100%-100%]
	UW	1	0	24	3770885	97.22% [91.8%-100%]	100% [100%-100%]
LowComplexity	BCM	0	0	36	4189864	100% [100%-100%]	100% [100%-100%]
	BI	0	0	36	4189864	100% [100%-100%]	100% [100%-100%]
	UW	0	0	36	4189864	100% [100%-100%]	100% [100%-100%]
LowGC	BCM	0	0	20	1675940	100% [100%-100%]	100% [100%-100%]
	BI	0	0	20	1675940	100% [100%-100%]	100% [100%-100%]
	UW	0	0	20	1675940	100% [100%-100%]	100% [100%-100%]
Heterozygous	BCM	11	15	1648	4188226	99.37%	100%

Category	Genome Center	Panel+/ WGS-	Panel-/ WGS+	Panel+/ WGS+	Panel-/ WGS-	PPA [95% CI]	NPA [95% CI]
Variants						[99.0%-99.7%]	[100%-100%]
	BI	13	16	1648	4188223	99.27% [98.9%-99.7%]	100% [100%-100%]
	UW	13	15	1648	4188224	99.27% [98.9%-99.7%]	100% [100%-100%]
Homozygous Variants	BCM	1	18	1025	4188856	99.92% [99.8%-100%]	100% [100%-100%]
	BI	0	18	1026	4188856	100% [100%-100%]	100% [100%-100%]
	UW	0	18	1026	4188856	100% [100%-100%]	100% [100%-100%]

Conclusion: The centers display high performance and concordance across a range of G/C and different variant types in blood-derived samples.

Genome-wide reproducibility studies were completed for the specific 66 genes included in this device. A total of 177 variants (84 common to both samples, 47 unique for NA12878, 46 unique for NA24385) across 44 genes (42 HDR and 2 PGX) are represented in these two control samples. Also for this analysis, we restricted the precision comparison to the NIST high confidence regions. An additional comparison of each replicate to NIST also confirmed these variant calls as TP without FN variant sites and one FP site for InterRun 7. Below are the list of genes with variants included in this Intra- and Inter-run analysis.

Study 2 - Human cell lines (used in the pathogenic variant accuracy assessment) were sequenced utilizing the clinical NGS panel described in Precision Part 2, Study 1 to define the 'truth.' Calls from WGS on all 30 samples at each GC were compared to the panel variants to determine concordance across sites by genomic context (**Table 14**). Across all samples, we identified 11020 SNPs (6749 Het, 4271 Hom), 136 Insertions (95 Het, 41 Hom, size range 1-5, mean 2.0), and 224 Deletions (141 Het, 83 Hom, size range 1-13, mean 2.7).

Table 14. Concordance of variant calling by genomic context in cell lines

Category	Genome Center	Panel+/ WGS-	Panel-/ WGS+	Panel+/ WGS+	Panel-/ WGS-	PPA [95% CI]	NPA [95% CI]
SNPs	BCM	24	1	3558	6071772	99.34% [99.0%-99.7%]	100% [100%-100%]

Category	Genome Center	Panel+/ WGS-	Panel-/ WGS+	Panel+/ WGS+	Panel-/ WGS-	PPA [95% CI]	NPA [95% CI]
	BI	24	1	3694	6281131	99.36% [99.1%-99.7%]	100% [100%-100%]
	UW	24	0	3694	6281132	99.36% [99.1%-99.7%]	100% [100%-100%]
Insertions	BCM	1	23	21	5237330	96.70% [90.1%-100%]	100% [100%-100%]
	BI	1	22	22	5446825	96.88% [90.8%-100%]	100% [100%-100%]
	UW	1	23	22	5446824	96.88% [90.8%-100%]	100% [100%-100%]
Deletions	BCM	1	9	63	5865787	99.11% [97.4%-100%]	100% [100%-100%]
	BI	1	9	65	6075280	99.14% [97.4%-100%]	100% [100%-100%]
	UW	1	10	65	6075279	99.14% [97.4%-100%]	100% [100%-100%]
SegDup	BCM	0	0	309	6075046	100% [100%-100%]	100% [100%-100%]
	BI	0	1	322	6284527	100% [100%-100%]	100% [100%-100%]
	UW	0	0	322	6284528	100% [100%-100%]	100% [100%-100%]
LowMap	BCM	0	0	16	2932915	100% [100%-100%]	100% [100%-100%]
	BI	0	0	18	3142407	100% [100%-100%]	100% [100%-100%]
	UW	0	0	18	3142407	100% [100%-100%]	100% [100%-100%]
LowComplexity	BCM	0	11	54	6075290	100% [100%-100%]	100% [100%-100%]
	BI	0	11	54	6075290	100% [100%-100%]	100% [100%-100%]
	UW	0	10	54	6075291	100% [100%-100%]	100% [100%-100%]

Category	Genome Center	Panel+/ WGS-	Panel-/ WGS+	Panel+/ WGS+	Panel-/ WGS-	PPA [95% CI]	NPA [95% CI]
LowGC	BCM	1	0	34	4399360	99.21% [97.7%-100%]	100% [100%-100%]
	BI	1	0	36	4608853	99.24% [97.8%-100%]	100% [100%-100%]
	UW	1	0	36	4608853	99.24% [97.8%-100%]	100% [100%-100%]
Heterozygous Variants	BCM	18	20	2225	6073092	99.22% [98.7%-99.7%]	100% [100%-100%]
	BI	18	20	2323	6282489	99.24% [98.8%-99.7%]	100% [100%-100%]
	UW	18	20	2323	6282489	99.24% [98.8%-99.7%]	100% [100%-100%]
Homozygous Variants	BCM	8	13	1417	6073917	99.52% [99.1%-100%]	100% [100%-100%]
	BI	8	12	1458	6283372	99.54% [99.1%-100%]	100% [100%-100%]
	UW	8	13	1458	6283371	99.54% [99.1%-100%]	100% [100%-100%]

Conclusion: The GCs display high performance and concordance across a range of genomic contexts and variant types in cell line samples.

Part 3 – Equivalence of cell lines and clinical samples

To demonstrate the equivalence of cell line derived DNA with that of clinical samples we present a summary of both performance measures and technical measures for selected assessments (**Table 15**). This table summarizes assessments from other sections within the analysis as a means to compare the results across clinical samples and cell lines. For example, PGx accuracy lists the results for PGX accuracy in clinical samples (Part 5 Table 8) and cell lines (Part 5 Table 9). The P/LP Accuracy reflects the evaluations from Part 2 Table 3 (Row that lists accuracy of P/LP variants in clinical samples) and Part 3 Table 5 (P/LP variants in cell lines).

Table 15. Equivalence of performance measures and technical metrics between cell line-derived DNA and blood-derived DNA

Performance Measures

	Accuracy (N)	P/LP Accuracy	PGx Allele Accuracy	Precision (N)	Interlab Concordance (BCM-UW) [95% CI]	Interlab Concordance (BI-BCM) [95% CI]	Interlab Concordance (UW-BI) [95% CI]
Clinical Samples	271	100%	100%	28	99.77% [99.55%-99.99%]	99.86% [99.71%-100%]	99.71% [99.46%-99.96%]
Cell Line Samples	30	100%	100%	175	99.8% [99.71%-99.89%]	99.86% [99.79%-99.93%]	99.81% [99.73%-99.90%]

Technical Metrics

	Number of Samples	% Aligned Bases [std dev]	% Duplicate Reads [std dev]	Insert Size (bp) [std dev]	% Q30 Bases [std dev]	% Chimeric Reads [std dev]	Genome Coverage [std dev]	% Covered ≥20X [std dev]	% Contam-ination [std dev]	Ti/Tv Ratio [std dev]
Clinical Samples	253	92% [±0.94%]	11.15% [±1.78%]	421 bp [±16 bp]	92.2% [±0.77%]	1.81% [±0.24%]	39.99X [±7.3X]	96.55% [±0.49%]	0.03% [±0.33%]	1.94 [±0.01]
Cell Line Samples	223	91.25% [±0.93%]	11.14% [±1.4%]	419 bp [±11 bp]	91.43% [±1.03%]	1.69% [±0.16%]	39.53X [±3.13X]	96.07% [±0.24%]	0.02% [±0.07%]	1.93 [± 0.01]

The **Performance Measures** section has several sub sections:

1. Accuracy in P/LP variants reported in clinical samples and cell line samples. This comparison is restricted to the HDR genes (as PGx genes do not have P/LP variants in this study).
2. PGx allele accuracy reported in clinical samples and cell line samples. This comparison is restricted to the 7 PGx genes.
3. Interlab concordances for variants called in a common set of samples (either blood-derived or cell line). This analysis does include both HDR and PGx genes however the PGx variants contribute only a small fraction of the overall variation considered (2% in both the blood and cell comparisons).

The **Technical Measures** section reports on several performance measures that indicate performance at the whole genome level (% Aligned Bases, % Q30 bases, Genome Coverage, % Covered ≥20X), the library level (% Duplicate reads, Insert size, % Chimeric bases, % Contamination), or the variant level (Ti/Tv ratio). These measures are not specific to the HDR or PGx gene intervals.

Conclusion: The use of characterized cell lines can provide an excellent proxy for clinical samples as performance and technical measures in cell lines closely mirror those in clinical samples.

Reproducibility - Reproducibility (inter and intra run, operator, day, and instrument) has been established in each laboratory. In all cases the laboratories utilize reference samples (NA12878 and other reference samples with truth data) with established truth data to assess reproducibility based on concordance of sample genotypes. In all cases inter and intra run precision are greater than 99%.

Reagent Lot management and qualification is performed as a matter of course and policy in each group. Each new lot, or new shipment of an existing lot, is tested for equivalent performance as part of controlled release into the production process. This includes testing of control samples (NA12878) for comparison of the results with the reference range defined for the test. Additionally assessed are the metrics for coverage and contamination.

Sequencing instruments and other primary instruments that contribute to a key role in the clinical test (NovaSeq 6000 sequencers and GC specific robotics platforms) are validated as part of clinical operations by using cell lines and other control samples for comparison between independent runs for inter-run reproducibility determination and/or within the same time for intra-run reproducibility determination.

Part 4 – Precision of AoUPGx calling

Precision of AoUPGx variant calling was assessed by processing 62 cell lines with known PGx alleles, as defined by Stargazer (Lee SB, et al., 2019), at each of the three AoURP GCs. Results are shown in **Table 16**. Samples were sequenced and star alleles were called for the samples according to the device procedures. Concordance of the star allele calls were evaluated across the centers and shown to be highly concordant (Table 16. - 298/300, 99.3%). We have selected a subset (28 out of 47) of alleles for our concordance validation across the three centers (Table 16 in IDE) that represent the allele frequency spectrum of Rare (<1% AF on avg), Somewhat common (1-10% AF on avg), and Common (>10% AF on avg), we expect to see in the AoU program. In addition to capturing the allele frequency spectrum, these cell lines represent all the various types of variants (SNP, Multiple SNPs and indels) that define PGx star alleles that will be reported by the AoURP.

As shown in Table 16 below, two GCs observed incorrect calls in a single sample (NA19226) due to a missing ploidy call in this case. Ploidy is necessary for automated zygosity determination in X-linked genes. In the actual study, samples with unknown ploidy will be flagged for manual review by the clinical director for resolution.

Table 16. Concordance of AoUPGx calling across AoU GCs

Gene and Star Allele	Coriell PGx Sample	BCM Correct	BCM Incorrect Call	UW Correct	UW Incorrect Call	Color Correct	Color Incorrect Call	Overall Concordance
DPYD c.1679T>G(*13)	HG00332	1/1	0	1/1	0	1/1	0	3/3
G6PD Kalyan-Kerala, Jamnaga, Rohini	NA20853	1/1	0	1/1	0	1/1	0	3/3
CYP2C19*17	HG02511 HG00332 NA19323 NA20337	4/4	0	4/4	0	4/4	0	12/12
UGT1A1*37	HG02511 NA12375 NA17102	3/3	0	3/3	0	3/3	0	9/9
UGT1A1*6	NA17012 NA23093 NA18617 NA18572	4/4	0	4/4	0	4/4	0	12/12
SLCO1B1*17	HG00332 HG01808	2/2	0	2/2	0	2/2	0	6/6
DPYD c.1129-5923C>G	NA17227 NA24217	2/2	0	2/2	0	2/2	0	6/6
G6PD Ube Konan	NA18998	1/1	0	1/1	0	1/1	0	3/3
G6PD A-968C_376G	HG02982 HG01083	2/2	0	2/2	0	2/2	0	6/6
TPMT*2	NA20337 NA20336 HG01083	3/3	0	3/3	0	3/3	0	9/9
DPYD c.2846A>T	NA07048 NA06991 NA07055	3/3	0	3/3	0	3/3	0	9/9
G6PD Union, Maewo, Chinese-2, Kalo	HG03756	1/1	0	1/1	0	1/1	0	3/3
CYP2C19*2	HG02131 HG02982 HG02511 HG02682 HG03756 HG02373 NA18998 HG00978 HG02649 HG01083	10/10	0	10/10	0	10/10	0	30/30
G6PD Mediterranean, Dallas, Panama, Sassari, Cagliari, Birmingham	HG02682 HG02649	2/2	0	2/2	0	2/2	0	6/6
G6PD A-202A_376G/A-202A_376G	NA19147 NA18484	2/2	0	2/2	0	2/2	0	6/6
G6PD Canton, Taiwan-Hakka, Gifu-like, Agrigento-like	HG00978 NA18617	2/2	0	2/2	0	2/2	0	6/6
NUDT15*2	NA18945 NA17660 NA18526	3/3	0	3/3	0	3/3	0	9/9
G6PD Viangchan, Jammu	HG02131 HG02373	2/2	0	2/2	0	2/2	0	6/6
G6PD Sibari	NA19323	1/1	0	1/1	0	1/1	0	3/3
UGT1A1*28	NA21781 NA07048 HG02131 HG02682 NA19917 HG00978 HG01800 NA19323 NA20336 NA06993 NA10846 NA10855 NA10865 NA12892 NA17061 NA17288 NA18563 NA19035 NA23296 NA23297 NA23872 NA23878 NA19785 NA18572	24/24	0	24/24	0	24/24	0	72/72
NUDT15*3	NA18998 NA18564 NA18992 HG01808	4/4	0	4/4	0	4/4	0	12/12
UGT1A1*28/*28	HG03756 HG02649 NA20337 NA18873 NA19238 NA23405 NA24027	7/7	0	7/7	0	7/7	0	21/21
G6PD Kaiping, Anant, Dhon, Sapporo-like, Wosera	HG01800 HG01808	2/2	0	2/2	0	2/2	0	6/6
G6PD Quing Yuan, Chinese-4	HG00654	1/1	0	1/1	0	1/1	0	3/3
G6PD Asahi	HG02511	1/1	0	1/1	0	1/1	0	3/3
CYP2C19*17/*17	NA20336	1/1	0	1/1	0	1/1	0	3/3
G6PD A-202A_376G	NA18861 NA19143 NA19226* NA19920 NA19917	4/5	1	4/5	1	5/5	0	13/15
CYP2C19*35	NA19122	1/1	0	1/1	0	1/1	0	3/3
SLCO1B1*15	HG00654 HG02649	2/2	0	2/2	0	2/2	0	6/6
CYP2C19*3	HG00654	1/1	0	1/1	0	1/1	0	3/3
TPMT*3C	HG00978 NA19323	2/2	0	2/2	0	2/2	0	6/6

Conclusion: AoUPGx calling from WGS analysis on a common set of samples across GCs is highly concordant (298/300, 99.3%).

Part 5 – Quality Metrics Selection

To define threshold statistics for ongoing acceptance of sample data, we simulated a set of WGS data of differing qualities tailored to the metric being evaluated in order to define thresholds for acceptance criteria.

- For the **coverage** metrics, we simulated samples with less data by using the Picard tool (Van der Auwera 2013) to uniformly remove read data from four NIST control samples

for which gold-standard variant data is available, to assess the impact of lower coverage on variant calling.

- For **contamination**, we simulated samples that were contaminated at various known levels by combining read data from a second sample to read data control sample, at progressively higher levels.
- For the **percent duplicate reads comparison** we simulated samples with redundant data by progressively adding more duplicates to a sample while holding the overall amount of data generated for that sample constant. We processed these samples with the harmonized Illumina Dynamic Read Analysis for GENomics (DRAGEN) pipeline to perform mapping, alignment, and variant calling, and assessed FP and FN variant calls. Figures 4 to 10 show the relationship between the metrics evaluated and data quality, with a red bar to mark the threshold. In general, the logic for our metric threshold selection was to identify an inflection point and select a threshold with better performance than that inflection point.

Conclusion: The following six metrics have been defined by the program as sample-level acceptance criteria:

- **Mean Coverage (threshold ≥ 30X)** - The total FPs and FNs show a gradual increase as mean coverage decreases, with a rapid increase below 20x coverage, supporting a stringent threshold selection of a minimum of 30x (**Figure 4**).
- **Genome Coverage (threshold ≥ 90% at 20X)** - The total FPs and FNs steadily increase as the percent of bases with at least 20x coverage drops. Drop-off of performance is initially gradual, supporting a threshold of 90% (**Figure 5**).
- **AoUHDR Coverage (threshold ≥95% at 20X)** - The total FPs and FNs increase gradually as the average coverage in AoUHDR regions decreases. The reduction in performance is slow initially, and then increases rapidly below 40%, showing that the genome center threshold of 95% is conservative (**Figure 6**).
- **Aligned Q30 Bases (threshold ≥8e10)** - FPs and FNs increase with lower base quality counts, with inflection points starting around 6e10 for both (**Figure 7**).
- **Contamination (threshold ≤1%)** - Variant calling performance as measured by both the number of FPs and the number of FNs decreases with increasing contamination, demonstrating that the 1% threshold is appropriate (**Figure 8**).

A. Mean Coverage (Threshold ≥30X)

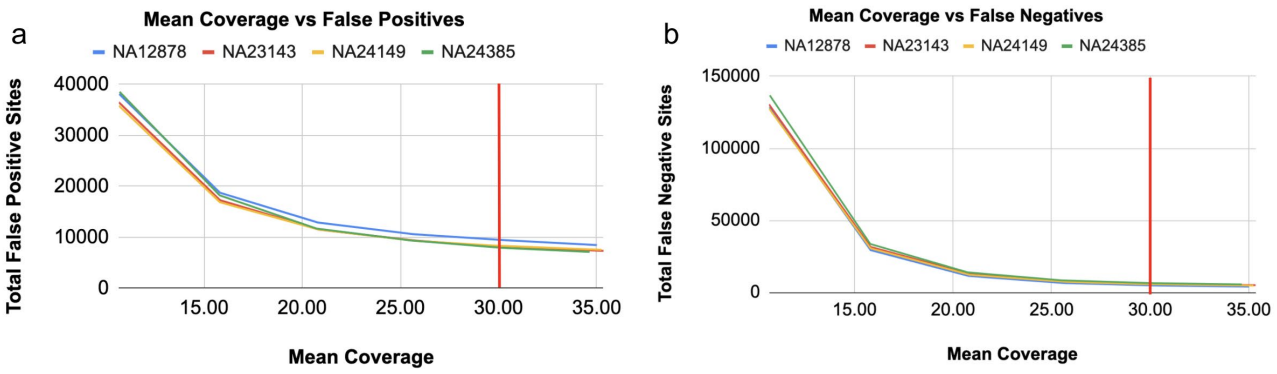


Figure 4. Relationship between mean coverage and performance (FP [a] and FN variant counts [b]).

B. Genome Coverage (Threshold ≥90% at 20X)

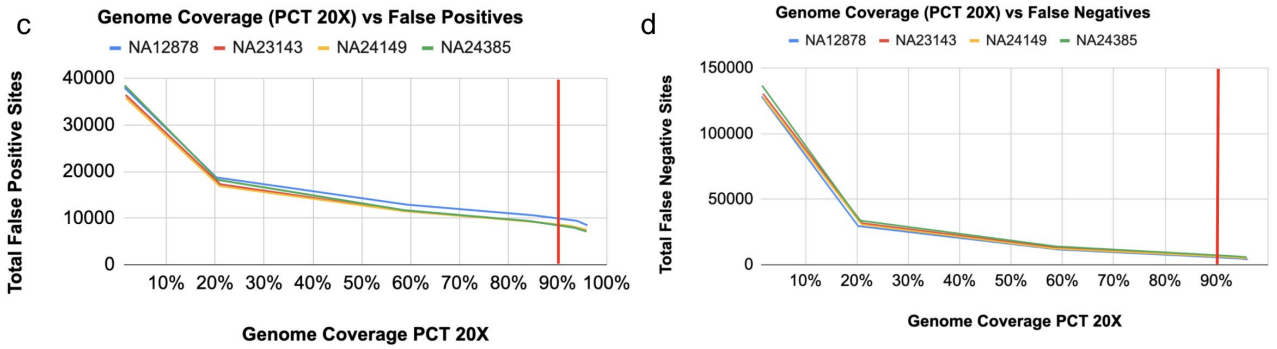


Figure 5. Relationship between genome coverage (% of the genome covered to ≥20X) and performance (FP [c] and FN variant counts [d]).

C. AoUHDR Coverage (Threshold ≥95% at 20X)

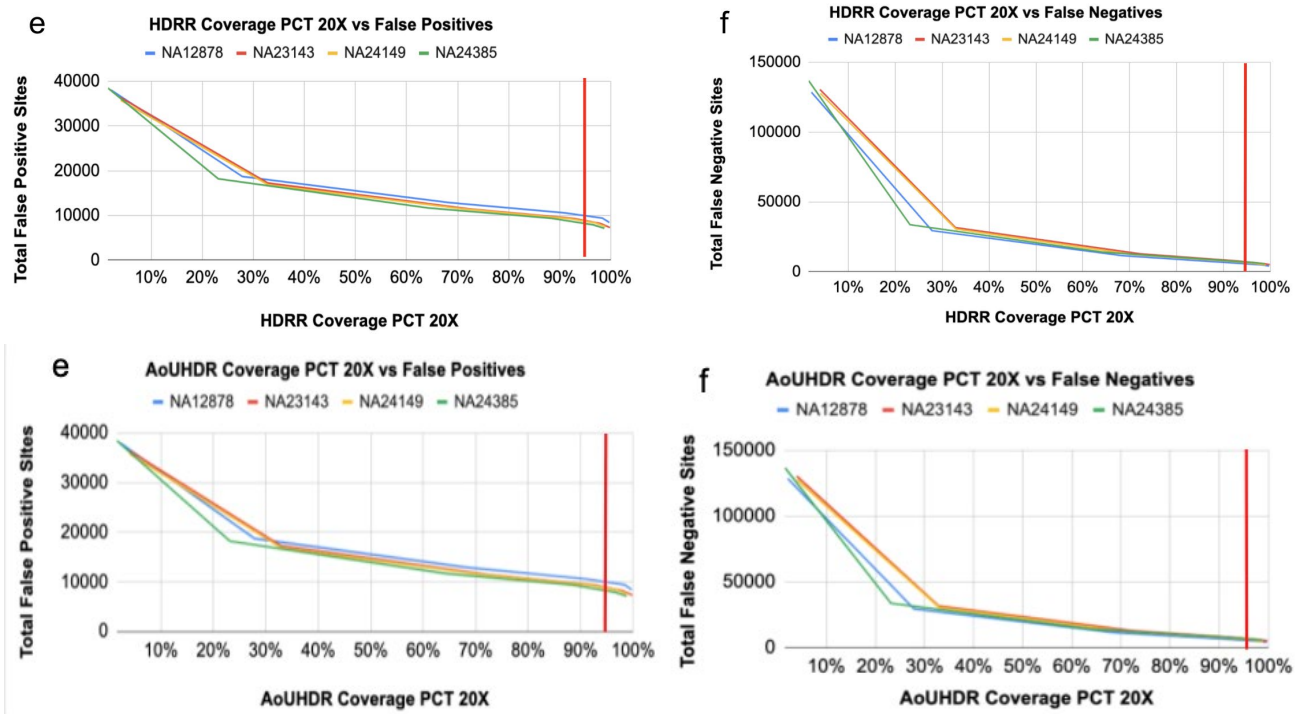


Figure 6. Relationship between AoUHDR coverage (% of the AoUHDR region covered to ≥20X) and performance (FP [e] and FN variant counts [f]).

D. Aligned Q30 Bases (Threshold ≥8e10)

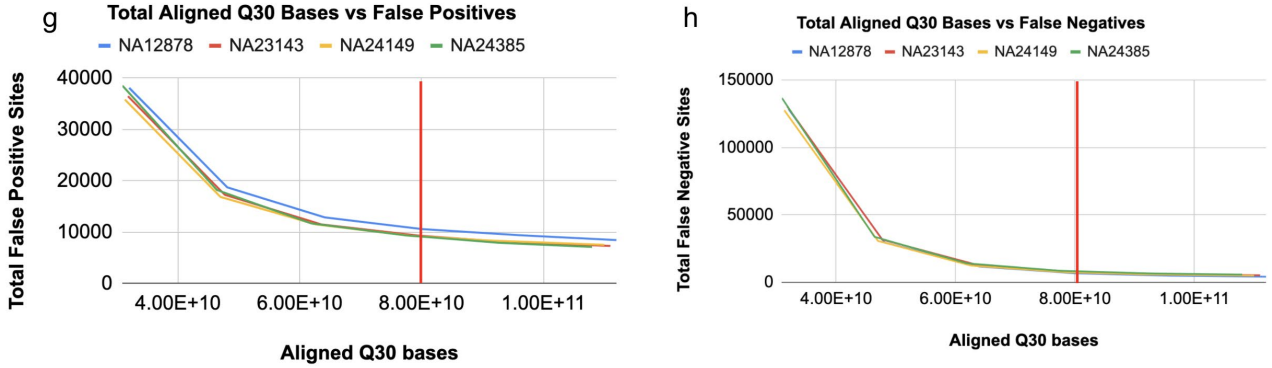


Figure 7. Relationship between Aligned Q30 bases and performance (FP [g] and FN variant counts [h]).

E. Contamination (Threshold ≤ 1%)

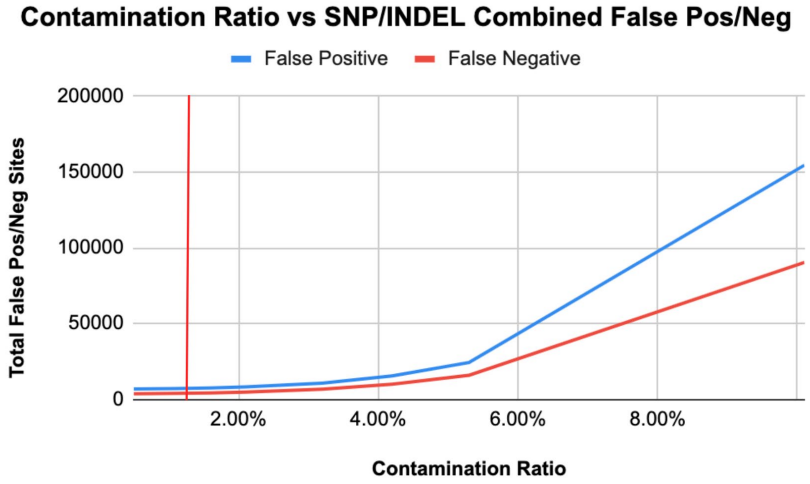


Figure 8. Relationship between estimated sample contamination and performance (FP and FN variant counts).

F. Duplicate Rate (Threshold ≤ 15%)

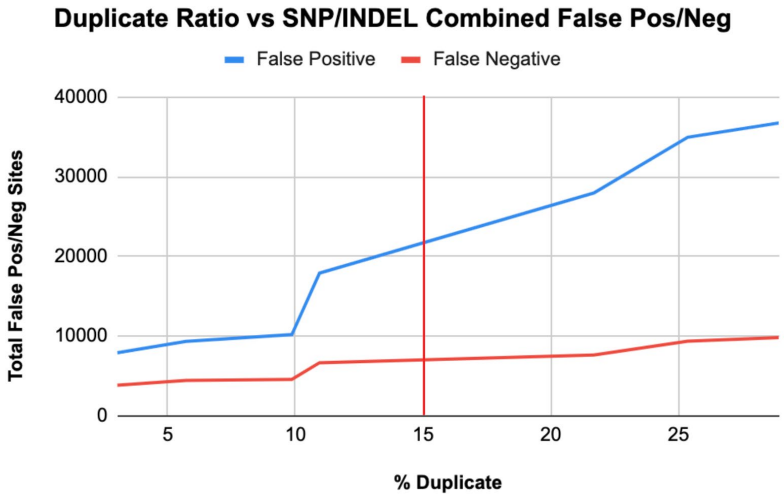


Figure 9. Relationship between duplicate rate and performance (FP and FN variant counts).

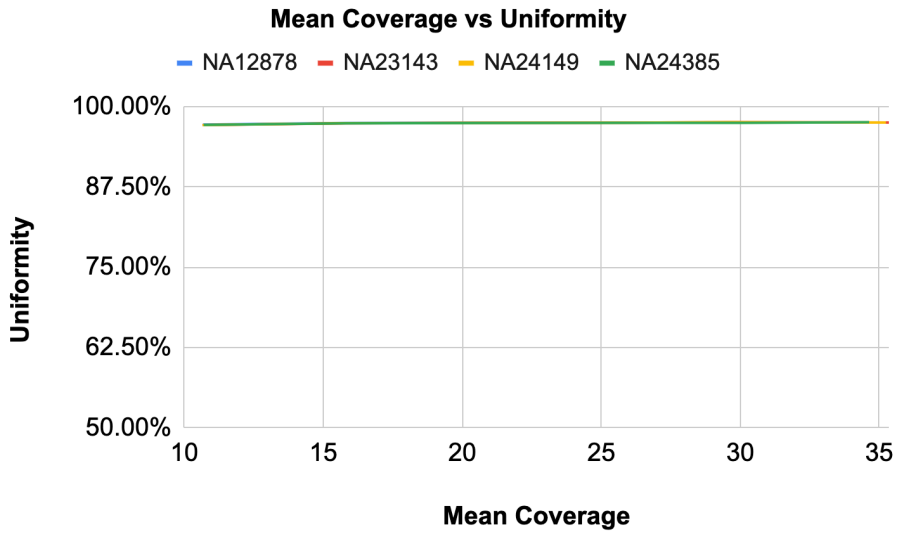


Figure 10. Relationship between mean coverage and uniformity.

Previous FDA feedback (Q190789 pre-Sub meeting; Appendix 16) suggested several additional metrics that are not used routinely as sample-level acceptance criteria. These have each been considered and described below. Although these metrics are not used as formal acceptance criteria for the IDE, many are either directly or indirectly used for sample or variant call acceptance criteria.

- **Percent Duplicate Reads** – A sample with a high percentage of duplicate reads will also have lower coverage because the pipeline includes duplicate marking which excludes duplicate reads from coverage calculations. Thus, samples with a high duplicate percentage may fail one or more of the genome coverage, mean coverage, or AoUHDR coverage metrics. A high percentage of duplicates does correlate to lower data quality. Therefore, samples are routinely rejected if the duplicate rate is greater than 15%. (**Figure 9**).
- **Median Insert Size** - The median insert size (i.e., the inner length of DNA between the sequencing adapters) is routinely assessed as part of clinical operations. Although not used as sample-level acceptance criteria, it can be useful for troubleshooting.
- **Sequence Length Distribution** - The sequence length distribution is not assessed bioinformatically, but DNA fragment size is assessed during the library preparation step using the Covaris instrument.
- **Uniformity** - The uniformity metric (which the GCs will monitor, but not use as acceptance criteria for samples) does not deviate as data quality is reduced, which supports its non-usage as acceptance criteria (**Figure 10**). In addition, large deviations in uniformity will impact the genome coverage metric.
- **Percent of PF (Pass-filter) Reads** - This metric is monitored during runs of the sequencing instruments, as it can indicate that a run will not meet the required yield. However, the percent of PF reads is a direct reflection of the amount of data that a

sequencing run will produce, and issues will be detected by coverage metrics such as the mean, genome, and AoUHDR coverage metrics.

- **GC Content** – GC content is evaluated during assay setup when assessing the reportable range. Regions with poor mappability will be excluded (See also Section 2.2.2.5 Reportable Range).
- **Cluster Density and % Passing Clusters Passing Filter** - These metrics are not applicable to the current assay due to the use of patterned flow cells on the Illumina NovaSeq instrument.
- **Percent Total Reads after Trimming** - This metric is not applicable to the current assay because trimming is not part of the protocol.
- **Percent Reads Mapped on Target** - In this whole genome sequencing assay, there is no capture target, so this metric is not applicable.
- **Strand Bias** - Although strand bias is not tracked on a per-sample level, it does form an integral part of variant calling. Variants receive a lower quality score if they show a particular stand bias.
- **Number of Reads Required for Exons** - Coverage in specific regions is assessed in multiple ways. First, the percent of bases at 20X or higher metric (i.e., genome coverage) is sensitive to localized regions that fall below acceptable coverage. Second, the AoUHDR coverage metric monitors whether coverage is acceptable within the reportable range.
- **Number of Reads Required for Variants** - The number of reads that supports a candidate haplotype are used by the DRAGEN variant calling model to assess confidence in that variant call.
- **% Unassigned Read Indices** - The number of unassigned reads is tracked as part of demultiplexing and can be used to troubleshoot low-coverage samples, but is not used as sample acceptance criteria.
- **Percent Reads for Non-Human DNA** - This metric is assessed by looking at the inverse of the percentage of aligned reads and large levels of contamination may impact other coverage metrics. It can be useful for troubleshooting low coverage samples, but is not used as sample acceptance criteria.
- **Global Imbalance Value (GIV) (G->T/C->A) Score** - This score, which is designed to detect DNA damage, is not routinely tracked. However, DNA damage is detected by pre-analytical Quality Control (QC) steps and false positives related to DNA damage can be detected during confirmatory testing in the CVL.

Conclusion: The metric thresholds selected for quality monitoring of AoURP samples are supported by the observed performance of NIST control samples. Additionally, explanations are given as to why previously suggested metrics are not appropriate for this assay.

Part 6 – Validation of multiplexing and barcoding

During library construction, each individual sample is labeled with unique dual-index molecular barcodes to allow for multiplex sequencing and analysis. The libraries are pooled together for sequencing, and the analysis pipeline de-multiplexes the data so that it can be aggregated and aligned by individual sample based on the molecular barcode index. Sample Identity QC and read assignment accuracy can be checked to monitor multiplexing and molecular barcode indexing.

Each GC has designed or has employed commercially available (Illumina Tech Note. 2017) 96 unique dual-index barcodes that are 8 or 10 base pairs in length. The barcodes are designed with an edit distance of 3, meaning that each index sequence can tolerate 3 errors before the index could potentially be incorrectly assigned to a different sample. The adapters contain different sequences, allowing for even more stringent control of index assignment errors, as well as detection of chimeric molecules. In addition, these unique dual index barcodes allow sequence reads with conflicting barcodes on either Read 1 or Read 2 to be discarded during data analysis. This measure therefore eliminates low level sample read contamination (barcode swapping) issues that occur on patterned flow cells on the NovaSeq instrument when pooling libraries with single indices during sequencing (Illumina Tech Note. 2017, Costello, Fleharty, Abreu, et al., 2018). Each barcoded library is quantified using quantitative polymerase chain reaction (qPCR) assays and normalized prior to pooling for multiplexed library sequencing.

Barcode Lot QC: Adapter oligonucleotides are ordered High Performance Liquid Chromatography- (HPLC-) purified in order to minimize non-conforming sequences (Integrated DNA Technologies and Illumina Tech Note. 2017). Barcode lots from all groups are functionally tested prior to use by generating test libraries (WGS or amplicon) to determine performance for library yield and contamination with other synthesized barcodes. Each adapter set is sequenced, the index counts are monitored for underperforming adapters and the missed index file is examined for unmatched index sequences with high read counts. Any adapter with >1% contamination is flagged, and the barcode is re-tested with another sample. If it fails the second test, the barcode is quarantined. Similar to contamination screening, quality of the new adapters for de-multiplexing through the pipeline is also monitored as incorrectly synthesized adapters can trigger barcode mismatch when reads are assigned to a barcode. Together, these two QC measures ensure that newly synthesized adapter lots are validated for production use.

Accuracy of Read Assignment: The DRAGEN cross-sample contamination module uses a probabilistic mixture model to estimate the fraction of reads in a sample that may be from another human source. This sample contamination fraction is estimated as the parameter value in the mixture model that maximizes the likelihood of the observed reads at multiple pileup locations. The mixture model accounts for the population allele frequencies and the inferred sample genotypes. This value is provided as a fraction, so a value of 0.011 the same as 1.1% estimated contamination.

Conclusion: These dual-index molecular barcode designs and multiplexing protocols were utilized for all AoU clinical samples at each of the GC's. The overall performance of the barcoding methodology is reflected in the high sensitivity, specificity, precision, limit of detection and equivalency between centers described in prior sections.

2.2.2.3 Limit of Detection

Part 1 – Performance as a function of DNA input

To determine the range of acceptable genomic DNA inputs into library construction, an input titration experiment using DNA derived from NA12878 with total input amounts of 25 ng, 100 ng, 250 ng, 375 ng, 400 ng, 500 ng, 600 ng, 750 ng, 1000 ng and 1500 ng into library construction was performed. All three of the 25 ng replicates failed to successfully produce a library. To assess the effect of lower input amounts on sensitivity and to confirm that the minimum input identified produces acceptable sensitivity and precision, the vcfeval tool (Cleary JG, et al. 2015) was leveraged to calculate sensitivity vs NIST for each titration point (**Figure 11**). Input amounts below 250 ng show reduced sensitivity. As such, to meet target quality metrics, we concluded that the minimum amount of input DNA into library construction for sequencing is 250 ng. Equivalent performance was found with similar titration done with clinical samples. Four donor blood samples were titrated across an input range from 25 ng to 1500 ng to library construction (see **Table 17**).

Figure 11. Analytical sensitivity for NA12878 input titration series vs NIST across all three GCs.

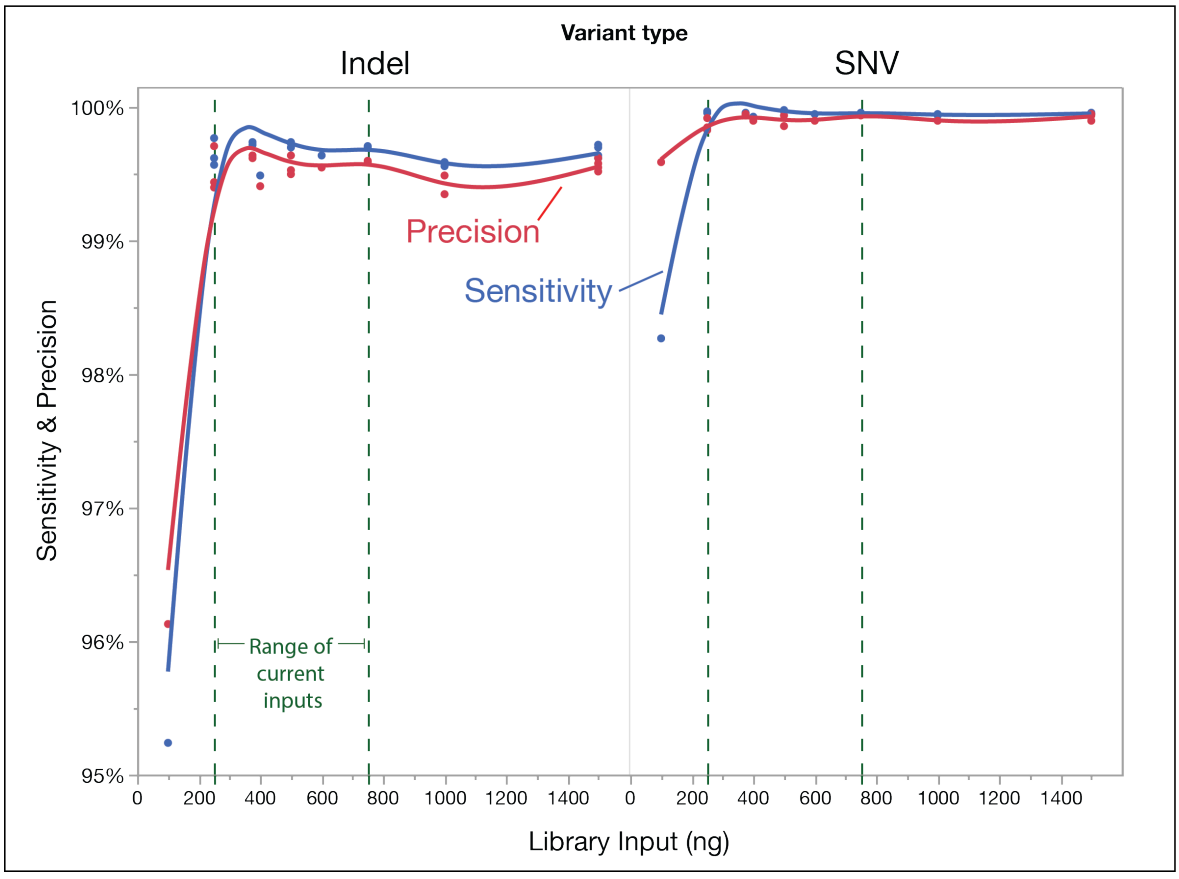


Table 17. Input titration results for four blood donor samples (performance of SNVs shown).

Input (ng)	Library Construction Success Rate	Mean Sensitivity	Std Dev Sensitivity	Mean Precision	Std Dev Precision
25	0%	-	-	-	-

100	17%	99%		100%	
250	100%	100%	0%	100%	0%
375	100%	98%	1%	99%	0%
500	100%	99%	1%	99%	1%
750	100%	98%	1%	99%	0%
1500	100%	98%	1%	98%	0%

Conclusion: Library construction success rates and variant calling performance achieve acceptable levels with a minimum DNA input of 250 ng. No significant difference in performance is seen between 250 ng and 1500 ng input DNA.

Part 2 – Performance as a function of allele fraction

To evaluate performance as a function of allele fraction, 7 replicates of NA12878 were sequenced with the *AoU* WGS pipeline and compared to results from the high confidence region of the GIAB v3.3.2 truth set. In **Figure 12** we show recall, precision, and the raw numbers of TP, FP, and FN variants binned by observed alternative allele fraction.

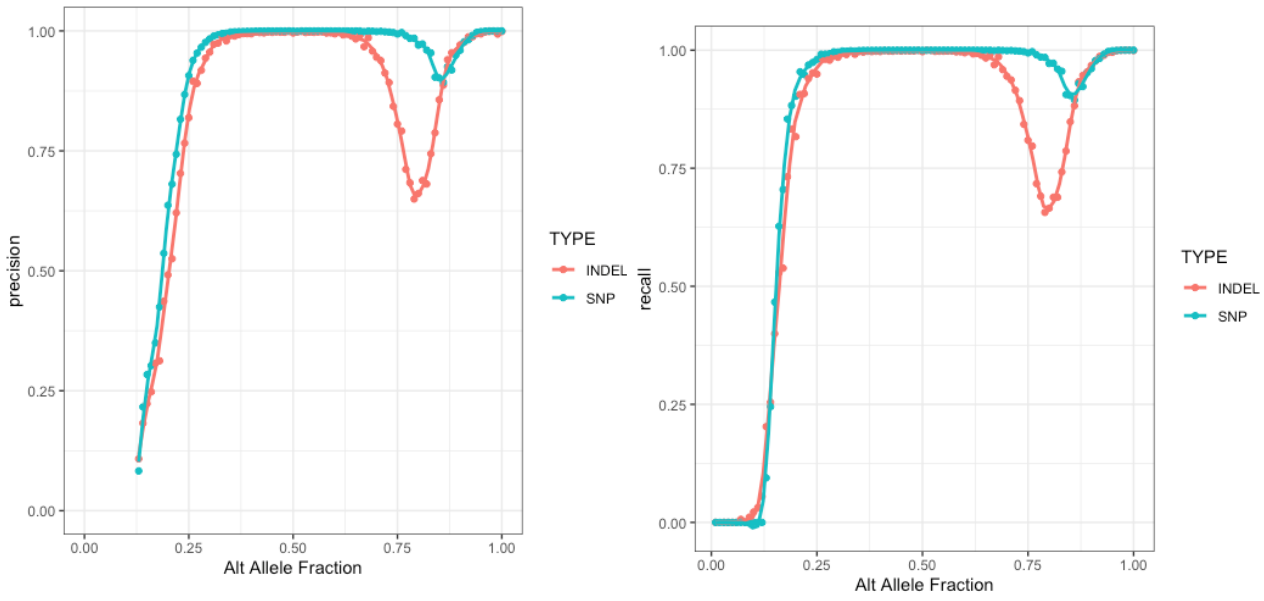


Figure 12. Precision and recall (*i.e.*, sensitivity) of SNV and indel calling as a function of alt allele fraction.

Conclusion: The analysis indicates that confident and accurate heterozygous calls are made between 30-75% allele fraction for SNVs and 30-65% allele fraction for indels.

2.2.2.4 Analytical Specificity – Interfering Substances

To establish the analytical specificity of the AoU WGS assay, other approved medical devices, such as K132750 were evaluated. K132750, the Illumina MiSeqDx Cystic Fibrosis Clinical Sequencing Assay is a targeted next generation sequencing assay which was performed on samples collected in K_2 Ethylenediaminetetraacetic acid (EDTA). In the validation summary, samples were tested with added exogenous and endogenous interfering substances as described below. All tested conditions gave 100% correct calls and no samples required a re-concentration to obtain the appropriate amount of DNA for the assay.

1) Interfering effects of **bilirubin, hemoglobin, cholesterol, and buffer component interference** on the performance of the assay were examined using 8 samples carrying 8 unique genotypes, and blood from the same individual was tested across all 4 interfering substances. Sample types included one PolyTG/PolyT variant, one indel (F508del), and 6 SNVs. Bilirubin (684 and 137 µmol/L), hemoglobin (2 and 0.4 g/L), and cholesterol (13 and 2.6 mmol/L) were spiked into blood aliquots prior to DNA extraction. Wash buffer from DNA extraction (15%) was spiked into genomic DNA samples prior to library preparation. For the assessment of each inhibiting substance, data for each spiked sample was compared to an untreated aliquot of the same blood/DNA sample. Impact on call rate, reproducibility, and sample first pass rate were determined. All 88 samples met the acceptance criteria of the test.
2) A study to assess the potential interference of triglycerides (37 mmol/L and 7.4 mmol/L) and high and low concentrations of K_2EDTA (7 mg/mL and 2.8 mg/mL) to mimic short blood draws was conducted. Eight whole blood samples were used for this study. Two samples were WT, two were F508del, and the remaining 4 were SNVs. With the exception of R75Q, all SNVs were replicates of those tested in Interference Study A (above), but were different blood samples. All tested conditions gave 100% correct calls, and no samples required a re-concentration to obtain the appropriate amount of DNA for the assay.

In addition, another FDA approved device, the FoundationOne CDx™ (RAL-0003-01), included results of an extensive interference study. In their Premarket Approval (PMA) submission, Foundation Medicine validated a total of 59 formalin-fixed paraffin-embedded (FFPE) samples with various potential interfering substances at various concentrations. Our main purpose is to demonstrate that even from FFPE, which is technically a much more challenging sample type, it was proven from Foundation's large interference study, none of the endogenous and exogenous substances demonstrated any inhibitory effects on the downstream procedures of DNA extraction, library prep and Next Generation Sequencing.

With these available evidence and data completed in specific interference studies as well as all the samples which were already completed by all the Genome Centers, there is no known interference substance which would cause inhibitory effects in our technology. Although these were FFPE samples, their study had evidently proven no interfering substances effect on the NGS assay even with the presence of exogenous (e.g., ethanol, proteinase K) and endogenous (e.g., melanin and molecular index barcodes) substances.

Conclusion: Interfering substance analyses of the specimen collection and DNA isolation methods have been performed previously and have demonstrated suitability for next generation sequencing.

2.2.2.5 Reportable Range

The AoUHDR interval list was created by first defining the relevant transcript(s) for each of the 59 genes of interest. Next, with regard to intronic padding surrounding the exons of interest, we decided to include up to the -15 intronic position upstream of the exon to the +6 intronic position downstream of the exon. Additional intervals were added to adequately cover P/LP variants in the genes of interest that fall outside the -15 to +6 regions. A participant will receive one report if they have P/LP variants in more than one HDR gene. The report will contain the details that are relevant for each of the gene(s) in which a P/LP variant has been identified. Lastly, the interval list includes 43 sites needed for accurate AoUPGx star allele reporting.

The following factors impact the sensitivity of variant detection for the Hereditary Disease Risk (HDR) panel and AoUPGx regions:

1. Some regions of the AoUHDR were determined to be technically challenging and have been excluded from the reportable range. First, PMS2 exons 12-15 cannot be analyzed using short-read sequencing technology due to their high homology with the PMS2CL pseudogene. The second category contains regions of high GC content (typically >75% across 100 bp) resulting in coverage dropout. The last category are regions with spurious variant calling artifacts due to the presence of micro-repeats (di-, tri-nucleotides) and long homopolymers.
2. Other regions were determined to not consistently meet a minimum quality standard defined as any site that did not have at least 20x coverage in 20% of the samples in the dataset. A per-site coverage analysis was performed across the entire range of AoUHDR and AoUPGx sites with a dataset of 104 samples from the 3 Centers; all with a whole-genome mean coverage of 30-35x. This analysis revealed 6 regions that included 56 bases across 4 genes that did not meet the 20x threshold (**Table 18**). In addition, a subset of blood samples was analyzed with the same minimum quality criteria set forth above. The results were similar to the analysis of the 104 samples reported above with even fewer regions failing the coverage criteria, showing consistency across the types of samples being collected.

Table 18. Frequently underperforming bases within the AoUHDR.

Gene	Total sites	GRCh37 low-coverage sites	GRCh38 low-coverage sites
MYH11	6699	8	8
MSH2	3148	14	14
KCNH2	3131	24	23
TSC1	3936	10	10

3. For these regions (Table 18), none were low enough quality to consider excluding the region from the reportable range; however, given the reduced quality, these regions will

be reported in our limitations section of the report to highlight the potential for reduced sensitivity.

4. For each patient analyzed, individual sites and regions can drop below the quality threshold. This will be handled by ensuring that each sample meets the minimum quality metrics defined for the overall assay before being passed on for analysis and reporting as well as stating the general limitations of WGS.

Conclusion: The AoUHDR and PGx reportable range covers the relevant transcripts for the 59 genes of interest, a few additional sites that fall outside the exonic regions and sites defining PGx star alleles, whereas technically challenging regions (e.g., pseudo genes and those with GC bias) have been excluded. Individual participant samples will be analyzed for <20x coverage across the reportable range, since we have identified six regions that did not consistently meet our quality threshold in the AoUHDR interval, but failed to warrant exclusion. These lower quality regions will be reported in the limitations section of the report to highlight the potential for reduced sensitivity.

2.2.2.6 DNA extraction performance

DNA is extracted from blood by two methods; A. salt-based precipitation method on Autogen FlexStar or B. a bead-based method on Chemagen 360. DNA samples are stored long term in a -80°C automated freezer.

DNA samples are checked for volume via a BioMicroLab volume check instrument. DNA samples are also quantified (spectrometric method) via Lunatic-Unchained Labs / Trinean DropSense 96 to obtain total DNA concentration as well as A260/280 and A260/230.

All samples must be within the following quality range:

a. minimum concentration = 50 ng/ul
b. A260/280 = 1.6-2.0

DNA samples are handled based on these criteria to determine pass/fail.

Based on data from 170,338 DNA extractions from the 4 mL whole blood tubes, and 2,859 DNA extractions from the 10 mL buffy coat tubes (all performed at the AoU Biobank on specimens collected between June 7, 2017 and August 31, 2019) the average DNA yield for whole blood is 109,091 ng with an A260/280 of 1.8. The average DNA yield for buffy coat is 139,889 ng with an A260/280 of 1.8. With a minimum input range for library construction of 250 ng to 750 ng, the extraction methods easily meet the requirements of the sample preparation process.

We assessed performance equivalency of two DNA extraction methods, Autogen and Chemagen, from both whole blood and buffy coat, using specimens from 5 donors. We present data demonstrating the equivalence of the Autogen and Chemagen extraction methods since both methods will be used in this study.

The DNA from these samples was sequenced using the AoU WGS assay and an orthogonal targeted panel assay. (**Table 19**).

Table 19. Performance of different extraction platforms and input material types.

Category	Type	Panel+/ WGS-	Panel-/ WGS+	Panel+/ WGS+	Panel-/ WGS-	PPA [95% CI]	NPA [95% CI]
Autogen	SNP	11	29	4340	3138045	99.8% [99.6%-99.9%]	100% [100%-100%]
Chemagen	SNP	9	11	3489	3138916	99.8% [99.6%-99.9%]	100% [100%-100%]
Autogen	INDEL	3	33	114	3142275	98.1% [96.1%-100%]	100% [100%-100%]
Chemagen	INDEL	0	26	72	3142327	100% [100%-100%]	100% [100%-100%]
WBC	SNP	11	11	3816	3138587	99.7% [99.5%-99.8%]	100% [100%-100%]
Whole Blood	SNP	12	29	4013	3138371	99.7% [99.5%-99.9%]	100% [100%-100%]
WBC	INDEL	3	29	90	3142303	98.5% [96.8%-100%]	100% [100%-100%]
Whole Blood	INDEL	0	30	96	3142299	100% [100%-100%]	100% [100%-100%]

Conclusion: All extraction platforms and input types assessed produced acceptable results.

2.2.2.7 Invalid rates

To illustrate fail rates at each step, we used historical data for the Genome Centers in addition to data from the samples run as part of the AoURP analytical validity cohort. CVL invalid rates were also calculated from the AoURP validation cohort for Color and UW. BCM data represents capillary sequence data from an internal cohort. For historical data, we used the National Heart, Lung, and Blood Institute (NHLBI) Trans-Omics for Precision Medicine (TOPMed) cohort as it represents a large number of samples run at all three AoU GCs (**Table 20**).

Table 20. Fail rates for samples at discrete parts of the process.

Genome Centers (WGS)

Lab	Cohort	Instrument	Sample Intake		Library		Sequencing		Data Quality		Total samples failed	Aggregate Sample Fail Rate
			Input #	# Failed (% of total Fails)	Input #	# Failed (% of total Fails)	Input #	# Failed (% of total Fails)	Input #	# Failed (% of total Fails)		
BCM	TOPMed	NovaSeq	1986	103 (90.4%)	1883	0 (0%)	1883	0 (0%)	1883	11 (9.6%)	114	5.74%
	AoURP	NovaSeq	584	0 (0%)	584	0 (0%)	584	0 (0%)	584	1 (100%)	1	0.17%
Broad	TOPMed	NovaSeq	6391	547 (99.5%)	5844	3 (0.5%)	5841	0 (0%)	5841	0 (0%)	550	8.61%
	AoURP	NovaSeq	424	3 (33.3%)	421	0 (0%)	421	0 (0%)	421	6 (66.7%)	9	2.12%
UW	TOPMed	HiSeqX	14587	1346 (88.3%)	13241	45 (3%)	13196	102 (6.7%)	13094	32 (2%)	1,525	10.45%
	AoURP	NovaSeq	291	0 (0%)	291	0 (0%)	291	2 (33.3%)	289	4 (66.7%)	6	2.06%

CVLs

Lab	Cohort	Instrument	Sample Intake		Library		Sequencing		Data Quality		Total samples failed	Aggregate Sample Fail Rate
			Input #	# Failed (% of total Fails)	Input #	# Failed (% of total Fails)	Input #	# Failed (% of total Fails)	Input #	# Failed (% of total Fails)		
BCM	Internal	ABI 3730/3500	1635	0 (0%)	NA	NA	1635	0 (0%)	1635	0 (0%)	10*	0.61%
Color	AoURP	NovaSeq	114	0 (0%)	114	1 (100%)	113	0 (0%)	113	0 (0%)	1	0.88%
UW	AoURP	NovaSeq	315	0 (0%)	315	0 (0%)	315	2 (50%)	313	2 (50%)	4	1.27%

*BCM sample fails are in amplicon generation.

Conclusion: We note that the incoming sample fail rate (i.e., samples that fail to meet acceptance criteria for quantity or quality upon receipt) represents a large number of the fails in the TOPMed cohort (on average 92.7%). The TOPMed study did not utilize a central biobank for sample extraction and QC. In addition, the research samples used in TOPMed varied in age and condition of storage. We anticipate that the CAP accredited biorepository at the Mayo Clinic being used by AoURP will dramatically improve success rates for incoming samples, as was evident in the samples received from them (donor blood samples and American College of Medical Genetics and Genomics (ACMG)cell line samples) as part of the AoURP validation study. Overall, we observed a 1.5% average fail rate for the AoURP validation samples at the Genome Centers and 1% at the CVLs.

2.2.2.8 Liftover

The AoU WGS will be aligned and reported for research use based on the GRCh38DH reference genome. However, the AoU clinical validation pipelines and results are reported in the GRCh37 reference format, since this reference is most trusted, validated, and used by the clinical genetics community. To convert variant VCF calls from the GRCh38 reference to the GRCh37 reference format, a liftover process will be performed as follows:

After the completion of whole genome sequencing, alignment, and variant calling, variant calls are converted from GRCh38 reference to GRCh37 reference. Variant calls that overlap sites where the reference alleles between GRCh38 and GRCh37 differ (see Table 22) are pre-processed to match the GRCh37 reference. Picard's LiftoverVcf 4.1.4.1—a bioinformatics software package—is then run on the GRCh38 VCFs to generate GRCh37 VCFs. The GRCh37 VCFs are then post processed to make them consistent with the GRCh37 reference genome.

Concordance analysis of SNPs and indels from 240 AoU harmonization samples was performed between lifted-over GRCh38-to-GRCh37 VCFs and the corresponding GRCh37 VCFs from DRAGEN analysis. Genotype concordance was 100% for all variants that passed DRAGENHardQUAL metric and had coverage ≥20x.

Annotations from the GRCh38 VCF and the lifted-over GRCh37 VCF were compared for 34 pathogenic and likely pathogenic variants. All annotations matched 100%. All pathogenic and likely pathogenic variants from the AoU Coriell validation study were used for this analysis.

Analysis was performed to compare the sequences for both references in the reportable range. Four sites were identified with mismatched reference alleles that need pre-processing before running Picard's LiftoverVcf to ensure 100% concordance. See **Table 21**.

Table 21. Mismatched bases after liftover between reference builds.

	GRCh37				GRCh38			
Gene	**Chr**	**Start**	**End**	**Sequence**	**Chr**	**Start**	**End**	**Sequence**
APOB	chr2	21,235,474	21,235,475	T	chr2	21,012,602	21,012,603	C
DSP	chr6	7,563,982	7,563,983	T	chr6	7,563,749	7,563,750	G
FBN1	chr15	48,807,636	48,807,637	C	chr15	48,515,439	48,515,440	T
TNNI3	chr19	55,665,583	55,665,584	A	chr19	55,154,215	55,154,216	C

Conclusion: Clinical variant calling results were 100% concordant between those that are natively aligned to GRCh37 and those that are lifted over from GRCh38 to GRCh37.

2.2.2.9 Report comprehension testing

The AoURP will return to participants results of incidental findings in the AoUHDR reportable region and AoUPGx alleles in the form of a research report. Five mock research reports are included as appendices: HDR Positive Report (*BRCA1, MSH2,* and *LDLR*) (Appendix 2, 3, and 4), HDR Uninformative Report (Appendix 1), and PGx Report (also referred to as "Medicine and Your DNA Report") (Appendix 5). The genes for the mock positive reports were chosen because they are each associated with one of the Tier 1 genomic applications (hereditary breast and ovarian cancer, Lynch syndrome, and Familial hypercholesterolemia, respectively) as defined by the Centers for Disease Control and Prevention's (CDC's) Office of Public Health Genomics and are expected to be some of the most common results in the population.

To evaluate participant comprehension of the AoURP reports, a mixed methods research approach was utilized to assess the content validity of survey items and participant understanding of report-specific concepts through a computer-administered survey. Detailed research methodology and findings from the quantitative arm of the study can be found in Appendix 8.

Briefly, participants were recruited through a user experience platform, and study objectives and procedures shared with participants through a virtual study information session. All participants who attended the information session were invited to take the survey. Survey response rates were

31.5% for the Positive HDR Report (e.g., *BRCA1*), 23.8% for the Uninformative HDR Report, and 31.0% for the Medicine and Your DNA Report.

Participants were 45 years of age or older (n=401, 48.9%), female (n=526, 64.1%), non-white (n=366, n=44.6%), Latino/a (n=125, 15.2%), had an associate's degree or less education (n=417, 50.9%), and earned $74,999 annually or less (n=498, 60.7%). Participant comprehension rates for the Positive HDR Report (n=347) were 96.9% (96.7% genetic knowledge, 97.5% self-efficacy concepts), 96.6% for the Uninformative HDR Report (n=287; 94.6% genetic knowledge, 98.6% self-efficacy concepts), and 98.1% for the Medicine and Your DNA Report (n=205; 97.6% genetic knowledge, 98.4% self-efficacy concepts).

Conclusion: Participants were able to understand the AoURP Positive HDR, Uninformative HDR, and the Medicine and Your DNA Reports.

Overall Conclusion from Non-Clinical Laboratory Prior Investigations: The device is highly accurate, highly precise, equivalent across GCs, and meets the Sponsors' needs for the genes and alleles of interest in the sample matrix proposed.

2.2.2.10 Future Validations

It is important to note that there are additional disorders/phenotypes that have been reported in association with the genes analyzed -- with a wide range of supporting evidence. The decision to focus on the phenotypes listed is based on the recommendations from the 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. The American College of Genetics and Genomics (ACMG) chose to restrict their list to gene-disease pairs in which the disorders are considered actionable and the genes represent major contributors to disease. Where there are multiple names for a syndrome, we chose to include the most referenced name for the syndrome (e.g. Long QT syndrome versus Romano-Ward syndrome). In addition to ACMG's guidelines, we also defer to ClinGen's evaluation for strength of evidence to ensure sufficient validity of these associations. If we become aware of additional gene-disease associations that are valid as determined by ClinGen and actionable as determined by ACMG, we will consider adding those and will adhere to our change management guidelines for informing the FDA of such a change.

Several additional gRoR modules/capabilities have been proposed for the future. These include but are not limited to:

- Adding copy number calling for specific alleles (e.g., CYP2D6 for PGx).
- Adding DNA derived from saliva as a specimen type.
- Adding additional genes to the AoUHDR region or additional alleles to the AoUPGx list.
- Adding polygenic risk score reports.

In each case we propose to follow the change control procedures outlined in this application and consult with the FDA via a pre-IDE process where appropriate to agree upon a validation strategy tailored to the specific new feature. Based on the initial IDE experience, we believe this will largely take the form of identifying appropriate clinical specimens, supplementing with cell lines or contrived samples where appropriate, and generating data to establish analytical validity

(accuracy, reproducibility, etc.) in addition to a risk analysis. Relevant updates to informed consent, reporting materials, etc., will also be made on a case-by-case basis in consultation with the FDA.

3. Investigational Plan

3.1 Purpose

Intended Use Statement

The *All of Us* WGS assay is a next generation sequencing based in vitro device for the detection of SNVs and indels in 66 genes using DNA isolated from fresh whole blood. The device is intended to provide AoURP participants with genetic information that could be relevant to their health, specifically HDRR and PGx variant status.

3.2 Protocol

The gRoR protocol is located as file 002_AoURP_Protocol on the USB drive containing this IDE.

3.3 Risk Analysis

A significant risk in vitro diagnostic (IVD) device is generally one of substantial importance in diagnosing, curing, mitigating, or treating disease, or otherwise preventing impairment of human health and presents a potential for serious risk to the health, safety, or welfare of a subject or otherwise presents a potential for serious risk to health, safety, or welfare of a subject (21 CFR 812.3(m)). For IVDs, FDA guidance defines "potential for serious risk" in relation to the nature of the harm that may result to the subject. Misdiagnosis and/or error in treatment caused by inaccurate test results would be considered a significant risk if the potential harm to the subject could be life-threatening or could result in permanent impairment of a body function or permanent damage to the body structure. In this study, we are not diagnosing or providing treatment suggestions in the health-related results that will be returned. The results and associated educational materials explicitly state they are research results that would require follow-up clinical testing to confirm and consultation with a healthcare professional before taking any medical action(s).

As described in Section 2, the MedSeq Project (Vassy et al., 2014), a randomized clinical trial utilizing WGS to return health-related genetic findings to apparently healthy adult participants, demonstrated that participants receiving these results did not experience undue distress (Lee et al., 2015; Roberts, et al., 2018). Further, Vassy et al. (2017) reported that primary care physicians managed these findings without serious errors and that short-term downstream healthcare costs were not significantly greater for those receiving these results (Christensen et al., 2018; Christensen et al., 2018; Perkins et al., 2018; Zoltick et al., 2019). Given these findings, we do not anticipate any serious risks or adverse events to be associated with returning genetic research results. Furthermore, result reports are clearly labeled as research results that should not be used to alter medical care unless verified in a clinical lab.

Table 22 describes potential risks, along with applicable general controls that the AoURP has identified, as well as various risk mitigation strategies to provide reasonable assurance of the safety and effectiveness of the program.

Table 22. Potential risks and AoURP mitigation strategies

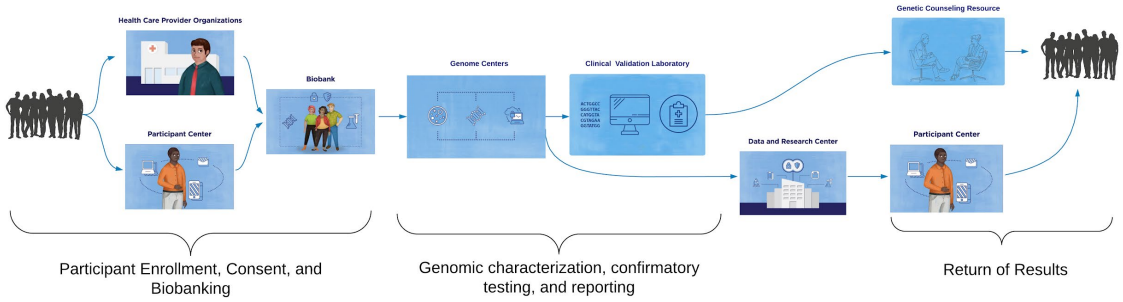
Risk	Mitigation Strategy
Standard risk of genetic testing - Incomplete or Incorrect Results due to 1. limitations of current sciences	• Limitations discussed in informed consent, educational materials, and reports (e.g., from the HDR report "Could my results change? Yes. *All of Us* could look at more genes or look again at these genes or DNA changes as science improves. Check your *All of Us* account to make sure this is the most up-to-date version of this report."). • Variant classifications will be updated over time; CVLs will re-issue AoUHDR reports whenever a reported variant changes classification and GCR will be notified to provide support as necessary (see study protocol section 6.1.4.1 *AoUHDR results - Updating results over time* for more details).
Standard risk of genetic testing - Incomplete or Incorrect Results due to 2. Malfunction of the device	• Analysis is performed in clinical labs with stringent regulatory and compliance programs, controlled under the total QMS based on CLSI Guidelines. • Analysis conducted using clinically validated LDTs. • All reports will be reviewed and signed out by a board-certified laboratory geneticist or molecular pathologist (consistent with practice of medicine). • All positive HDR results will be confirmed with an orthogonal and medically established method.
Standard risk of genetic testing - Incomplete or Incorrect Results due to 3. Human error (e.g., mislabeling, contamination)	• To the extent possible, processes are automated to minimize the opportunity for human error.
Standard risk for genetic testing – Misuse of information [e.g., participant changes medication or medical care without first consulting a healthcare provider]	• Reports and educational material clearly instruct how to use results (e.g., from HDR report "This report comes from a research program so it is a research result. Your doctor will need to confirm these results with a clinical genetics test before using them in your care. Do not change your medical care before this result is confirmed by your doctor.")

Risk	**Mitigation Strategy**
	• Comprehension testing for reports demonstrated high participant comprehension (>95%) • Reports include disclaimer "Results provided are from an investigational device. An 'investigational device' is a device that is the subject of a clinical study." • Positive results for P/LP variants are returned during a meeting with a licensed genetic counselor who can help to reinforce safety issues. • Genetic counselors available at no cost (regardless of type of result) through the Genetic Counseling Resource (GCR) to answer questions and to remind participants of the limitations of these results.
Low risk for study - Emotional/psychological distress from 1. Fear for health of self/loved ones	• Consent, educational materials, and reports undergo IRB review and approval to ensure they clearly convey and contextualize the intents, benefits, risks, and expected outcomes of this study. • Program offers additional educational materials and support from GCR to assist participants in these situations. • Genetic counselor will disclose any P/LP HDR result directly to the participant, during which they will monitor the participant for distress. In the unlikely event of significant anxiety or distress, the genetic counselor will refer the participant to a mental health professional and provide a "warm handoff," when possible.
Low risk for study - Emotional/psychological distress from 2. Results that make participant question group/family membership	• Topic is discussed and contextualized in the consent and educational materials. • GCR is on hand to assist with interpreting the meaning of results regarding health, relatedness, and identity, • Program is not returning familial results (i.e. no relatedness linkages or information) to minimize occurrence of this outcome.

Risk	**Mitigation Strategy**
Low risk for study - Emotional/psychological distress from 3. Navigating family discussions of results	• Language on the potential need for family discussion in both the consent, educational materials and reports. • GCR will facilitate conversations with participants and their family members at the request of the participant and only with the participant present in the conversation.
Standard risk of genetic testing - Privacy and Security Incidents 1. Breach of data	• Security controls derived from NIST Special Publication 800-53 (Security and Privacy Controls for Federal Information Systems and Organizations) to meet or exceed the Federal Information Security Management Act (FISMA) moderate baseline; controls are implemented in accordance with the Precision Medicine Initiative Data Security Policy Principles and Framework. • The program and all awardees adhere to the *HHS Policy and Plan for Preparing and Responding to a Breach of Personally Identifiable Information (PII).* • See Section 2.3.1.2 *Privacy and Security Incidents* of the study protocol for full details.
Standard risk of genetic testing - Privacy and Security Incidents 2. Disclosures and misuse	• The program has been issued Certificates of Confidentiality, which limit the allowable disclosures and uses of data generated under its auspices. • Results are not returned to a participant's care provider without: 1) requirement by state law; and 2) explicit consent/request of the participant.
Standard risk for genetic testing - Clinicians will be unprepared for medical management of participants with HDR results (positive or uninformative)	• Educational materials are available to participants and providers. • Reports prominently feature contact information for GCR with notice to clinicians of availability of genetic counselors to answer questions.
Standard risk of genetic testing -	• All HDR P/LP findings are based on ACMG professional guidelines.

Risk	Mitigation Strategy
Providing pleiotropic or uncertain HDR P/LP results	• All results are reviewed and approved board-certified laboratory geneticist or molecular pathologist. • No Variants of Unknown Significance (VUSs) will be reported. • Licensed genetic counselors from GCR will facilitate disclosure of all P/LP HDR results.

3.4 Description of Investigational Device



In this "device", constituting a small but important feature of the much more substantive AoURP, results from clinical interpretation of pre-defined HDR and PGx associated genes will be returned to participants who consent to receive results, with incorporation of all appropriate disclaimers on limitations and use of the results.

The sections immediately following describe the genome sequencing and analysis methods that form the foundation of the device and that occur prior to returning genomic results to a participant.

3.4.1 Summary of reagents, software and instruments (Table 23)

Table 23. Summary of reagents, software and instruments used in AoUWGS device.

	Baylor	Broad	UW
Sample Accessioning & QC			
Quantitation	Picogreen (Synergy) or DropQuant	Picogreen	Invitrogen Quant-it
Automation / Liquid Handler	Biomek FXp	Dynamic Devices Lynx	SPT LabTech Mosquito and Perkn Elmer Janus
Library Construction			
Library prep	PCR Free Kapa HyperPrep	PCR Free Kapa HyperPrep	PCR Free Kapa HyperPrep
Barcodes	96 8-bp Illumina TruSeq DNA UD Indices	8-bp unique dual indices (Roche)	576 10-bp unique dual indices
Automation / Liquid Handler	Biomek FXp	Agilent Bravos	Perkin Elmer Janus
Sonicator	Covaris E220	Covaris LE220-Plus	Covaris LE220
Library QC - Quantitation	AB QuantStudio 6 Flex	Viia7 qPCR machine	Biorad CFX384
Library QC - Size estimation	Agilent Bioanalyzer 2100 or Agilent Fragment Analyzer	Agilent Bioanalyzer 2100 or Agilent Fragment Analyzer	Agilent Fragment Analyzer
Whole Genome Sequencing (WGS)			
Sequencer	NovaSeq 6000	NovaSeq 6000	NovaSeq 6000
Multiplexing & Sequencing Strategy	2 pooling methods: - 24-plex on S4 flowcell - calibration pool with re-pool: 75-plex on 12 lanes of S4 flowcell	- 24-plex pool on 24 flowcell - adaptive pooling	- 192-plex pool for NovaSeq XP QC run - 26-plex pool on S4 flowcell

Software			
Bioinformatics	DRAGEN v3.4.12	DRAGEN v3.4.12	DRAGEN v3.4.12

3.4.2 Sample Receipt, Accession and QC

Upon receipt of DNA sample shipments from the Mayo Biobank, the GCs perform an inspection of the packaging and sample containers to ensure that sample integrity has not been compromised during transport and to verify that the sample containers correspond to the shipping manifest. QC of the submitted samples also includes DNA quantification, using routine procedures to confirm volume and concentration.

Any issues or discrepancies are recorded, and affected samples are put on hold until resolved. Samples that meet quality thresholds are accessioned in the Laboratory Information Management System (LIMS), and sample aliquots are prepared for library construction processing (normalized with respect to concentration and volume).

3.4.3 Library Construction

The DNA sample is first sheared using a Covaris sonicator and is then size-selected using AMPure XP beads to restrict the range of library insert sizes. Using the PCR Free Kapa HyperPrep library construction kit, enzymatic steps are completed to repair the jagged ends of DNA fragments, add proper A-base segments, and ligate indexed adapter barcode sequences onto samples. Excess adaptors are removed using AMPure XP beads for a final clean up. Libraries are quantified using qPCR with the Illumina Kapa DNA Quantification Kit and then normalized and pooled for sequencing.

3.4.4 Sequencing and Data QC

Pooled libraries are loaded on the Illumina NovaSeq instrument (see Section 2.2.2.2 Part 7 for detail on molecular barcoding and multiplexing). The data from the initial sequencing run are used to QC individual libraries and to remove non-conforming samples from the pipeline. The data are also used to calibrate the pooling volume of each individual library and re-pool the libraries for additional NovaSeq sequencing. WGS uses Illumina reagents and follows the manufacturer's best practices.

3.4.5 Bioinformatic Analysis

After demultiplexing (see Section 2.2.2.2 Part 7 for detail on barcoding and multiplexing), WGS analysis occurs on the DRAGEN platform. The DRAGEN pipeline consists of highly optimized algorithms for mapping, aligning, sorting, duplicate marking, and haplotype variant calling and makes use of platform features such as compression and BCL conversion. In the 2017 Precision FDA Hidden Treasures – Warm Up Challenge, the DRAGEN Platform received the highest score in five out of six accuracy measures for whole-genome variant calling among platforms that recognized all 50 variants (*https://precision.fda.gov/challenges/1/view/results*). Additionally, DRAGEN is able to perform both intra- and inter-flow cell event merging of fastq data files, providing the required flexibility for optimal NovaSeq multiplexing. Alignment uses the GRCh38dh reference genome. QC data are collected at every stage of the analysis protocol, providing high-resolution metrics required to ensure data consistency for large-scale

multiplexing. The DRAGEN pipeline produces a large number of metrics that cover lane, library, flow cell, barcode, and sample-level metrics for all runs as well as assessing contamination and mapping quality. Bioinformatic tools are summarized in **Table 24**.

Table 24. Bioinformatics Tools and Versions

Site	Sequencer Real Time Analysis (RTA)	Demultiplex	Map & Align	Variant Call
Broad Institute	v3.4.4	DRAGEN v3.4.12	DRAGEN v3.4.12	DRAGEN v3.4.12
BCM HGSC-CL	v3.4.4	Picard (2.6.0)	DRAGEN v3.4.12	DRAGEN v3.4.12
UW - NWGC	v3.4.4	Picard (2.20.5)	DRAGEN v3.4.12	DRAGEN v3.4.12

3.4.6 Facilities and CVLs

The Program uses the term CVL to encompass all clinical interpretation activities. More specifically, CVL may refer to: 1. variant interpretation activities at each Genome Center, or 2. the specific function of validating an HDR+ result. In brief, the AoURP Genome Centers are each responsible for the genome interpretation of participant samples received at that Center. The interpretation results from each Center are processed through the Reporting Harmonization Platform where, as the name indicates, harmonization of results is checked. Although we refer to the interpretation process as occurring in CVLs, this is simply an extension of activities at each Center. The three GCs, as well as the CVLs hold the relevant clinical lab licenses to perform molecular testing on clinical samples. Confirmatory tests at the three sites are used routinely to perform clinical testing for patients. Depending on the license, these facilities are inspected on a regular basis by their respective states (CLIA), the CAP and/or the State of New York. Each of the centers has designed an internal process for quality management and assay validation that complies with these agencies.

The UW CVL is deploying a validated, multi-gene NGS panel test as described previously (Pritchard CC, et al. 2014; Pritchard CC, et al. 2012). Laboratory procedures, bioinformatic analyses, and variant interpretation are performed at the U W (Seattle, WA) under CLIA license #50D250662.

The Color Genomics CVL deploys a multi-gene NGS panel Color test as described (Neben, et al. 2019; Berger, et al. 2020). Laboratory procedures, bioinformatics analysis, and variant interpretation are performed at Color (Burlingame, CA) under CLIA (#05D2081492) and CAP (#8975161) compliance. Analysis, variant calling, and reporting focus on the complete coding sequence and adjacent intronic sequence of the primary transcript(s), unless otherwise indicated.

The Human Genome Sequencing Center Clinical Laboratory (HGSC-CL) CVL assay utilizes standard Sanger sequencing methods to confirm sequencing variants that have been detected in the WGS assay. Modified dye-terminator protocols enable sequencing through regions containing GC-rich, di- and tri-nucleic repeats, inverted and Alu repeats, or long homopolymer stretches. In some cases, long-range primer design may be employed due to segmental duplications, pseudogenes and other high-similarity regions. Samples are loaded and sequenced on the Applied Biosystems 3500 Genetic Analyzer and analyzed utilizing Mutation Surveyor (MS, SoftGenetics) that assigns base calls and quality values. All Sanger laboratory procedures,

analysis, and variant interpretation are performed at the HGSC-CL CVL under CLIA (# 45D2027450) and CAP (# 8004250) compliance.

3.4.7 Control Information

The AoURP WGS assay incorporates both programmatic controls and process controls. Programmatic controls consist of study duplicates, while process controls are known control samples consisting of the seven NIST-GIAB samples. These controls are provided by the Mayo Biobank at 2- month intervals. Each GC uses NA12878-NIST as an internal library construction control and as part of their WGS CLIA validation. One control is added to each plate of 94 samples in the same plate position for every production run. The NA12878-NIST control serves as an internal QC at several steps throughout the process. At 2- month intervals, the process control is replaced with one of the seven GIAB samples and fully sequenced. If control samples fail at any point during the WGS assay, a deviation is reported and an investigation is initiated. To monitor WGS data equivalency, programmatic and processes control sequencing data files are compared between centers.

In addition to the positive controls one well on each plate is left empty and serves as a negative control. WGS samples are pooled together for sequencing on the Illumina NovaSeq and 1% PhiX is added to each library pool and serves as a process control for cluster generation and alignment.

3.4.8 Comprehensive Procedural Quality Metrics

Every step through the WGS procedure is rigorously controlled by pre-defined quality control measures. Various control mechanisms and acceptance criteria are established during assay validation. Outliers to the established acceptance criteria must be investigated and resolved prior to result reporting. Specific metrics for reviewing and releasing genome data are listed here:

- **Mean coverage (threshold ≥30X)** - The total FPs and FNs show a gradual increase as mean coverage decreases, with a rapid increase below 20x coverage, supporting a stringent threshold selection of a minimum of 30x.
- **Genome coverage (threshold ≥90% at 20X)** - The total FPs and FNs steadily increase as the percent of bases with at least 20x coverage drops. Drop-off of performance is initially gradual, supporting a threshold of 90%.
- **AoUHDR coverage (threshold ≥95% at 20X)** - The total FPs and FNs increase gradually as the average coverage in AoUHDR regions decreases. The reduction in performance is slow initially, and then increases rapidly below 40%, showing that the genome center threshold of 95% is conservative.
- **Aligned Q30 bases (threshold ≥8e10)** - FPs and FNs increase with lower base quality counts, with inflection points starting around 6e10 for both.
- **Contamination (threshold ≤1%) -** Variant calling performance as measured by both the number of FPs and the number of FNs decreases with increasing contamination, demonstrating that the 1% threshold is appropriate.

For variant calling, all GCs have harmonized on the following set of Dragen parameters, which were locked prior to the IDE validation studies.

DRAGEN Parameter	Parameter Value	Description
-f	n/a	Overwrite if output exists
-r <hg38-ref-dir>	<hg38-ref-dir>	The reference to use
--fastq-list	<path-to>/fastq_list.csv	A list of fastq files to use as input for this sample
--fastq-list-sample-id	<sampleID>	The sample ID to use for naming this sample
--output-directory	<output-dir>	The location of the final output files
--intermediate-results-dir	<int-results-dir>	The location to write intermediate outputs
--output-file-prefix	[CenterID]_[Biobankid_Sampleid]_[LocalID:optional]_[Rev#]	Standardized naming prefix for each output file
--enable-variant-caller	true	Turn on variant call outputs
--enable-duplicate-marking	true	Mark duplicate reads during alignment
--enable-map-align	true	Produce an alignment from unaligned read input
--enable-map-align-output	true	Store the output of the alignment
--output-format	CRAM	Store the alignment as a CRAM file
--vc-hard-filter	DRAGENHardQUAL:all:QUAL<5.0;LowDepth:all:DP<=1'	This parameter setting changes the threshold on the quality to 5.
--vc-frd-max-effective-depth	40	Setting this parameter puts a limit on the penalty value that is applied for variant calls that deviate from the expected 50% allele fraction for heterozygous variants.
--qc-cross-cont-vcf	<path-to/SNP_NCBI_GRCh38.vcf>	Marker sites to use for contamination estimation
--qc-coverage-region-1	<path-to/wgs_coverage_regions.bed>	Regions to use for coverage analysis (whole genome)
--qc-coverage-reports-1	cov_report	The type of reports requested for qc- coverage-region-1
--qc-coverage-region-2	<path-to/HDRR_regions.bed>	Regions to use for coverage analysis (HDRR reportable regions)
--qc-coverage-reports-2	cov_report	The type of reports requested for qc- coverage-region-2
--qc-coverage-region-3	<path-to/PGx_regions.bed>	Regions to use for coverage analysis (PGx reportable regions)

--qc-coverage-reports-3	cov_report	The type of reports requested for qc- coverage-region-3

Notes regarding above table: The DP is the filtered depth of coverage and QUAL is the phred-scaled probability that the described variant exists at this site given the sequencing data. DRAGEN has simplified hard filtering rules compared to other conventional callers, and the QUAL score alone gathers most of the evidence of the variant being present or not. This is because rather than relying on hard filtering rules downstream of the variant caller, DRAGEN characterizes systematic artifacts and correlated pileup errors with mathematical models inside the genotyper. These models were developed to characterize errors, and help distinguish true variants from noise. The genotyping algorithm exploits certain properties of these artifacts (such as low MAPQ, skewed AF, strand bias, mean base quality at a site, position of the variant in the read) and incorporates this evidence into the probability calculation in a mathematically rigorous manner. The QUAL score incorporates all these effects into a single score at the output of the variant caller. These models have been extensively benchmarked and evaluated on a wide range of control samples, which mitigates risk. In particular, we note high accuracy on control samples in section 2.2.2 Test Performance Characteristics of the IDE.

The vc-hard-filter and the vc-frd-max-effective-depth parameters were changed from the defaults to improve the limit of detection and extensive benchmarking was conducted to find parameters that minimize the overall impact of variant calling performance. Variants with a skewed allele fraction tend to have a lower confidence than variants with an allele fraction closer to 50%. To improve performance on a limit-of-detection control we lowered the QUAL confidence threshold from 10 to 5.

3.5 Monitoring Procedures

3.5.1. On-going Monitoring

The study uses a risk-based monitoring procedure to ensure all pre-analytical, analytical and post-analytical testing procedures are conducted fully in compliance to CAP, CLIA, New York State Clinical Laboratory regulations, Health Insurance Portability and Accountability Act (HIPAA), Occupational Safety and Health Administration (OSHA) and all applicable laws and regulations. Each site will comply with its established quality management and compliance management programs.

3.5.2 Change Control Procedure

The AoURP established a program-wide change control procedure to manage changes during the course of the research program. Proposed changes will be reviewed internally and then communicated to the FDA appropriately, for FDA information and/or approval. This process is based upon the Final FDA Guidance entitled “Changes or Modifications During the Conduct of a Clinical Investigation” issued on May 29, 2001. All changes will be managed as follows:

Step 1. Initiation of A Change Request Within the AoURP

When a potential change is identified during the conduct of the study, the requestor will submit a change request to the AoURP Regulatory Compliance Office.

Step 2. Review of Change Request by the AoURP

The AoURP Change Control Committee comprised of NIH staff will conduct a risk analysis to identify potential risks that might be introduced by the change. Supporting information, such as preliminary testing results, will also be reviewed to determine the risk level. The Committee will review the evidence and recommendations, then make a final determination of the risk category. Once the risk analysis is completed, the description of the change, its risk analysis and recommendation will be shared with the AoURP Regulatory Compliance Office.

The AoURP Regulatory Compliance Office will determine if an IRB approval should be obtained for the change. If an IRB approval is necessary, the Sponsor will initiate an amendment (or request an exemption) to the IRB.

Potential changes will be classified into one of three possible categories based on the anticipated risk:

I) Major Change

Major changes are significant (i.e., high risk) changes made to the design or principles of operation of the device or manufacturing process. Major changes may include modification to the study protocol or investigational plan which affect the validity of data, participant risk, scientific soundness or rights, safety, and welfare of the participants.

II) Moderate Change

Moderate changes include non-significant (i.e., low risk) changes to the device or manufacturing process. Moderate changes may include study protocol or investigational plan changes which do not affect the validity of data, participant risk, scientific soundness, or rights, safety, welfare of participants.

III) Minor Change

Minor changes are non-significant changes (i.e., no known risk) to the device/manufacturing process or for which sufficient mitigation is in place to eliminate the risk. These changes also include minor investigational plan changes which do not affect the validity of data, participant risk, scientific soundness or rights, safety, welfare of participants.

Step 3. Submission to the FDA (Prior to Implementation; Only Applicable to Major Changes)

For major changes, the AoURP will work with the Responsible Investigators to prepare an IDE supplement. The supplement will include the IRB approval letter, if applicable, and results from any initial feasibility studies.

The Sponsor will review the package and submit the IDE supplement to the FDA.

Only after an IDE supplement is approved by the FDA, will the AoURP gRoR study proceed with the implementation of the respective change(s).

Step 4. Implementation of A Change Within the AoURP gRoR study

All proposed changes (major, moderate and minor) will require validation to comply with FDA, CAP, CLIA and all applicable laws and regulations. Depending on the type of change(s), the Responsible Investigators will prepare a project plan to include a validation strategy/plan, a clinical validation study and roll-out plan. Validation results and findings will be reviewed by the Sponsor. When the validation results meet the acceptance criteria as defined by the plan, the validation will be accepted and reported to the Sponsor for approval. Once approved, the requested change will be implemented accordingly. Validation results will also be reviewed and approved by the CLIA director or designee, at each site prior to implementation.

Major changes will be implemented after FDA concurs with the information provided in the IDE supplement.

Moderate changes will be synchronized with the Responsible Investigators to ensure sufficient time is allotted to prepare the 5-day notice submission to the FDA. Within 5 days after the implementation, FDA will be notified as described in Step 5.

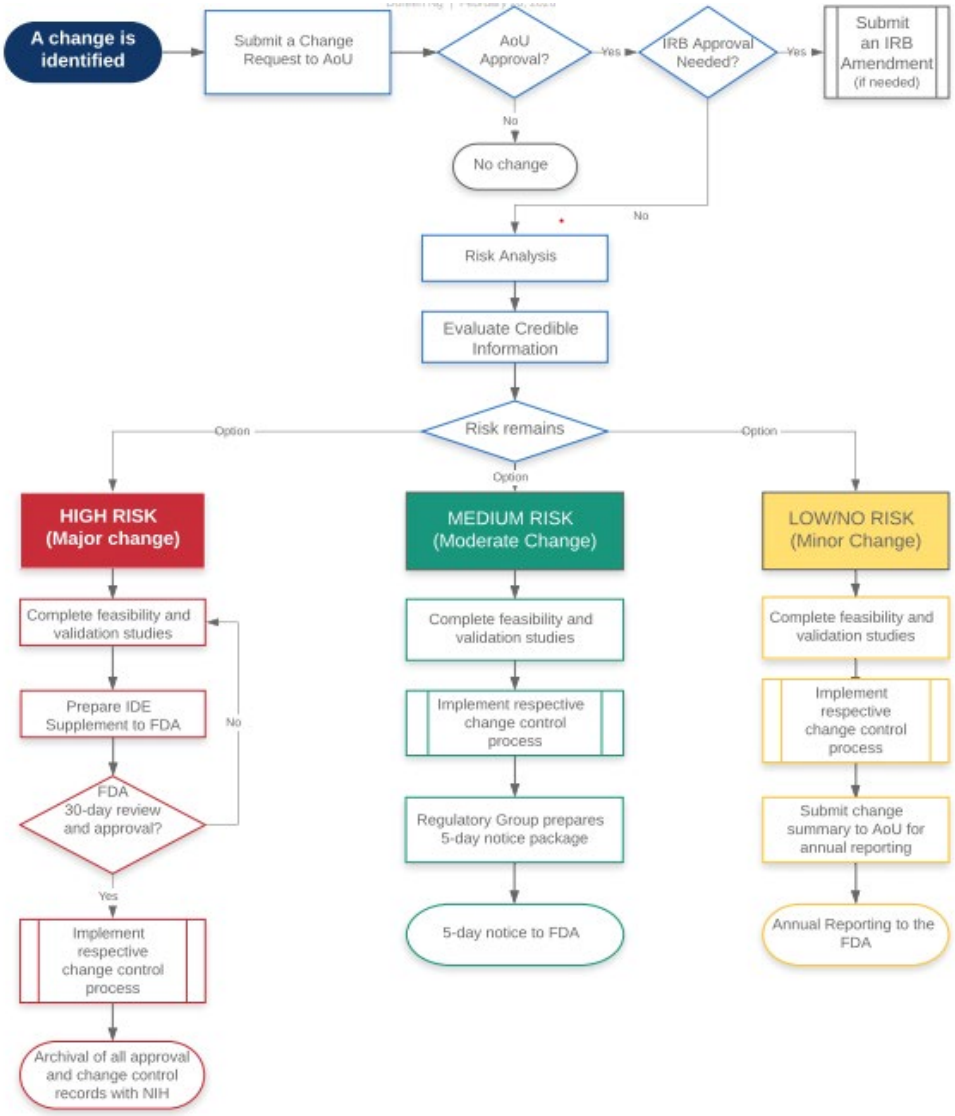
Minor changes will be implemented according to the change control procedure at the respective institution and after approval by the Sponsor.

Step 5. Subsequent Change Notification to the FDA

Notification for major changes will be provided to the FDA through an IDE supplement.

All moderate changes will be reported to the FDA within 5 days using the 5-day notice process.

All minor changes are reported by the respective centers in the annual progress report to the Sponsor who will include the information in the IDE annual report.



3.6 Name and address of the individual(s) who will monitor the study

1. Stephanie Devaney, PhD, AoURP Chief Operating Officer
2. John Horigan, AoURP Director, IRB and Protocol Compliance
3. Bradley Ozenberger, PhD, AoURP Genomics Team Lead
4. Justin Hentges, AoURP Budget and Awards

All individuals are located at the National Institutes of Health, AoURP, 6001 Executive Boulevard, Suite 8112, Rockville, MD 20852. The phone number is (301) 451-5996

4. Methods, Facilities and Control Information

The “device” will not be manufactured or marketed. Methods and facilities, including standard equipment, for operation of the “device” are described in Section 3.4.

5. Example and Certifications of Investigator Agreements and Financial Interest

The AoURP gRoR IDE study Investigator Agreement template is shown below and is provided as Appendix 9.

All of Us Research Program: Investigator Agreement

FDA Investigational Device Exemption Study – Return of Genetic Results in the *All of Us* Research Program

Sponsor: National Institutes of Health *All of Us* Research Program

This Investigator Agreement provides acknowledgement of the signatory of his/her responsibilities as a co-Responsible Investigator in the referenced study, per requirements specified by 21 CFR 812.43.

Instructions in italics.

Name / Title / Institution

a. Curriculum Vitae

b. Relevant experience

c. Information on any terminated studies you were involved in, including an explanation of the circumstances that led to termination

d. Investigator statement: I certify that I will conduct the study in accordance with this agreement, the Investigational Plan (IDE study description), the IDE and other applicable FDA regulations, and conditions of approval imposed by the reviewing IRB or FDA; supervise all testing involving human subjects; and ensure requirements for informed consent are met, as applicable.

e. Financial disclosure: I certify that based on information obtained from the sponsor or from other participating investigators, the listed clinical investigators (list of names contained in email request) did not participate in any financial arrangement with the sponsor of a covered study whereby the value of compensation to the investigator for conducting the study could be affected by the outcome of the study (as defined in 21 CFR 54.2(a)); had no proprietary interest in this product or significant equity interest in the sponsor of the covered study (as defined in 21 CFR 54.2(b)); and was not the recipient of significant payments of other sorts (as defined in 21 CFR 54.2(f)). I affirm that I will notify the Sponsor if new financial conflict should become relevant, for a time period to extend to one year after the conclusion of the study.

Note that the information provided will not be submitted in the IDE application to the FDA. Its collection by the Sponsor is required for submission in any marketing application involving the device. AoURP has no intention to market or license the device described in the study.

Signature _____

Printed Name

Date _____

End Investigator Agreement

This study includes multiple co-Responsible Investigators (co-Is) to ensure safety and compliance at all stages of the protocol. These co-Is direct respective federal awards for the NIH AoURP that comprise the Biobank, DRC, GCR, Genome Centers, Regional Medical Centers, Federally-qualified Health Centers, Veteran Affairs Medical Centers, and The Participant Center. Those responsible investigators are listed below. The Sponsor attests that each co-I has affirmed agreement to:

· protecting the rights, safety, and welfare of participants

· obtaining informed consent/ensuring informed consent has been appropriately obtained

· permitting the use of the investigational device only with participants under his/her/their supervision, not to any person not authorized under the IDE to receive it

·conducting the investigation in accordance with the signed agreement with the Sponsor, the investigational plan, the IDE regulations and other applicable FDA regulations, and any conditions of approval imposed by an IRB and FDA

· providing sufficient accurate financial information to allow the Sponsor to submit certification or disclosure of financial interests, as required in the IDE application

·maintain accurate and complete records relating to the investigation including: all correspondence including required reports, the protocol and documentation (date and reason) for each deviation from the protocol, any other records that FDA requires to be maintained by regulation or by specific requirement for a category of investigation or a particular investigation.

The study involves disparate awardee/institutions with discreet responsibilities for the execution of the study. All awardees are monitored for safety and compliance by the Sponsor (NIH). In particular, the term within the investigator agreement stating the investigator affirms *agreement to obtaining informed consent/ensuring informed consent has been appropriately obtained* does not pertain to some operational components. Healthcare Provider Organizations (HPOs), (Regional Medical Centers [RMCs], Federally Qualified Health Centers [FQHCs], and the Veterans Administration [VA]) conduct in-person sessions with participants and conduct identification (ID) verification and consent checks. The Participant Center (TPC) also oversees partners that may have in-person contact with participants. However, as described in the Protocol, the status of informed consent is recorded electronically in the Data and Research Center (DRC) and status is conveyed to trained staff at enrollment sites using a web-based tool. Similarly, the Biobank and the GCs are dependent on informed consent verification status as provided by the DRC. The DRC maintains records of informed consent (including withdrawal of consent) that are updated daily.

Responsible Investigator Name, Institution, and Address

Biobank

Stephen N. Thibodeau, PhD
Professor of Laboratory Medicine and Pathology
Mayo Clinic Rochester
Rochester, MN

DRC—Data and Research Center

Paul A. Harris, PhD
Professor, Dept of Biomedical Informatics
Vanderbilt University Medical Center
Nashville, TN

GCR - Genetic Counseling Resource

Alicia Y. Zhou, PhD
Chief Science Officer
Color Genomics, Inc.
Burlingame, CA

Genome Centers

Stacey B. Gabriel, PhD
Senior Director, Genomics Platform
Broad Institute of MIT and Harvard
Cambridge, MA

Richard A. Gibbs, PhD
Director of the Human Genome Sequencing Center
Baylor College of Medicine
Houston, TX

Deborah A. Nickerson, PhD
Professor of Genome Sciences
University of Washington
Seattle, WA

Healthcare Provider Organizations

RMCs - Regional Medical Centers

David B. Goldstein, PhD
Director, Institute for Genomic Medicine
Columbia University Irving Medical Center
New York, NY

Lucila Ohno-Machado, MD, PhD
Professor of Medicine
University of California-San Diego School of Medicine
La Jolla, CA

Christine D. Cole Johnson, PhD
Chair, Department of Public Health Sciences
Henry Ford Health System
Detroit, MI

Jordan W. Smoller, MD, ScD
Professor of Psychiatry
Massachusetts General Hospital
Boston, MA

Philip Greenland, MD
Professor of Preventative Medicine
Northwestern University at Chicago
Chicago, IL

Irving L. Kron, MD
Senior Associate Vice President
University of Arizona
Tucson, AZ

Steven E. Reis, MD
Professor of Medicine
University of Pittsburgh

Pittsburgh, PA

Bruce R. Korf, MD, PhD
Medical Director
University of Alabama at Birmingham
Birmingham, AL

Narayana S. Murali, MD
Executive Director of Marshfield Clinic
Marshfield Clinic Research Institute
Marshfield, WI

Stephan L. Zuchner, MD, PhD
Professor and Chair
University of Miami Health System
Coral Gables, FL

FQHCs - Federally Qualified Health Centers

Parinda Khatri, PhD
Clinical Psychologist
Cherokee Health Systems
Knoxville, TN

Yashoda Sharma, PhD
Senior Research Scientist
Community Health Center, Inc.
Middletown, CT

Eric M. Schlueter, MD
Family Medicine Specialist
Eau Claire Cooperative Health Center
Columbia, SC

Carmen Chinea, MD, MPH
Chief of Clinical Strategy and Research
Hudson River Health Care Community Health
Peekskill, NY

Donna Antoine-LaVigne, MSEd MPH, PhD
Principal Investigator
Jackson Heart Study Community Outreach Center, Jackson-Hinds Comprehensive Health Center
Jackson, MS

Fatima A. Muñoz
Director in the Research and Health Promotion Dept.
San Ysidro Health Center
San Ysidro, CA

VAMCs - Veterans Affairs Medical Centers

Christopher J. O'Donnell, MD, MPH
Chief of Cardiology
United States Department of Veterans Affairs Boston Healthcare System
Boston, MA

Phillip S. Tsao, PhD
Associate Chief of Research & Development
United States Department of Veterans Affairs Palo Alto Health Care System
Livermore, CA

TPC - The Participant Center

Eric Topol, MD
Executive Vice President
Scripps Translational Science Institute
La Jolla, CA

6. Sponsor's Certification

All Responsible Investigators participating in this investigation have signed the Investigator Agreement in accordance with 21 CFR 812.20(b)(5). Provided in Section 5 is the current list of Responsible Investigators. The federal awards to these investigators and their institutions confer responsibilities to all other affiliated investigators as described further in Section 8. No additional Responsible Investigator will be permitted to participate in the study without signing the Investigator Agreement.

7. Reviewing Institutional Review Boards

Name of Reviewing IRB: All of Us (AoU) IRB

Address: c/o Emmes Corporation
401 N. Washington St., Suite 700
Rockville, MD 20850

Chairperson: Nancy E. Kass, Sc.D.
Johns Hopkins University
Baltimore, MD

8. Name and Address of Other Investigational Institutions

It is imperative, even in the complex organization of program partners involved in this program, that all enrollment and specimen collection personnel are adequately trained in the protocol and procedures to ensure all human subject protections are enforced and to ensure data quality and integrity are established throughout. Furthermore, the responsibilities of supervisory investigators must be articulated and suitably enforced. This establishment of responsibilities and the mechanisms of enforcement are provided via the terms and conditions specified by the federal awards to the partners listed in Section 5. The terms and conditions as specified in the primary awards are required to flow down to any and all sub-contractors employed by the primary awardee institution. Illustrative terms and conditions are provided below to demonstrate the contractual obligations of all institutions participating in the program as enrollment partners and their obligations to oversee sub-contractors. The Sponsor affirms that these terms and conditions ensure the compliance of all investigators involved in the program with the policies and procedures established by the AoURP.

Protection of Human Subjects – Conditions of Federal Award	
From Notice of NIH Award	**HUMAN SUBJECTS** - No funds may be drawn down from the payment system and no obligations may be made against Federal funds for research involving human subjects by any site engaged in such research for any period not covered by both an OHRP-approved Assurance and approval from the AoURP IRB, as required, consistent with 45 CFR Part 46. This award requires the institution to ensure that all key personnel and partners who engage in human subjects research have completed education on the protection of human subjects, in accordance with NIH policy, "Required Education in the Protection of Human Research Participants," found here: (*http://grants.nih.gov/grants/guide/notice-files/NOT-OD-00-039.html*). Any individual involved in the design and conduct of the study must satisfy this requirement prior to participating in the project. Failure to comply can result in the suspension and/or termination of this award, withholding of support of the continuation award, audit disallowances, and/or other appropriate action.
IRB Reliance	**SPECIAL AWARD CONDITION: AOURP CENTRAL Institutional Review Board (IRB)** Human subject research conducted under this award will be under the authority of the AOURP Central IRB. No other IRB will have authority and Awardees use of another IRB will be considered non-compliant and Enforcement shall be invoked in accordance with the NIH Grants Policy Statement.

Sub-contractor Term	
Enrollment Subcontractor link to Primary Award	This award is subject to the conditions set forth in OT-PM18-001, "Limited Competition: All of Us Research Program Regional Medical Center Healthcare Provider Organizations (OT2)" which are hereby incorporated by reference as special terms and conditions of this award. Copies of this funding opportunity can be found at the following link: *https://allofus.nih.gov/sites/default/files/hpo_2018_ot_final.pdf*.
Term in DV Partner (Scripps) Award Binding Subcontractors	**CONSORTIUM/CONTRACTUAL COSTS** This award includes funds awarded for consortium activity with Sage Bionetworks, PatientsLikeMe, Leidos Innovations Corp, Computer Science Corp, EMSI, WebMD, BCBS Association, San Diego Blood Bank, Montage, UC San Diego, Univ of Southern CA, Walgreen Company, and CareEvolution. The recipient, as the direct and primary recipient of NIH grant funds, is accountable to NIH for the performance project, the appropriate expenditures of grant funds by all parties, and all other obligations of the recipient, as specified in the NIH Grants Policy Statement. In general, the requirements that apply to the recipient, including the intellectual property requirements also apply to consortium participant (s).

The following comprises the name and address of current primary and subordinate sites involved in investigations for the AoURP device protocol, including all principal participant enrollment partners (marked with a *).

Biobank

- Mayo Clinic
  Rochester, MN 55905
    - Mayo Clinic
      Jacksonville, FL 32224

Data and Research Center (DRC)

- Vanderbilt University Medical Center
  Nashville, TN 37232
    - The Broad Institute
      Cambridge, MA 02142
    - Verily Life Sciences
      South San Francisco, CA 94080

Genetic Counseling Resource (GCR)

- Color Genomics, Inc.

  Burlingame, CA 94010

Genome Centers

- Broad Institute, Inc.

  Cambridge, MA 02142

  - Massachusetts General Hospital

    Boston, MA 02114

    (LMM is part of MGH which is a subcontractor of the Broad). LMM's involvement is primarily to contribute work in variant classification and harmonization with some advisory engagement based on experience in managing interpretation of results for a biobank as well as variant harmonization experience for large consortia such as eMERGE. Matt Lebo, the current laboratory director of the LMM, co-chairs, along with Steven Harrison, the Variant Harmonization Subcommittee of the Clinical Interpretation and Reporting Work Group for the AoURP. In regard to investigator responsibility structure, LMM is contractually bound to AoURP policies and principles through the federal award to the Broad Institute.)

  - Color Genomics, Inc.

    831 Mitten Road #100
    Burlingame, CA 94010

- Baylor College of Medicine

  One Baylor Plaza
  Houston, TX 77030

- University of Washington
  Box 355065
  Seattle, WA 98195

Health Care Provider Organizations

Regional Medical Centers (RMCs)

- Columbia University*
  New York, NY 10032
  - Weill Cornell Medicine*

New York, NY 10065
- NYC Health and Hospitals/Harlem Hospital Center*
  New York, NY 10037
- Hunter College of CUNY*
  New York, NY 10065

- University of California, San Diego*
  Department of Biomedical Informatics

  La Jolla, CA 92093

  - University of California, Irvine*
    Department of Medicine, School of Medicine
    Irvine, CA 92697
  - University of Southern California*

    Health Sciences Campus
    Los Angeles, CA 90033
  - Cedars-Sinai*

    West Hollywood, CA 90069
  - University of California, San Francisco*

    Department of Epidemiology and Biostatistics
    San Francisco, CA 94143
  - University of California, Davis*

    Department of Medical Pathology and Laboratory Medicine
    PATH Building
    Sacramento, CA 95817
  - San Diego Blood Bank *

    San Diego, CA 92102
  - The Henne Group

    San Francisco, CA 94107

- Henry Ford Health System *
  Detroit, MI 48202
  - Spectrum Health*
    Grand Rapids, MI 49503
  - Baylor Scott and White Research Institute*
    Dallas, TX 75201
  - Essentia Institute of Rural Health *
    Duluth, MN 55805
  - University of Massachusetts Medical School *

North Worcester, MA 01655
- HealthPartners Institute*
  Bloomington, MN 55425
- Reliant Medical Group*
  Worcester, MA 01605

- Massachusetts General Hospital*
  55 Fruit Street
  Boston, MA 02114
  - Brigham and Women's Hospital*
    Boston, MA 02115
  - Boston Medical Center*
    Boston, MA 02118
  - Boston University*
    Boston, MA 02118
  - Codman Square Health Center*
    Dorchester, MA 02124

- Northwestern University*
  Chicago, IL 60611
  - Northwestern University*
    Northwester Medicine, Northwestern Memorial Hospital
    Chicago, IL 60611
  - University of Chicago*
    University of Chicago Medical Center
    Chicago, IL 60637
  - University of IL at Chicago*
    University of Illinois Hospital and Health Sciences System
    Chicago, IL 60612
  - University of IL at Chicago*
    University of IL College of Medicine Peoria
    Peoria, IL 61605
  - University of IL at Chicago*
    John H. Stroger, Jr. Hospital of Cook County
    Chicago, IL 60612
  - RUSH University Medical Center*
    Chicago, IL 60612
  - NorthShore University Health System*
    NUH Hospitals & Clinics
    Evanston, IL 60201

- University of Arizona*
  Tucson, AZ 85721
  - Banner Health*
    Phoenix, AZ 85012
  - Mariposa Community Health Center*
    Nogales, AZ 85621

- University of Pittsburgh*
  Pittsburgh, PA 15260
  - Temple University*
    Philadelphia, PA 19122
  - Chartis Group LLC

    Chicago, IL 60654

- University of Alabama at Birmingham*
  Birmingham, AL 35294
  - UAB-Selma Family Medicine*
    Selma, AL 36701
  - UAB-Huntsville Regional Medical Campus*
    Huntsville, AL 35801
  - CooperGreen Mercy Health Services*
    Birmingham, AL 35233
  - University Medical Center, Tuscaloosa*
    Tuscaloosa, AL 35401
  - University of South Alabama*
    Mobile, AL 36688
  - University of Mississippi Medical Center*
    Jackson, MS 39216
  - Louisiana State University Health Sciences Center*
    New Orleans, LA 70112
  - Tulane University*
    New Orleans, LA 70112

- Marshfield Clinic Research Institute/Marshfield Clinic Health System*
  Marshfield, WI 54449
  - University of Wisconsin – Madison*
    Madison WI, 53705
  - Medical College of Wisconsin*
    Milwaukee, WI 53226
  - Gundersen Health System*
    LaCrosse, WI 54601

- University of Miami Health System*
  Coral Gables, FL 33124
  - University of Florida*
    Gainesville, FL 32611
  - Emory University*
    Atlanta, GA 30322
  - Morehouse School of Medicine*
    Atlanta, GA 30310
  - Grady Hospital*
    Atlanta, GA 30303

FQHCs- Federally Qualified Health Centers

- Cherokee Health Systems*
  Knoxville, TN 37921
- Community Health Center, Inc.*
  Middletown, CT 06457
- au Claire Cooperative Health Center, Inc. *
  DBA Cooperative Health
  Columbia, SC 29203

- Hudson River Health Care Community Health*
  Peekskill, NY 10566
- Jackson-Hinds Comprehensive Health Center*
  Jackson, MS 39213
- San Ysidro Health Center*
  San Diego, CA 92173

VAMCs- Veterans Affairs Medical Centers

- United States Department of Veterans Affairs *
  Boston Healthcare System
  Boston, MA 02130
- United States Department of Veterans Affairs *
  Palo Alto Health Care System
  Livermore, CA 94550

The Participant Center (TPC)

- Scripps Translational Science Institute

La Jolla, CA 92037

- American College Health Association
  Silver Spring, MD 20910
- Blue Cross Blue Shield Association
  Chicago, IL 60601
- CareEvolution LLC
  Ann Arbor, MI 48104
- DXC Technologies
  Tysons, Virginia, 22102
- Examination Management Services, Inc. (EMSI)*
  Irving, TX 75063
- Harvard Medical School
  Boston, MA 02115
- Leidos Innovations Corporation*
  Reston, VA 20190
- Montage Marketing Group, LLC
  Bethesda, MD 20814
- Owaves
  Encinitas, CA 92023
- PatientsLikeMe
  Cambridge, MA 02142
- Sage Bionetworks
  Seattle, WA 98121
- San Diego Blood Bank*
  San Diego, CA 92102
- Scripps Health
  San Diego, CA 92121
- Sensis, Inc.
  Los Angeles, CA 90014
- Urban One
  Silver Spring, MD 20910
- Walgreen Co.*
  Deerfield, IL 60015-5614
- Quest Diagnostics*
  Secaucus, NJ 07094

9. Device Charges

There is no intention to commercialize or distribute this device.

10. Device Labeling

There is no intention to market this device and the final reports provided to participants will be labeled with explicit disclaimers regarding their intended use as described below.

The AoURP HDR and PGx Reports are labeled with the following sentences (see also Appendices 1 - 5):

> **Results provided are from an investigational device.** An "investigational device" is a device that is the subject of a clinical study.

The AoURP HDR Reports include the following limitations (see also Appendices 1-4):

Limitations

- **Results provided are from an investigational device.**
- **Because this report is based on data derived from a research study, this information cannot be used to diagnose, cure, mitigate, treat, or prevent disease.**
- This test may not detect all variants in the analyzed genes. The *All of Us* Research Program only reports findings within the genes that are on the panel; variants in other genes are not reported. Larger chromosomal events will also not be reported.
- In very rare cases, such as allogeneic bone marrow transplant, or recent blood transfusion (within 7 days of providing the sample), the results of this analysis may reflect the DNA of the donor. DNA quality may be affected if a participant has received chemotherapy within 120 days of providing the sample. In addition, certain organ transplants or diseases (liver, kidney, heart) may limit the relevance of the results.

The AoURP PGx Reports include the following limitations (see also Appendix 5):

Limitations

- **Results provided are from an investigational device.**
- **Because this report is based on data derived from a research study, this information cannot be used to diagnose, cure, mitigate, treat, or prevent disease.**
- This analysis does not detect all possible variants in the tested genes. When *1 (or B in the case of *G6PD*) is reported, it indicates that none of the alleles listed above were identified; it does not rule out the presence of an allele not analyzed by this test and does not rule out the possibility that a non-normal allele is present. This analysis cannot phase variants.
- The reported result may be refined as new alleles are added to the analysis.
- In some cases, observed data can be consistent with more than one possible diplotype, and in these cases the diplotype may be reported as "indeterminate".
- This analysis cannot distinguish between the more common *1/*3A and the more rare *3B/*3C diplotypes in *TPMT*; clinical phenotypic testing can distinguish between these alleles.

- In very rare cases, such as allogeneic bone marrow transplant, or recent blood transfusion (within 7 days of providing the sample), the results of this analysis may reflect the DNA of the donor. DNA quality may be affected if a participant has received chemotherapy within 120 days of providing the sample. In addition, certain organ transplants or diseases (liver, kidney, heart) may limit the relevance of the results.

11. Consent Forms and All Materials Provided to the Participant

The informed consent forms can be found in Appendices 10 through 13. Videos related to the consents can be found on the USB drive in the MISC Folder under Consent Videos.

12. Other Relevant Information

The final report from the Precision Medicine Working Group of the Advisory Committee to the NIH Director can be found in Appendix 15.

FDA comments and Meeting Minutes during pre-Submission activities are found in Appendix 16.

12.1 List of References

Berger MJ, Williams HE, Barrett R, Zimmer A, et. al. Color Data v2: a user-friendly, open-access database with hereditary cancer and hereditary cardiovascular conditions datasets. bioRxiv. 2020; doi: https://doi.org/10.1101/2020.01.15.907212

Carere DA, VanderWeele TJ, Vassy JL, et al. Prescription Medication Changes Following Direct-To-Consumer Personal Genomic Testing: Findings From the Impact of Personal Genomics (PGen) Study. Genet Med. 2017 May;19(5):537-545. doi: 10.1038/gim.2016.141.

Carey DJ, Fetterolf SN, Davis FD, et al. The Geisinger MyCode community health initiative: an electronic health record-linked biobank for precision medicine research. Genet Med. 2016;18(9):906–13.

Christensen KD, Phillips KA, Green RC, Dukhovny D. Cost Analyses of Genomic Sequencing: Lessons Learned from the MedSeq Project. Value Health 2018;21:1054-61.

Christensen K, Vassy J, Phillips K, et al. Short-term costs of integrating whole-genome sequencing into primary care and cardiology settings: a pilot randomized trial. Genet Med 2018;Epub ahead of print.

Cleary JG, et al. Comparing Variant Call Files for Performance Benchmarking of Next-Generation Sequencing Variant Calling Pipelines. bioRxiv. 2015.

Costello, M., Fleharty, M., Abreu, J. et al. Characterization and remediation of sample index swaps by non-redundant dual indexing on massively parallel sequencing platforms. BMC Genomics 19, 332 (2018). https://doi.org/10.1186/s12864-018-4703-0}

Effects of index mismanagement. Illumina Tech Note. 2017. https://www.illumina.com/content/dam/illumina-marketing/documents/products/whitepapers/index-hopping-white-paper-770-2017-004.pdf

Gordon AG, Zouk H, Venner E et al. Frequency of genomic incidental findings among 21,915 eMERGE network participants. Genet Med. In press.

Hart MR, Biesecker BB, Blout CL, et al. Secondary findings from clinical genomic sequencing: Prevalence, patient perspectives, family history assessment, and healthcare costs from a multi-site study. Genet Med. 2019 May;21(5):1100–1110. doi:10.1038/s41436-018-0308-x.

Lee SB, et al. Calling Star Alleles With Stargazer in 28 Pharmacogenes With Whole Genome Sequences. Clinical pharmacology and therapeutics. 2019;106(6):1328-37. Epub 2019/06/18. doi: 10.1002/cpt.1552. PubMed PMID: 31206625; PMCID: PMC6896231.)

Neben CL, Zimmer AD, Stedden W, et al. Multi-Gene Panel Testing of 23,179 Individuals for Hereditary Cancer Risk Identifies Pathogenic Variant Carriers Missed by Current Genetic Testing Guidelines. J Mol Diagn. 2019;21(4):646-657. doi:10.1016/j.jmoldx.2019.03.001

Perkins BA, Caskey CT, Brar P, et al. Precision medicine screening using whole-genome sequencing and advanced imaging to identify disease risk in adults. Proc Natl Acad Sci U S A 2018;115:3686-91.

Pratt VM, et al. Characterization of 137 Genomic DNA Reference Materials for 28 Pharmacogenetic Genes: A GeT-RM Collaborative Project. J Mol Diag. 2016 18:109-123 PMID: 26621101.

Pritchard CC, et al. ColoSeq provides comprehensive lynch and polyposis syndrome mutational analysis using massively parallel sequencing. J Mol Diagn. 2012;14(4):357-66. doi: 10.1016/j.jmoldx.2012.03.002. PubMed PMID: 22658618; PMCID: PMC3391416.

Pritchard CC, et al. Validation and implementation of targeted capture and sequencing for the detection of actionable mutation, copy number variation, and gene rearrangement in clinical cancer specimens. J Mol Diagn. 2014;16(1):56-67. doi: 10.1016/j.jmoldx.2013.08.004. PubMed PMID: 24189654.

Roberts J, Robinson J, Diamond P, et al. Patient understanding of, satisfaction with, and perceived utility of whole-genome sequencing: findings from the MedSeq Project. Genet Med. 2018 September; 20(9): 1069–1076. doi:10.1038/gim.2017.223.

Sanderson SC, Linderman MD, Suckiel SA, et al. Psychological and behavioural impact of returning personal results from whole-genome sequencing: the HealthSeq project. Eur J Hum Genet. 2017;25(3):280–92.

The eMERGE Consortium , Harmonizing Clinical Sequencing and Interpretation for the eMERGE III Network. AJHG, 2019; 105(3):588-605 https://doi.org/10.1016/j.ajhg.2019.07.018

Van der Auwera, et al. From FastQ data to high confidence variant calls: The Genome Analysis Toolkit Best Practices Pipeline. Curr Protoc Bioinformatics. 2013. 43. PMID: 25431634

Vassy JL, Brunette CA, Majahalm N, et al. The Integrating Pharmacogenetics in Clinical Care (I-PICC) Study: Protocol for a Point-Of-Care Randomized Controlled Trial of Statin Pharmacogenetics in Primary Care. Contemp Clin Trials 2018;75:40-50. doi:10.1016/j.cct.2018.10.010.

Vassy JL, Christensen KD, Schonman EF, et al. The impact of whole-genome sequencing on the primary care and outcomes of health adult patients: a pilot randomized trial. Ann Intern Medicine. 2017;167(3):159-169. doi: 10.7326/M17-0188.

Zoltick ES, Linderman MD, McGinniss MA, et al. Predispositional genome sequencing in healthy adults: design, participant characteristics, and early outcomes of the PeopleSeq Consortium. Genome Med. 2019;11:10. *https://doi.org/10.1186/s13073-019-0619-9*.

Zook JM, et al. An open resource for accurately benchmarking small variant and reference calls. Nature biotechnology. 2019;37(5):561-6. Epub 2019/04/03. doi: 10.1038/s41587-019-0074-6. PubMed PMID: 30936564; PMCID: PMC6500473.

13. List of Appendices (and Files in MISC Folder)

Appendix 1 – HDR Uninformative Report

Appendix 2 – HDR Positive *BRCA1* Report

Appendix 3 – HDR Positive *MSH2* Report

Appendix 4 – HDR Positive *LDLR* Report

Appendix 5 – Medicine and Your DNA Report

Appendix 6 – PGx Translations Table

Appendix 7 – PGx Rubric / Drug Evidence Table

Appendix 8 - Brief Report of Comprehension Testing

Appendix 9 – AoURP FDA Investigator Agreement

Appendix 10 – Research eConsent

Appendix 11 – gRoR eConsent

Appendix 12 – Full-length Research Consent

Appendix 13 – Full-length gRoR Consent

Appendix 14 – Compliance with 21 CFR 11

Appendix 15 – ACD Working Group Report

Appendix 16 – FDA Comments and Meeting Minutes

MISC Files Folder:

Consent Videos:

001_Research eConsent Video 1-Welcome

002_Research eConsent Video 2-Keeping in Touch

003_Research eConsent Video 3-Health Data

004_Research eConsent Video 4-Health Records

005_Research eConsent Video 5-Physical Measurements

006_Research eConsent Video 6-Samples

007_Research eConsent Video 7-Fitness Tracker

008_Research eConsent Video 8-Other Health Data

009_Research eConsent Video 9-What will we do with your samples

010_Research eConsent Video 10-Data Sharing

011_Research eConsent Video 11-Potential Risks

012_Research eConsent Video 12-Research Results

013_Research eConsent Video 13-Seeing your Data

014_gRoR eConsent Video 1-What are DNA changes

015_gRoR eConsent Video 2-How will you check my DNA and how long will it take

016_gRoR eConsent Video 3-What will you check for

017_gRoR eConsent Video 4-What will you tell me

018_gRoR eConsent Video 5-How could finding out about my DNA results help me

019_gRoR eConsent Video 6-What are the risks

All of Us Research Program[1] (AoURP): Return of Genetic Results (gRoR)

IDE Sponsor: Stephanie Devaney, Ph.D., Chief Operating Officer of the *All of Us* Research Program, National Institutes of Health (NIH)

Funded by: *All of Us* Research Program, National Institutes of Health

Version Number: 1.0

June 3, 2020

As amended November 18, 2020 to include changes requested by FDA during review.

Additions to original submission are highlighted.

[1] Precision Medicine Initiative, PMI, *All of Us*, the *All of Us* logo, and "The Future of Health Begins with You" are service marks of the U.S. Department of Health and Human Services.

Table of Contents

STATEMENT OF COMPLIANCE

The study will be conducted in accordance with National Institutes of Health (NIH) Terms and Conditions of Award to Investigator organizations. The Responsible Investigators will assure that no deviation from, or changes to, the protocol will take place without prior agreement from the Investigational Device Exemption (IDE) Sponsor and documented approval from the *All of Us* Research Program (AoURP) Institutional Review Board (IRB), except where necessary to eliminate an immediate hazard(s) to the research participants. All personnel involved in the conduct of this study are required to complete Human Subjects Protection Training annually.

The protocol and all participant materials specific to this study will be submitted to the IRB for review and approval prior to the Return of Genetic Results study launch. Any amendment to the protocol will require review and approval by the IRB before the changes are implemented to the study. All consent forms have been IRB approved (*AoU* IRB approval v9: Mar 20, 2020).

1 PROTOCOL SUMMARY

1.1 SYNOPSIS

Title:	*All of Us* Research Program (AoURP): Return of Genetic Results (gRoR)
Study Description:	The study involves clinical interpretation of pre-defined heritable disease and pharmacogenetic associated genes and return of those interpretations to participants who consent to receive results, with all appropriate disclaimers on the use of the results. There is no intervention nor is there an assessment of safety or efficacy outcomes. This is an investigational device study because results being returned to participants are not generated from a cleared or approved assay validated for clinical use. The study is embedded in a broader program of research, the *All of Us* Research Program (AoURP). The AoURP is a longitudinal observational cohort program with repeated engagement of participants to create a research resource (registry/repository) that will enable a wide range of scientific questions on health and disease. A core value of the AoURP is to return value to participants, including potentially important health-related genetic results gleaned from whole genome sequencing (WGS). The AoURP enrolled its first participant prior to the revisions to the Common Rule in 2018 and complies with the pre-2018 regulation which is harmonized with current FDA regulations 21 CFR 50 and 21 CFR 56. The gRoR protocol will be implemented as a sub-protocol of the AoURP.
Objective:	**To ethically and responsibly return health-related genetic results to AoURP participants.**
Endpoints:	There are no endpoints.
Study Population:	One million participants. Participation in the gRoR protocol is open to participants of the AoURP which aims to enroll 1 million participants or more throughout the United States.
Phase:	Not applicable
Description of Sites/Facilities Enrolling Participants:	All participants in the gRoR protocol must be participants in the AoURP. Participants in the AoURP are recruited and enrolled through hundreds of sites overseen by selected Health Care Provider organizations (HPOs), including Regional Medical Centers (RMCs), Federally Qualified Health Centers (FQHCs), Veteran's Affairs Medical Centers (VA) and other contract partners. All sites are in the United States.
Description of Study Intervention:	No invervention is being provided.
Study Duration:	The AoURP is expected to last for decades, with active enrollment anticipated for the first five years.
Participant Duration:	Participation is expected to last for the duration of the Program.

1.2 SCHEMA

Figure 1: High level overview of the flow of samples and information for the AoURP:gRoR protocol

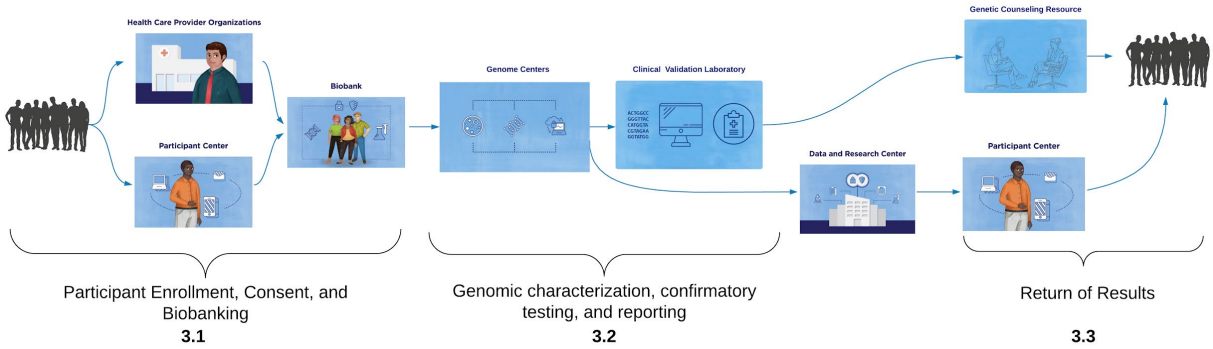
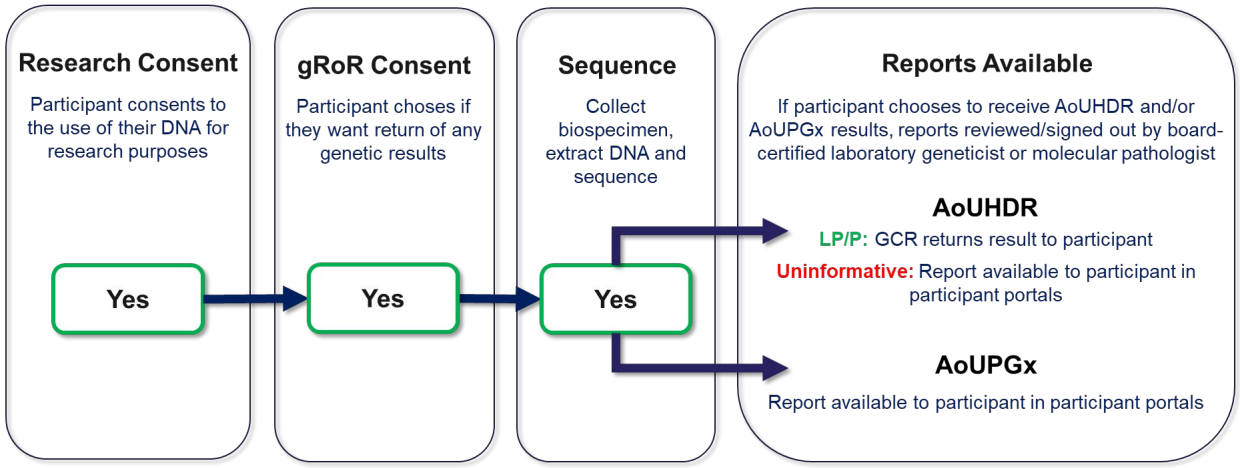


Figure 2: Participant perspective for consent to return of genetic results. Abbreviations: AoUHDR – *All of Us* hereditary disease risk, AoUPGx – *All of Us* pharmacogenetics, GCR – Genetic Counseling Resource, P/LP –pathogenic/likely pathogenic.



1.3 SCHEDULE OF ACTIVITIES (SOA)

The AoURP does not institute a set schedule of activities as the innovative design of the study provides each participant with the opportunity to choose a custom experience. For example, a participant may choose to donate data from a personal health tracker to the study or choose whether to receive individual genetic results or decline to receive such results. There are, however, a core set of activities at the time of initial enrollment that are necessary for a participant to complete to be eligible for the return of genetic results.

The core activities for participants for this protocol are as follows:

1. Create account via website (*http://joinallofus.org*) or mobile application
2. Review and electronically sign informed consent forms: 1. research consent and 2. return of genetic results consent
3. Complete Basics Questionnaire (contact information; sociodemographic information)
4. Biospecimen collection (i.e., blood draw)
5. Respond to informing loops to receive specific results

2 INTRODUCTION

2.1 STUDY RATIONALE

The AoURP seeks to return value to participants, specifically, the interpretation of important health-related genetic results gleaned from whole genome sequencing (WGS).

This return of genetic results (gRoR) protocol describes the AoURP's strategy for ethically and responsibly returning individual genetic results, specifically hereditary disease risk and pharmacogenetics, directly to participants if they choose to receive them. All reports will include appropriate disclaimers that these are RESEARCH RESULTS from an investigational device.

2.2 BACKGROUND

In March 2015, the National Institutes of Health formed the Precision Medicine Initiative Working Group of the Advisory Committee to the Director. The group produced a report (Appendix 15: ACD Working Group Report) that provided the framework for the creation of the *All of Us* Research Program – a collaboration of health care providers, researchers, technology experts, community partners and the public creating a new era of individualized healthcare (See *https://allofus.nih.gov*).

The return of genetic information to patients within the clinical context falls into two broad categories. First, the return of primary findings from targeted clinical genetic assessment either for disease diagnosis or for medication sensitivity (pharmacogenetics). Second, the return of "incidental" findings – results gleaned from clinical genetic assessments that are unrelated to the primary reason for the clinical test.

Advancements in genetic analysis and the integration of genetic results into clinical care set off a robust debate regarding the clinical return of incidental findings for disease susceptibility that continued for more than a decade. The (arguable) denouement arrived in 2013 with the publication of a recommended disclosure list for incidental findings by the American College of Medical Genetics and Genomics (ACMG). Laboratories are now encouraged to return these select incidental findings in the clinical context (see Green et al., 2013; ACMG Board of Directors, 2015; Kalia et al., 2017). This has been described as a form of "opportunistic screening," considered consistent with other types of medical practice like x-rays or skin examinations but continues to be challenged as an unproven form of "population screening." Nevertheless, offering patients the option of incidental findings from clinical genetic assessment has largely become standard, with more than 90% of adult patients choosing to receive them.

In parallel, results are already being returned to patients within primary pharmacogenetics implementations at several large academic medical centers and pharmacogenetic information is included in US Food and Drug Administration- (FDA-) approved labeling for over 190 medications. In several instances (e.g., abacavir, codeine, clopidogrel), there are boxed warnings recommending the use of genetic information to "guide medication or dosage selection when results are available." (see Table of Pharmacogenomic Biomarkers, 2018). The Clinical Pharmacogenetics Implementation Consortium (CPIC, https://cpicpgx.org/) has published over 20 guidelines for 38 medications that have the highest level of evidence (CPIC-A), defined as medications where the "preponderance of evidence is high or moderate in favor of changing prescribing" based on genetic test results. FDA and CPIC data sources are largely congruent, but some differences exist.

In the research context, many stakeholders now urge that all results of medical relevance be made available to adult participants who want them out of respect for participant autonomy (see Shalowitz &

Miller, 2005; Angrist, 2011; and Terry & Terry, 2011). This view is consistent with the AoURP Core Values (https://allofus.nih.gov/about/about-all-us-research-Program). However, the return of genetic results has been more controversial than return of physical measurements, assays, or imaging in the research context for a number of reasons, including concerns related to:

- Risks to individual privacy posed by genetic data/results
- Relational risks of genetic data/results (i.e., risks posed to family or communities who share genetic information)
- Risks of genetic discrimination
- Novelty of genetic results in most clinical care settings
- Uncertainty regarding the long-term risks and benefits of providing health-related genetic results to otherwise healthy individuals
- Unknown long-term costs of providing health-related genetic results to otherwise healthy individuals

Additionally, critics have raised concerns that in the research context, the return of genetic results may distress participants, motivate unnecessary (and potentially dangerous) follow-up care, and place untenable time and financial demands upon researchers.

Empirical data refute the critics. Specifically, the MedSeq Project (Vassy et al., 2014), a randomized clinical trial utilizing WGS to return health-related genetic findings to apparently healthy adult participants, demonstrated that participants receiving these results did not experience undue distress (Roberts, et al., 2018). Further, Vassy et al. (2017) reported that primary care physicians managed these findings without serious errors and that short-term downstream healthcare costs were not significantly greater for those receiving these results (Christensen, Phillips et al., 2018; Christensen, Vassey et al., 2018; Perkins et al., 2018; Zoltick et al., 2019).

Based on these findings and others, consensus has gradually emerged supporting the return of health-related results to adult research participants, although concerns about comprehension and inappropriate response to such information remain. The ACMG recommendations for the return of incidental heritable disease risk findings in the clinical context, while not originally designed for return in the research context or for population screening, have become a convenient, concrete starting point to use for these purposes; likewise, the list of CPIC-A genes offers a foundation for considering the return of evidence-based pharmacogenetic results.

With the rapid developments in genomic science and medicine, the AoURP is determined to offer genetic results to its participants and, toward that end, has established a consortium of College of American Pathologists/Clinical Laboratory Improvement Amendments (CAP/CLIA) certified genomics laboratories to analyze the DNA from participants who affirmatively consent to receiving their genetic results. Specifically, this protocol describes the procedures for the use of WGS data to return hereditary disease risk and pharmacogenetic results directly to AoURP participants.

2.3 RISK/BENEFIT ASSESSMENT

2.3.1 KNOWN POTENTIAL RISKS

Participation in the AoURP:gRoR protocol offers participants the opportunity to explore one's genetic information in a manner that may illuminate or impact certain health-related experiences. It is an opportunity not often available to many people and presents significant potential for individual and group benefits, both in terms of education and downstream effects on quality of life.

While there are significant benefits, participation is not without risk, and these risks should be scrutinized. *All of Us* differs from many traditional studies, as it does not include activities typically associated with risk of physical harm. As a component of the AoURP, the gRoR protocol inherits the risks associated with participation as a whole, including the risk of discomfort or reaction to the blood draw. However, there are risks that are unique to the gRoR element of AoURP, such as incomplete or incorrect results, potential privacy and security incidents associated with the generation and storage of results, emotional or psychological harms, and physical harm associated with the misuse of result information. These risks are detailed below.

2.3.1.1 INCOMPLETE OR INCORRECT RESULTS

However unlikely, participants may receive inaccurate or incomplete results for several reasons, some of which are beyond the scope of the device.

Return of incomplete or incorrect results due to limitations of the current science. Because the scientific evidence that has formed the basis for our current knowledge of health-related genetic variants is based on study populations with limited ancestral diversity, we may be unable to determine the full scope of expectations and risk associated with those variants for individuals who do not descend from largely European ancestry. Without more investigation into variant effects in diverse populations, or other factors that may impact penetrance, we are unable to assure that all AoURP participants will receive the same level of benefit from their results.

These limitations are discussed throughout consent and return materials. Participants are made aware that their results may change over time as we learn more. Furthermore, the program may update and expand findings as scientific advances in this area are made and our understanding changes (see Section 6.1.4.1 AoUHDR results - Updating results over time for further details).

Return of incomplete or incorrect results due to malfunctions in the device. As with any study, there are risks of device malfunction that could result in participants receiving less than complete or accurate results. The following steps are taken to mitigate these risks:

- Analysis is performed in clinical labs with stringent regulatory and compliance programs, controlled under the total Quality Management Systems (QMS) based on Clinical and Laboratory Standards Institute (CLSI) Guidelines.
- Analysis is conducted using clinically validated laboratory developed tests (LDTs)
- All reports will be reviewed and signed out by a board-certified laboratory geneticist or molecular pathologist (consistent with practice of medicine).
- All positive Hereditary Disease Risk (HDR) results will be confirmed with an orthogonal and medically established method.

Return of incomplete or incorrect results due to **human error** (including mislabeling, contamination, etc.). To the extent possible, processes are automated to minimize the opportunity for human error. Centralized data tracking and exchange systems, laboratory information management systems, and robotic sample handling for AoURP incorporate the most advanced tools available today, as further described below.

2.3.1.2 PRIVACY AND SECURITY INCIDENTS

Risk of breach. As with any system that stores, maintains, or transfers data, there is a risk that sensitive participant information will be exposed.

The AoURP security approach is a combination of applying consistent standards through compliance with all applicable federal laws and regulations and industry best practices supporting multiple

technology platforms and interconnections. AoURP is committed to safeguarding participant data on systems transmitting, distributing, and storing highly sensitive information. Security controls derived from National Institute of Standards and Technology (NIST) Special Publication 800-53 (Security and Privacy Controls for Federal Information Systems and Organizations) and are implemented in accordance with the Precision Medicine Initiative Data Security Policy Principles and Framework (PMI-DSPP). AoURP systems use the NIST Risk Management Framework to govern a structured authorization process that determines the risk profile so the appropriate level of security controls can be implemented and thoroughly tested. All systems supporting AoURP must implement NIST 800-53 security controls at MODERATE for minimum assurance and allow for tailoring of security controls and configurations to meet unique mission requirements.

The AoURP adheres to the HHS Policy and Plan for Preparing and Responding to a breach of Personally Identifiable Information (PII). All awards follow NIH guidelines for preparing for (via security authorization process) and responding to a breach. With the advent of new technologies to support the gRoR, the NIH must be prepared for breach scenarios requiring the implementation of security and privacy safeguards accordingly. gRoR security and privacy safeguards must support the following data states and will include the specific safeguards described.

- AoURP data in transit - Example breach scenario - Malicious actor is able to eavesdrop on gRoR between *All of Us* systems and participant. Safeguard - The Program implements strong encryption techniques while adhering to NIST federal standards.
- AoURP data in storage - Example breach scenario - Hacker exfiltrated data from database after gaining fraudulent access. Safeguard - The AoURP relies on our large cloud partners to encrypt data at several layers while ensuring the platform is secure and implementing defense in depth controls to apply access configurations and audit logging.
- AoURP data in distribution - Example breach scenario - Unintended access to incorrect recipients could result from poor identity verification standards. Safeguard - The AoURP implements rigorous identity verification steps and subsequent account management access rules that ensure the correct recipient is provided the correct gRoR report(s).

Disclosures and Misuse (by law enforcement, insurers, employers, etc.). The program has been issued Certificates of Confidentiality, which limit the allowable disclosures and uses of data generated under its auspices. The Certificates of Confidentiality extends to data held by AoURP awardees and sub-contractors and covers all copies of AoURP-generated identifiable, sensitive data.

The program will not disclose a participant's results to any person or entity other than the participant, with the following exceptions.

1. At the request of the participant to have their family be present for or to take part in conversation with an *All of Us* genetic counselor. The participant must also be present; or
2. Results may be returned to a participant's care provider only if: 1) required by state law; and 2) the participant has explicitly consented/requested this disclosure.

2.3.1.3 BIOSPECIMEN COLLECTION

Blood sampling risks include bruising of the arm and fainting. The modest amount of blood drawn, up to 50 mL, should not have any adverse physiological effects.

2.3.1.4 PSYCHOLOGICAL DISTRESS

Receipt of genetic results that indicate potentially serious health risks could cause psychological distress. As mentioned in Section 2.2, recent research indicates that this is a minimal risk. Even so, it is an

important risk to consider and provide mitigation for. This risk is mitigated by extensive informed consent, educational materials, and disclosure of P/LP HDR results over the phone by a genetic counselor.

Emotional/psychological distress from fear for health of self/loved ones. Materials germane to consent for, and return of, results contextualize results as within the confines of their meaning and limitations. Additionally, prior to choosing to receive specific types of health-related results (e.g., HDR or pharmacogenetic [PGx]) the program provides tailored educational materials on that type of result. This includes an overview of the risks of receiving such results, such as emotional or psychological distress. Participants can choose not to receive results if they are unwilling to assume that risk.

Participants may also reach out for support from the Genetic Counseling Resource (GCR). The GCR is empowered to help people find clinical or psychological care to assist them, if necessary, and will provide a "warm handoff" of the participant to this care when possible.

Emotional/psychological distress from receiving results that make people question group/family membership. This is a relatively low risk from health-related results. This risk is discussed in the consent and educational materials. Participants can choose not to receive results if they are unwilling to assume that risk.

The GCR is available to assist with interpreting the meaning of results with regards to, not just health, but also relatedness and identity.

Importantly, the program is not returning familial results (i.e., no relatedness linkages or information) to minimize occurrence of this outcome. Educational materials clearly state the results will not tell participants information about who they are related to.

Emotional/psychological distress from navigating family discussions of results. The informed consent form, educational materials, and reports contain language presaging the potential need for family discussion of health-related results. The GCR will facilitate conversations with participants and their family members at the request of the participant and only with the participant present as part of the conversation.

2.3.1.5 MISUSE OF INFORMATION

The greatest risk associated with gRoR is the possibility that a participant decides to make changes in their medical care without clinical confirmation or consultation with their healthcare provider.

Changes made to care without consult of a healthcare provider. Consent, educational materials, and result reports emphasize the nature of these results as research results and not for use in clinical care. This includes multiple disclaimers emphasizing these points, such as, "these are research results and your doctor will need to confirm with a clinical test before using them in your care." See Appendices 2, 3, 4, and 5 for examples of positive reports.

Further mitigation is provided by return of HDR P/LP results through a meeting with a licensed genetic counselor at the GCR, who can help to reinforce safety issues. Whenever interacting with a participant, genetic counselors will emphasize that the AoURP is providing research results that will need to be confirmed by a doctor or other healthcare provider prior to making any changes in medical care and offer to connect the participant to a provider, if needed.

2.3.2 KNOWN POTENTIAL BENEFITS

Potential benefits of participation include:

- Learning about possible serious health risks. Health conditions may be prevented or discovered which could lead to earlier intervention and more appropriate treatment.
- Potential to learn about previously unknown health resources (e.g., tests for certain types of medication reactions).
- Opportunity to increase general genetic and scientific literacy.

2.3.3 ASSESSMENT OF POTENTIAL RISKS AND BENEFITS

We acknowledge potential risks that may be incurred by study participants and we have strategies in place to minimize these risks. Further, participants may, in fact, derive benefits from learning about their health-related genetic information. Taken together we have determined that overall benefits outweigh the risks of participation.

3 OBJECTIVES AND ENDPOINTS

The primary objective of this protocol is to ethically and responsibly return individual genetic results directly to AoURP participants who have affirmatively consented to receive this information. Delivery of the genetic reports to participants would be considered the endpoint.

4 STUDY DESIGN

4.1 OVERALL DESIGN

The AoURP is a large longitudinal cohort program, aimed at enrolling 1 million people or more across the United States, with repeated engagement of participants **to create a research resource** that enables a variety of studies. A vast and varied network of enrollment partners have been established and funded through competitive federal awards to achieve the goals of the aims of the AoURP, creating a research cohort of unparalleled diversity. On the date of submission of this application, the AoURP has over 300 enrollment sites across the country where participants can complete the informed e-consent and be assessed for simple physical measurements and donation of biospecimens (primarily blood and urine). All sites are directed by AoURP awardees or inter-agency partners (VAMCs). The Principal Investigators from these awardees/partners are included as co-Responsible Investigators in the study. Throughout the life of the Program, we anticipate that the number of enrollment sites will fluctuate depending on site performance and programmatic needs. However, we anticipate that the number of enrollment sites will not exceed 500 throughout the duration of the Program. In addition, participants can enroll in the program directly through the same online process at *joinallofus.org*.

This protocol describes the process of ethically and responsibly returning genetic results to AoURP participants who affirmatively consent to wanting to receive these results. For these participants, WGS data will be analyzed for clinical interpretation and reports will be provided directly to them.

4.2 SCIENTIFIC RATIONALE FOR STUDY DESIGN

Not applicable. The sole purpose of this study is to return value to AoURP participants as data are generated for the AoURP research resource.

4.3 JUSTIFICATION FOR DOSE

Not applicable. No intervention is being provided.

4.4 END OF STUDY DEFINITION

The AoURP is a longitudinal study that is anticipated to last for decades, with both active and passive participation opportunities throughout the span of the study. The option to receive genetic results is offered by way of an additional gRoR informed consent process after completion of the initial Research Consent that constitutes enrollment in the AoURP.

5 STUDY POPULATION

5.1 INCLUSION CRITERIA

All AoURP participants who provide biospecimens and affirmatively consent that they want to receive their individual genetic results are eligible for this protocol. Eligibility for inclusion in the AoURP is as follows:

1. Participants, aged 18 or older, with the legal authority and decisional capacity to provide a signed and dated informed consent forms
2. Currently residing in the US or a US territory

5.2 EXCLUSION CRITERIA

AoURP participants who indicate that they do not want to receive their genetic results will be excluded from this protocol.

For the AoURP as a whole, an individual who meets any of the following criteria will be excluded from participation in the study:

1. Persons who are currently incarcerated at the time of enrollment
2. Individuals who rely on a legally authorized representative to provide consent (this could be for physical or decisional reasons)
3. Persons under the age of 18
4. Individuals who cannot provide informed consent in either English or Spanish

Additionally, if it is learned that a participant has become incarcerated, their participation will be suspended until such time as the AoURP is permitted to allow for incarcerated individuals to participate or the participant is no longer incarcerated.

5.3 LIFESTYLE CONSIDERATIONS

None. All persons meeting the criteria above will be eligible for participation.

5.4 SCREEN FAILURES

Not applicable. There is no screening for study participation beyond initial eligibility requirements of the broader AoURP.

5.5 STRATEGIES FOR RECRUITMENT AND RETENTION

Not applicable. Participants are only recruited for the overall longitudinal observational AoURP, not this gRoR protocol.

6 STUDY PROCEDURES

6.1 STUDY ADMINISTRATION

This is not an interventional study; therefore, this section describes the study procedures.

The procedures are outlined as follows with operational details contained in subsequent sections.

- Biospecimens are received and processed at the AoURP Biobank.
    - Genomic DNA from whole blood is extracted, aliquoted, plated and shipped to the AoURP Genome Centers (GCs) for genomic analysis.
- All samples and workflows are tracked by the Data and Research Center (DRC) and monitored by the Sponsor.
- AoURP GCs run samples through WGS pipelines and submit resultant data files to the DRC.
- Participants are presented with a series of informing loops, reminding participants of risks and limitations of the results as also described in the informed consent (which may have been signed months prior), and allowing the participant to choose the specific results they want to receive. For example, someone may want their hereditary disease risk (AoUHDR) results but not the pharmacogenetic (AoUPGx) results, or vice versa. Participants cannot choose to receive a report on specific genes. The AoUHDR results will be contained in a single report and the PGx results will be contained in a separate, complete report.
- After a participant has completed the informing loops indicating which specific results they would like to receive, the WGS data files are sent to the clinical validation laboratories (CVL) for AoUPGx and/or AoUHDR variant interpretation.
- Interpretation results are compiled in a study-specific environment, the Reporting and Harmonization Platform (RHP). The RHP is a secure automated system that checks consistency of interpretation across the CVLs to ensure that any specific variant is classified with the same interpretation. Any discrepancy in interpretation is flagged for attention by the originating CVL. Reports are rendered in the RHP and laboratory directors are provided workspaces to review and sign-out reports for release to participants.
- After reports are released, participants will be notified that results are available in their participant portal **except** when a P/LP variant is found in a hereditary disease risk gene(s).
    - If a participant tests positive for a P/LP variant in one or more of the hereditary disease risk genes, a replicate sample from that participant will be retrieved from the AoURP Biobank and sent to the original CVL to run an orthogonal confirmatory assay.
    - If the original result is confirmed, the participant will be contacted to schedule an appointment with a genetic counselor at the GCR who will communicate the results and answer any questions or concerns the participant may have.
- Additional support will be offered to all participants through extensive educational materials and unlimited live call center support. Genetic counseling appointments will be available to all participants who receive AoUHDR or AoUPGx results if desired.

6.1.1 BIOSPECIMEN PROCESSING AND PLATING

Additional information on rigorous procedures for labeling and tracking biospecimens can be found in *Section 10.5.4 Quality Assurance and Quality Controls*. The AoURP Biobank receives, processes, and extracts DNA from 4 mL Ethylene diamine tetraacetic acid (EDTA, whole blood) or 10 mL EDTA (buffy coat). DNA is extracted from blood by two methods; salt-based precipitation method on Autogen FlexStar or bead-based method on Chemagen 360. DNA samples are checked for volume via BioMicroLab volume check instrument. DNA samples are also quantified (spectrometric method) via Lunatic-Unchained Labs / Trinean DropSense 96 to obtain total DNA concentration as well as A260/280 and A260/230. For samples to meet GC Quality Control (QC) criteria, all samples must have a minimum concentration of 50 ng/uL **and** A260/280 in the range of 1.6-2.0. Samples meeting these criteria are aliquoted using the Perkin Elmer JANUS Automated Liquid Handler into Thermo Scientific 0.75 mL 2-D matrix tubes (catalog # 3732) with SepraSeals. Aliquots are diluted with water as necessary to reach a

concentration of 60 ng/uL and a volume of 40 uL. Samples requiring clinical validation are aliquoted into 96-well Bio-Rad Hard Shell Full Skirt plates (catalog# HSP-9601) at a concentration of 60 ng/uL and a volume of 50 uL. Upon completion of plating, sample tubes are labeled via the Scinomix Automated Tube Labeler and stored at -80°C until time of shipment. Samples collected in New York State will be sorted into separate boxes and the aliquots sequestered into specific storage locations until time of shipment to ensure they are only sent to the appropriately accredited sites.

6.1.2 BIOBANK SHIPPING TO GENOME CENTERS

The Biobank will ship frozen boxes and plates containing DNA aliquots stabilized with dry ice to the GCs via FedEx Priority Overnight shipping services. Sample packages will be International Air Transport Association (IATA) compliant using shipping materials validated by Mayo Clinic, including Styrofoam containers, 95 kPa biohazard bags, and applicable shipping placards. Prior to packaging, an automated box scan of the 2-D barcodes will be conducted to verify sample location and the presence or absence of New York state samples using the Research Laboratory Information Management System (RLIMS). Each package shipped will contain a condensed paper manifest and a unique barcoded package ID for tracking. Additionally, a detailed electronic manifest will be created based on the samples and boxes in the package and dropped into the appropriate site's Google bucket on the night of shipment. Manifests are named based on the recipient GC, the intended test for the samples, and the associated package ID for matchup. The GCs will receive automated FedEx tracking notifications and updates for all shipments.

6.1.3 GENOMIC CHARACTERIZATION

Sample preparation for WGS follows the validated processes at each of the AoURP GCs. Several QC checks are built into the processes including, but not limited to, DNA quantity and quality measurements; library yields; molecular barcode addition; pipeline cross-checks for barcode demultiplexing; and sequencing data quality checks with established and harmonized performance specifications across sites. All sites are using the Illumina NovaSeq system for sequence data generation. Details on data generation and analyses and test performance characteristics are provided in IDE Section 2.2 and are summarized here.

Steps in the WGS pipeline include:

- Extracted DNA is fragmented, adapter ligated, and barcoded without PCR amplification.
- Library fragments are sequenced (2 x 150 base paired end) using Sequencing-By-Synthesis (SBS) chemistry and the Illumina NovaSeq 6000 sequencer.
- Sequence data are aligned to the specified National Center for Biotechnology Information (NCBI) human reference sequence after discarding low quality sequences.
- Reads that are aligned to more than one region of the reference genome, reads with low alignment scores, and bases with low quality scores are excluded from variant calling.
- All GC sites use a harmonized version and parameter set for variant calling (currently the GATK-DRAGEN v3.4.12).

6.1.4 VARIANT INTERPRETATION AND REPORTING

Variant call files (VCFs) are analyzed by pipelines at each of the CAP/CLIA CVLs using a common bed file inclusive of AoUHDR and AoUPGx loci to identify and interpret all variants within these predefined reportable genomic regions (described in more detail in the IDE).

6.1.4.1 AOUHDR RESULTS

All variants in the 59 genes listed below that are identified in a given case by a CVL are classified as Pathogenic (P), Likely Pathogenic (LP), Uncertain Significance (VUS), Likely Benign (LB), Benign (B) or Non-Reportable (used when a variant is determined to not be P or LP but is not further classified as VUS/LB/B which is not a requirement). All CVLs follow professional practice guidelines for variant classification published by the ACMG and the Association for Molecular Pathology (Richards et al 2015), including additional guidance provided by the ClinGen Sequence Variant Interpretation Working Group as well as disease and gene-specific specifications provided by ClinGen Expert Panels, all maintained on ClinGen's website (*https://www.clinicalgenome.org/working-groups/sequence-variant-interpretation*).

The Hereditary Disease Risk Report (HDRR) comprises analyses of 59 genes: ACTA2, ACTC1, APC, APOB, ATP7B, BMPR1A, BRCA1, BRCA2, CACNA1S, COL3A1, DSC2, DSG2, DSP, FBN1, GLA, KCNH2, KCNQ1, LDLR, LMNA, MEN1, MLH1, MSH2, MSH6, MUTYH, MYBPC3, MYH11, MYH7, MYL2, MYL3, NF2, OTC, PCSK9, PKP2, PMS2, PRKAG2, PTEN, RB1, RET, RYR1, RYR2, SCN5A, SDHAF2, SDHB, SDHC, SDHD, SMAD3, SMAD4, STK11, TGFBR1, TGFBR2, TMEM43, TNNI3, TNNT2, TP53, TPM1, TSC1, TSC2, VHL, and WT1.

To ensure consistency of results being returned by the CVLs, plus the Partners Laboratory for Molecular Medicine (LMM) at Massachusetts General Hospital, which also is participating in this activity for AoURP, all four labs shared all variant classifications in the ACMG 59 genes from their respective databases (49,943 unique variants), of which 23.7% (11,813 variants) were classified by at least two laboratories. The pre-test launch data exchange showed 98.5% (11,631/11,813) concordance in variant classification at the AoU reporting level with 0.68% (80/11,813) discrepancies in pathogenicity confidence (P vs LP) and 0.86% (102/11,813) discrepancies in variant reportability (P/LP vs VUS/LB/B). To date, 56% (57/102) of reportability discrepancies have so far been resolved, with remaining discrepancies accounting for only 0.4% (45/11,813) of all variants in common across laboratories. We continue to work to resolve the remaining differences in interpretation between CVLs and LMM, to ease the burden of real-time resolution once the return of results begins.

Regardless of the state of pre-harmonization efforts, all variants returned to participants will not be returned unless there is concordance on the interpretation across all CVLs. To ensure continuous classification harmonization within the AoURP, a database of all variant classifications in the ACMG 59 genes from CVLs used in pre-harmonization will be stored within the AoURP RHP. The RHP used by all CVLs will track all reported variants and the reported classification within AoURP to ensure consistency at the time of reporting. If a new classification differs from the classification already on file in the database, the CVLs will first try to resolve the difference by exchanging evidence and rationale.

The CVLs are all experienced and mature clinical testing laboratories committed to an evidence-based approach to variant classification, using the structures of the ACMG/AMP variant classification criteria. The process for harmonization of variant classification discrepancies is not open to bias as it follows the same steps regardless of the individual site. The 3 steps are: 1) accruing and sharing the fundamental data and evidence that contributes to the classification, 2) mapping of that evidence against the ACMG criteria using the same standards, regardless of the site, and 3) application of a final classification. Using this protocol, the evidence used to classify a variant is open and transparent. The logic then applied for the classification is standardized across the sites. Together this provides protection against bias. In the rare case of a disagreement that cannot be resolved through this evidence-based method, the committee will take variants to the ClinGen expert panels to clarify gene or disease specific specifications of the ACMG/AMP rules. Pre-program harmonization studies have demonstrated that this

standardized process provides an efficient and objective path to a correct and harmonized classification (Harrison et al. 2017; Harrison et al. 2018 PMID: 28301460, 30311378). If unresolved, the variants will be presented to the AoURP Clinical Interpretation and Reporting Working Group for resolution. If there remains disagreement, the relevant ClinGen Expert Panel will be consulted to classify the variant through expert consensus.

Participants in whom no P/LP variants are found in AoUHDR genes will receive their results via the participant portal that no significant hereditary disease risk results were identified (Appendix 1, Uninformative Report). All reports are reviewed and signed out by a board-certified laboratory geneticist or molecular pathologist and uses carefully designed language to address the uninformative nature of the testing (rather than a negative result), the limitations of testing, and directs participants to follow-up with their medical provider and/or the GCR if they have questions or concerns about their health or results. Unlimited access to the GCR will be provided free of charge.

For participants in whom a P/LP variant is identified in AoUHDR genes, a replicate DNA sample will be sent from the AoURP Biobank to the original AoURP CVL for genotype confirmation using an orthogonal assay (see Section 3.4.6 in the IDE). For HDR+ variant validation, the replicate DNA sample will be sent to the CVL site which ran the original genome annotation, i.e., at Color, U of Washington, or Baylor College of Medicine. AoURP DNA samples for whole genome sequencing data generation are randomly distributed among the three Centers, with one exception. Only the Baylor laboratory has NY State certification so all DNA samples from participants residing in NY at the time of enrollment are shipped to the Baylor Genome Center. If the presence of the P/LP variant is confirmed by the CVL in the participant's sample, the CVL result report will be signed by a board-certified laboratory geneticist or molecular pathologist and made available to the AoURP GCR. Participants will be notified to schedule a genetic counseling appointment for phone or video consultation with the GCR to disclose the result. A genetic counselor who is licensed in the jurisdiction where the participant resides will be assigned the case. S/he will review information available in the participant's report to contextualize the finding(s). Both a written and electronic version of the report will be made available to the participant at the time of disclosure (Appendices 2, 3, 4, HDR Positive Reports). AoUHDR P/LP results will be released through consultation with the GCR.

In addition to disclosing the finding(s) and offering a first level of explanation and contextualization, the GCR can also help connect participants to local resources. Please note that the GCR staff do not offer medical advice. The GCR will monitor and report to the AoURP on the barriers and facilitators of responsible return of results to support future procedural improvements.

For the duration of the AoURP, the GCR will be available to provide consultation at no cost to all participants, their care providers, and their families, when calling with a participant to answer questions.

Updating results over time: As the evidence base for genetic variation is constantly evolving, variant classifications will get updated over time. As such, the CVLs will re-issue AoUHDR reports whenever a reported variant changes classification. For all significant re-classifications, a genetic counselor will deliver the new result to the participant before the report is made available to the participant. A significant re-classification is defined as one that changes the actionability of the result (P/LP vs. Variants of Uncertain Significance (VUS) or Likely Benign (LB) or Benign (B)). Of special concern is reported pathogenic or likely pathogenic variants that are subsequently reclassified to VUS, likely benign or

benign. It is possible to estimate the rate of such cases, based on the expected frequency of P or LP results, and the expected likelihood that those variants will be reclassified downward. Based on estimated population prevalence, the expected rate for identification of P/LP variants in any the AoUHDR genes in any individual is approximately 2%. While variant classification is not yet a truly quantitative process, ACMG recommends minimum "confidence" thresholds of 90% for LP and 99% for P. Empirically observed reclassification down-grade rates are 2-4% for LP and 0.1-0.6% for P. Taken together, this suggests an expected false positive range of 1:2,000 to 1:10,000.

Note that the informed consent includes an alert to participants that results may change over time. From gRoR module: *Over time, we may learn more about DNA changes. We may learn new information that changes your results. As we learn more about DNA changes, we may go back and look at your DNA again. We will tell you if we find anything new. We will tell you if we find anything that changes your results. The list of what we will check for may change as researchers make new discoveries. You can find the most updated list of what we check for in your All of Us account.*

The DRC will be alerted when an AoUHDR report has been re-issued and will ensure that the latest reports are provided to participants and the GCR is notified to provide support as necessary. A significant re-classification is defined as one that changes the actionability of the result (P/LP vs. VUS or LB or B). Of special concern is reported pathogenic or likely pathogenic variants that are subsequently reclassified to VUS, likely benign or benign. It is possible to estimate the rate of such cases, based on the expected frequency of P or LP results, and the expected likelihood that those variants will be reclassified downward. Based on estimated population prevalence, the expected rate for identification of P/LP variants in any the AoUHDR genes in any individual is approximately 2%. While variant classification is not yet a truly quantitative process, ACMG recommends minimum "confidence" thresholds of 90% for LP and 99% for P. Empirically observed reclassification down-grade rates are 2-4% for LP and 0.1-0.6% for P. Taken together, this suggests an expected false positive range of 1:2,000 to 1:10,000.

6.1.4.2 AOUPGX RESULTS

The AoUPGx report constitutes an analysis of a subset of PGx genes (*CYP2C19, DPYD, G6PD, NUDT15, TPMT, SLCO1B1, and UGT1A1*) that have an associated CPIC guideline with at least one gene-drug association at CPIC level A criteria. These genes and specific alleles to be interrogated were selected for analysis and return based upon the availability of control samples and their performance in validation studies (Appendix 6). Not all genes that are included in CPIC guidelines will be included due to performance characteristics; for example, human leukocyte antigen (HLA) haplotypes and genes known to have copy number variants (*e.g.*, *CYP2D6*) require additional assays or pipelines that may be validated later and submitted as a supplement to this protocol. The phenotypic interpretation of reported allele diplotypes are determined using CPIC translation tables as described in Appendix 6.

Medications are listed on the AoUPGx report if there is evidence of an actionable phenotype-drug association. To be included, a gene-phenotype-drug combination must appear in FDA-approved drug product labeling (minimally in the Boxed Warning, Dosage and Administration, or Indications section), appear in the FDA Table of Pharmacogenetic Associations (Table of Pharmacogenetic associations for which the data support therapeutic management recommendations specifically) or have a moderate/strong recommendation for alternative medication or dosing modification within a CPIC guideline. For G6PD associations, all FDA-approved drug product labeling sections and CPIC guideline supplement table "Drug and compound safety reviews for G6PD deficient patients" were considered.

The full list of medications is shown in Figure 3 and the detailed evidence table is provided in Appendix 7.

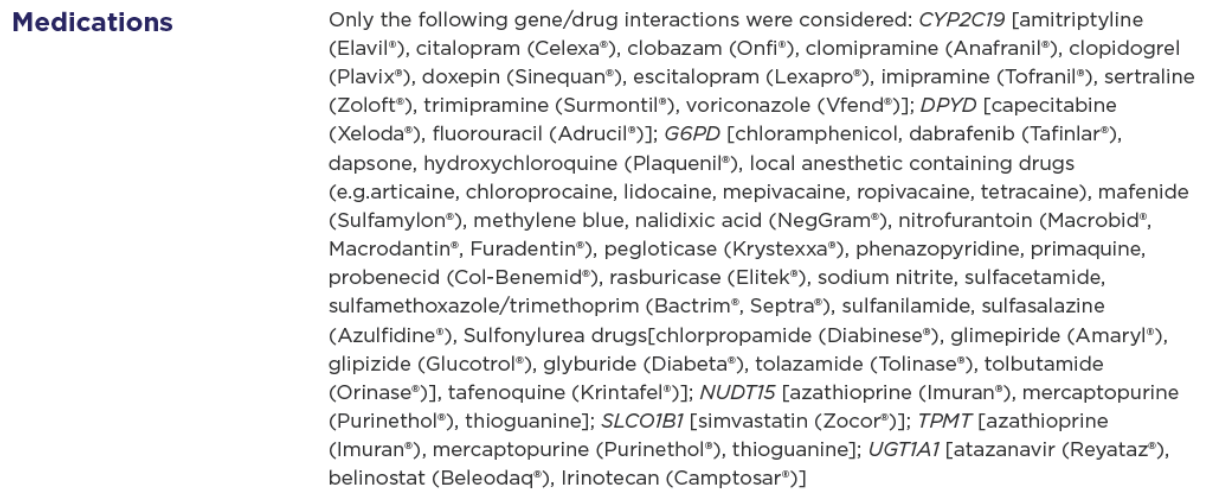


Figure 3: Full list of genes and potentially affected drugs. The full list of medications potentially affected by PGx alleles tested in this device is included in the Methods and Limitations section of the report provided to participants (Appendix 5; excerpted above).

The AoUPGx report, signed off by the laboratory or medical director, will be made available to participants through the Participant Portal (see example in Appendix 5, Medicine and Your DNA Report). Participants will be notified through push notifications, email, and/or similar direct notification pathways when reports are available. Reports include notice to the participant that results may change over time as new evidence becomes available. Standard AoURP practices for participant notifications will be utilized to alert participants of revised reports. AoUPGx reports will not be delivered through genetic counselors at the GCR, although participants, their families (with the participant on the line), and providers (as directed by the participant), will be able to access the GCR for free to address any questions they have about the AoUPGx report and speak with a genetic counselor if desired.

6.1.5 GENETIC COUNSELING

Availability of trained genetic counselors is an important risk mitigation factor for the study. Genetic counselors are available to participants throughout the protocol including to answer questions that might arise during the informed consent process. A contact phone number and email address are prominently displayed in participant notifications. Staff in the GCR are instructed to emphasize to all AoURP participants that the AoURP does not offer medical care and is providing research results that will need to be confirmed by a doctor prior to making any changes in medical care.

6.1.6 SUPPLEMENTAL EDUCATION FOR PARTICIPANTS AND PROVIDERS

The AoURP is working with multiple partners to deploy a variety of engaging, multimedia genomics-related educational materials for AoURP research participants. The AoURP is committed to providing participants with a knowledge base for understanding core genetic terms and concepts, scaling from rudimentary to multifaceted, through interactive engagement tools that spark a desire for knowledge. These educational resources augment reference materials such as frequently asked questions and fact sheets that are currently available as public resources. These materials will be approved by the IRB

before making them available to participants and providers in an interactive web-based learning environment.

6.2 PREPARATION/HANDLING/STORAGE/ACCOUNTABILITY

Not applicable; no exportable product is incorporated in the study.

6.3 MEASURES TO MINIMIZE BIAS: RANDOMIZATION AND BLINDING

Not applicable. There is no intervention provided in this study.

6.4 STUDY INTERVENTION COMPLIANCE

The Sponsor, i.e., NIH AoURP staff, will monitor activities throughout all elements of the study. Sponsor interactions with investigators include multiple web-enabled meetings each week, quarterly progress reports, and annual reporting requirements. The IRB is provided summary information for review on a quarterly schedule. In addition, operational metrics, such as sample failure or system downtime, are collected by the Sponsor on a continual basis through the AoURP DRC.

6.5 CONCOMITANT THERAPY

Not applicable. There is no intervention provided in this study.

7 STUDY INTERVENTION DISCONTINUATION AND PARTICIPANT DISCONTINUATION/ WITHDRAWAL

7.1 DISCONTINUATION OF STUDY INTERVENTION

Not applicable, there is no intervention provided in this study.

7.2 PARTICIPANT DISCONTINUATION/WITHDRAWAL FROM THE STUDY

The opportunity for AoURP participants to receive their genetic results is entirely optional. This is precisely the reason that this AoURP gRoR protocol requires a separate informed consent from the main research consent.

7.3 LOST TO FOLLOW-UP

Retrieval of an uninformative HDR report or PGx report is a passive process and at the complete discretion of the participant. A finding of an P/LP HDR result triggers a notification to the participant to schedule an appointment with a genetic counselor at the GCR. First, an email will be sent to the participant, which will contain a web-link to an online scheduling tool. If an appointment is not scheduled within seven days, a second email will be sent. If an appointment is still not scheduled after seven days, the GCR support staff will make three attempts to reach the participant by phone, leaving a HIPAA-compliant voicemail for each unsuccessful attempt. Finally, a letter will be sent to the participant's home address with instructions to call the GCR support line to schedule an appointment. After 30 additional days of inaction, a status update of "no response to GCR outreach" will be returned to the DRC.

8 STUDY ASSESSMENTS AND PROCEDURES

8.1 EFFICACY ASSESSMENTS

Not applicable; Outcomes will not be assessed in this study.

8.2 SAFETY AND OTHER ASSESSMENTS

Not applicable; Outcomes will not be assessed in this study.

8.3 ADVERSE DEVICE EFFECTS

8.3.1 DEFINITION OF ADVERSE DEVICE EFFECTS

An adverse device effect is any untoward or unfavorable medical occurrence in a human subject, including any abnormal sign (e.g., abnormal physical exam or laboratory finding), symptom, or disease, temporally associated with the subject's participation in the research, regardless of whether it is considered related to the subject's participation in the research. Adverse effects encompass both physical and psychological harms (Unanticipated Problems Involving Risks & Adverse Events Guidance, Office for Human Research Protections [2007]).

For the purposes of the *All of Us* gRoR protocol, we anticipate adverse device effects to be rare. At present, AoURP does not have plans to collect information from participants after results are returned.

8.3.2 TIME PERIOD AND FREQUENCY FOR ADVERSE EFFECTS ASSESSMENT AND FOLLOW-UP

Not applicable; Outcomes will not be assessed in this study.

8.3.3 ADVERSE EFFECTS REPORTING

AoURP Principal Investigators (PIs) are responsible for maintaining records and for reporting, as appropriate, to other research review committees all incidents, experiences, and outcomes, and their resolution, per the IRB-approved protocol, even if a report is not required to the AoURP IRB as a potential unanticipated problem. The AoURP IRB does not review individual adverse event or serious adverse event reports unless they meet the definition of unanticipated problems. Reporting of unanticipated problems is described in the relevant Section 8.4.

8.3.4 SERIOUS ADVERSE EVENT REPORTING

In compliance with 21 CFR 812.150(b)(1), AouRP will report any serious unanticipated adverse events to the FDA within 10 working days after the program first receives notice. However, at present, AoURP does not have plans to collect information from participants after results are returned.

8.3.5 REPORTING EVENTS TO PARTICIPANTS

AoURP does not anticipate any reportable events outside of potential events related to privacy and security. AoURP has an Incident Notification Board (INB) to oversee program responses to data security incidents and risks to participant privacy resulting from such incidents. In the case of an incident, including a breach, the INB, IRB, and the Program will determine whether participants need to be notified, and if so, the program will work with program partners, as necessary, to notify participants according to their preferred method of contact. For incidents for which notification is recommended, participants will be notified within 30 calendar days of incident discovery. For individuals whose access to notification cannot be confirmed electronically within 5 days, the program will send a letter to an individual's home address if available.

8.3.6 EVENTS OF SPECIAL INTEREST

Not applicable.

8.3.7 REPORTING OF PREGNANCY

Not applicable. Women of child-bearing potential are not excluded from participating in this study.

8.4 UNANTICIPATED PROBLEMS

8.4.1 DEFINITION OF UNANTICIPATED PROBLEMS (UP)

Unanticipated problems, in this case an unanticipated adverse device effect, means any serious adverse effect on health or safety or any life-threatening problem or death caused by, or associated with, a device, if that effect, problem, or death was not previously identified in nature, severity, or degree of incidence in the investigational plan or application (including a supplementary plan or application), or any other unanticipated serious problem associated with a device that relates to the rights, safety, or welfare of subjects (21 CFR 812.3(s)).

AoURP also abides by the Office for Human Research Protection's (OHRP) definition of an unanticipated problem. OHRP considers unanticipated problems, in general, to include any incident, experience, or outcome that meets all of the following criteria:

1. unexpected (in terms of nature, severity, or frequency) given (a) the research procedures that are described in the protocol-related documents, such as the IRB-approved research protocol and informed consent document; and (b) the characteristics of the subject population being studied;
2. related or possibly related to participation in the research (in this guidance document, possibly related means there is a reasonable possibility that the incident, experience, or outcome may have been caused by the procedures involved in the research); and
3. suggests that the research places subjects or others at a greater risk of harm (including physical, psychological, economic, or social harm) than was previously known or recognized.

8.4.2 UNANTICIPATED PROBLEM REPORTING

In compliance with 21 CFR 812.150(b)(1), AoURP will report any unanticipated adverse device effects to the FDA within 10 working days after the program first receives notice. However, AoURP does not currently have plans collect information from participants after results are returned.

AoURP will also report any unanticipated adverse device effects to the IRB, and other required reporting will be done in compliance with 45 CRF 46.103(b)(5). AoURP's specific procedures include preparing a letter that outlines the nature of the event, the findings of the program and the IRB, actions taken by the program or IRB, the reason for the actions, and plans for continued investigation or action. The letter will then be sent to program officials, OHRP, the FDA, and relevant institutional officials within 15 days of the completion of an investigation or determination.

8.4.3 REPORTING UNANTICIPATED PROBLEMS TO PARTICIPANTS

Depending on the severity of the observed unanticipated problem, the AoURP will make a determination if reporting to participants is warranted. The IRB will be informed of any serious adverse event and consulted on notification of the participants.

9 STATISTICAL CONSIDERATIONS

Not applicable. Statistical analysis is not being performed.

9.1 STATISTICAL HYPOTHESES

Not applicable.

9.2 SAMPLE SIZE DETERMINATION

Not applicable.

9.3 POPULATIONS FOR ANALYSES

Not applicable.

9.4 STATISTICAL ANALYSES

Not applicable.

10 SUPPORTING DOCUMENTATION AND OPERATIONAL CONSIDERATIONS

10.1 REGULATORY, ETHICAL, AND STUDY OVERSIGHT CONSIDERATIONS

10.1.1 INFORMED CONSENT PROCESS

10.1.1.1 CONSENT/ASSENT AND OTHER INFORMATIONAL DOCUMENTS PROVIDED TO PARTICIPANTS

AoURP takes a modular approach to informed consent, asking participants to complete a research consent to participate in the core facets of the research program and to subsequently complete additional consent modules for other program components. To participate in the gRoR protocol, all participants must first enroll in the program by completing the research consent (Appendix 12). They are then given the option to complete the "Consent to Get DNA Results" (referred to here as the gRoR consent; Appendix 13). Only those participants who answer affirmatively to the gRoR consent are eligible to receive genetic results. The modular approach to consent provides participants with a flexible experience to meet their individual needs and to decide if and when they would like to participate in different program components.

While the necessary information for informed decision-making is provided in the gRoR consent, the program will employ informing loops to remind participants of salient risk and benefit information, as well as topical educational content, for participants who consent to the return of their genetic results. Participants will be able to opt-in to each type of result return (HDR and PGx) through these informing loops, allowing for greater granular control to meet their individual needs.

Together, these informed consent forms encompass all general risks and benefits of the gRoR protocol. All informed consent forms comply with the requirements specified in FDA regulation 21 CFR Part 50 and the pre-2018 Common Rule (45 CFR Part 46). AoURP's participants sign each consent module with an electronic signature, and the program is compliant with 21 CFR Part 11 as described in Appendix 14.

10.1.1.2 CONSENT PROCEDURES AND DOCUMENTATION

Participants complete an interactive e-consent process that includes a series of screens with text, animated videos, and formative assessments (ungraded quizzes to facilitate understanding) before moving on to review and electronically sign a longform consent document. The purpose of the e-consent is to present key information from the longform document in an engaging and easy-to-understand way. The consent process may be self-navigated or completed with support from trained AoURP staff.

Participants can ask questions anytime as they proceed through the consent modules via in-person support (if present at an AoURP site) or chat and call prompts throughout the electronic materials. All of the program's informed consent longform documents are drafted at or below a 5th grade reading level, facilitating comprehension by ~80% of the US population.

The participant portal records the date and time that participants sign each informed consent module. Documentation of signed consents is maintained at the DRC and is also available through the participant portal. Participants can download and print a copy of their signed consent at any time.

The AoURP is committed to helping potential participants make informed decisions about whether to participate. By giving potential participants information about how the Program operates, reasonable expectations, and participants' rights, the AoURP ensures that people who decide to join do so because it's right for them. The research consent form, the gRoR consent form, and all accompanying materials, including animated videos, are provided in Appendices 10 -13 with video modules located in the submitted MISC FILES.

10.1.2 STUDY DISCONTINUATION AND CLOSURE

The AoURP is a longitudinal study expected to last for decades. Currently, the AoURP does not have procedures for discontinuation or closure but would develop such a protocol in the unlikely event that the NIH and/or US Congress discontinues support for the Program.

10.1.3 CONFIDENTIALITY AND PRIVACY

Maintaining data security and privacy within the *All of Us* Research Program is paramount to maintaining participants' trust and engagement. Extensive regulations, policies, governance, compliance, and technical safeguards are implemented to ensure that participant data security and privacy are appropriately protected. Details of the AoURP security policy can be found at *https://allofus.nih.gov/protecting-data-and-privacy/precision-medicine-initiative-data-security-policy-principles-and-framework-overview*.

10.1.4 FUTURE USE OF STORED SPECIMENS AND DATA

This is not applicable to this protocol. Only study participants will have access to their reports.

10.1.5 KEY ROLES AND STUDY GOVERNANCE

Leadership

The AoURP is a large collaborative initiative sponsored by the NIH. The research Program functions as a consortium of awardees from multiple institutions. Its governance involves representation from each awardee and participant representatives. The consortium also includes the Program Chief Executive Officer (CEO) and project scientists/specialists from NIH. Each awardee has responsibilities commensurate with expertise.

Dr. Paul Harris of Vanderbilt University Medical Center serves as the *AoU* IRB Protocol Principal Investigator on behalf of the AoURP Consortium.

Governance

The Steering Committee (SC) is the primary governing body of AoURP. The SC recommends strategic directions for the Program and oversees planning, coordination, and implementation of the Program's overall operations. Its 50 voting members include PIs from each awardee as designated in the notice of award; representation from NIH, comprising of the Chief Operating Officer (COO) and other Chief

Officers of AoURP; representation from community partners and participants; and additional representation as needed to ensure balanced representation of stakeholders.

The SC may approve the formation of additional governance bodies – committees, task forces, boards, etc. – as necessary to fulfill the mission of AoURP. The purpose of these additional governance bodies is to alleviate the bandwidth constraints of the SC by gathering subject-matter experts from within the consortium to oversee discussion and develop policies, recommendations, or guidelines related to their assigned topic.

10.1.6 SAFETY OVERSIGHT

The Sponsor will provide safety oversight.

10.1.7 CLINICAL MONITORING

Not applicable.

10.1.8 QUALITY ASSURANCE AND QUALITY CONTROL

Quality assurance approaches in the operations of the GCs are described throughout Section 2.2.2 Test Performance Characteristics of the IDE application. The Biobank for AoURP is operated by Mayo Clinic. The Mayo Clinic operates a CAP-accredited biobank that receives, stores, and processes participant biospecimens. employs industry best practices for sample collection, pre-processing, and shipment, as well as CAP-certified quality control procedures. All participating institutions and investigators contributing to this protocol must make records available for inspection at any time per the terms of the NIH awards.

10.1.9 DATA HANDLING AND RECORD KEEPING

Digital records from all AoURP operational activities are retained at the DRC. Data security incidents and human subjects adverse events, and description of their resolution, are reported to dedicated email boxes in the NIH AoURP offices.

10.1.9.1 DATA COLLECTION AND MANAGEMENT RESPONSIBILITIES

The AoURP DRC holds responsibilities to secure, host, and manage all transactions throughout the operational workflows of the Program. Importantly, the DRC secures and manages participant and biospecimen IDs. In this role, the DRC serves as the gatekeeper for informed consent or withdrawal status. At all decision steps in the protocol that require verification of informed consent status of the participant before initiation, e.g., prior to clinical annotation of genomic data, the DRC confirms affirmative consent / non-withdrawal status of participants.

10.1.9.2 STUDY RECORDS RETENTION

In accordance with 21 CFR 812.140(b), the Sponsor will maintain the following accurate, complete, and current records relating to this protocol:

a. Correspondence (including reports) with investigators, AoU IRB and FDA

b. Signed investigator agreements

e. Adverse device effects (whether anticipated or unanticipated) and complaints

10.1.10 PROTOCOL DEVIATIONS

Protocol deviations are not allowed without prior approval of the Sponsor and regulatory bodies. Penalties are specified in the conditions of federal awards to investigators escalating to termination of the award if warranted.

10.1.11 PUBLICATION AND DATA SHARING POLICY

As a large NIH-funded consortium of investigators, the AoURP has established publication policies and procedures to ensure appropriate authorship of publications. More importantly, the AoURP, as noted in the NIH Director's Precision Medicine Initiative working group report (Appendix 15, ACD Working Group Report), has a core principle to share data to the research community as broadly as possible, while protecting the privacy of Program participants. More information is located at the AoURP Research Hub located at *www.researchallofus.org*.

10.1.12 CONFLICT OF INTEREST POLICY

The independence of this study from any actual or perceived influence is critical. Therefore, any actual conflict of interest of persons who have a role in the design, conduct, analysis, publication, or any aspect of this trial will be disclosed and managed. Furthermore, persons who have a perceived conflict of interest will be required to have such conflicts managed in a way that is appropriate to their participation in the design and conduct of this study. In conjunction with the NIH Office of the Director, the AoURP leadership has established policies and procedures for all study group members to disclose all conflicts of interest and has established a mechanism for the management of all reported dualities of interest as enforced by the terms and condition of the NIH federal award to each investigator's institution.

10.2 ADDITIONAL CONSIDERATIONS

None.

10.3 ABBREVIATIONS

Abbreviation	Definition
ACMG	American College of Medical Genetics and Genomics
AoUHDR (AoU Hereditary Disease Risk)	The list of genes underlying serious monogenic conditions, based on the American College of Medical Genetics and Genomics' policy statement
AoUPGx	The list of pharmacogenetic genes to be interpreted and returned as part of this study. The list is based on a subset of pharmacogenomic variants selected using guidance of the US Food and Drug Administration (FDA) and the Clinical Pharmacogenomics Implementation Consortium (CPIC)
AoURP	*All of Us* Research Program
B	Benign
CAP	College of American Pathologists
CEO	Chief Executive Officer
CLIA	Clinical Laboratory Improvement Amendments
COO	Chief Operating Officer
CPIC	Clinical Pharmacogenomics Implementation Consortium
CPIC-A	The list of medications for which, "the preponderance of evidence is high or moderate in favor of changing prescribing" based on genomic test results per the Clinical Pharmacogenomics Implementation Consortium, https://www.pharmgkb.org/page/clinAnnLevels
CVL	Clinical Validation Laboratories
DNA	Deoxyribonucleic Acid
DRC	Data and Research Center
EC	Ethics Committee
EC	Executive Committee
EDTA	Ethylenediaminetetraacetic Acid
FDA	US Food and Drug Administration
FQHC	Federally Qualified Health Center
GCR	Genetic Counseling Resource
gRoR	Return of Genomic Results
HDRR	Hereditary Disease Risk Report
HIPAA	Health Insurance Portability and Accountability Act
HLA	Human Leukocyte Antigen
HPO	Healthcare Provider Organization
IATA	International Air Transport Association
IDE	Investigational Device Exemption
INB	Incident Notification Board
IRB	Institutional Review Board
LB	Likely Benign
LMM	Library for Molecular Medicine
NCBI	National Center for Biotechnology Information
NIH	National Institutes of Health
NIST	National Institute of Standards and Technology

Abbreviation	Definition
OHRP	Office of Human Research Protections
P/LP	Pathogenic/Likely Pathogenic genetic variants
PGx	Pharmacogenomics/Pharmacogenetics
PII	Personally Identifiable Information
PMI-DSPP	Precision Medicine Initiative Data Security Policy Principles and Framework
RHP	Reporting and Harmonization Platform
RLIMS	Research Laboratory Information Management System
RMC	Regional Medical Center
SBS	Sequencing-By-Synthesis
SC	Steering Committee
VA	Veteran's Administration
VCF	Variant Call File
VUS	Variant-Uncertain Significance
WGS	Whole Genome Sequencing

10.4 PROTOCOL AMENDMENT HISTORY

Version	Date	Description of Change	Brief Rationale
1.0		Original protocol	

11 REFERENCES

1. ACMG Board of Directors. ACMG policy statement: updated recommendations regarding analysis and reporting of secondary findings in clinical genome-scale sequencing. *Genet Med*. 2015;17(1):68-69. doi:10.1038/gim.2014.151.
2. Angrist M. You never call, you never write: why return of 'omic' results to research participants is both a good idea and a moral imperative. *Per Med*. 2011;8(6):651-657. doi:10.2217/pme.11.62.
3. Christensen KD, Phillips KA, Green RC, Dukhovny D. Cost Analyses of Genomic Sequencing: Lessons Learned from the MedSeq Project. Value Health 2018;21:1054-61.
4. Christensen K, Vassy J, Phillips K, et al. Short-term costs of integrating whole-genome sequencing into primary care and cardiology settings: a pilot randomized trial. Genet Med 2018;Epub ahead of print.
5. Green RC, Berg JS, Grody WW, Kalia SS, Korf BR, Martin CL, et al. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. Genet Med. 2013;15:565–574. doi: 10.1038/gim.2013.73.
6. Joint Task Force Transformation Initiative, ‘‘Risk management framework for information systems and organizations: A system life cycle approach for security and privacy,’’ Nat. Inst. Standards Technol., Gaithersburg, MD, USA, Tech. Rep. NIST SP 800-37r2, Dec. 2018. [Online]. Available: *https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800- 37r2.pdf*.
7. 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 [published correction appears in Genet Med. 2017 Apr;19(4):484]. *Genet Med*. 2017;19(2):249-255. doi:10.1038/gim.2016.190.
8. Perkins BA, Caskey CT, Brar P, et al. Precision medicine screening using whole-genome sequencing and advanced imaging to identify disease risk in adults. Proc Natl Acad Sci U S A 2018;115:3686-91.
9. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015;17:405-24.
10. Roberts J, Robinson J, Diamond P, et al. Patient understanding of, satisfaction with, and perceived utility of whole-genome sequencing: findings from the MedSeq Project. Genet Med 2018;[Epub ahead of print].
11. Terry SF, Terry PF, Power to the People: Participant Ownership of Clinical Trial Data. SCI. TRANSL. MED. **3**, 69cm3 (2011).
12. Shalowitz DI, Miller FG. Disclosing individual results of clinical research: implications of respect for participants. *JAMA*. 2005;294(6):737-740. doi:10.1001/jama.294.6.737.
13. Vassy JL, Christensen KD, Schonman EF, et al. The impact of whole-genome sequencing on the primary care and outcomes of health adult patients: a pilot randomized trial. Ann Intern Medicine 2017;Epub ahead of print.
14. Vassy JL, Lautenbach DM, McLaughlin HM, et al. The MedSeq Project: a randomized trial of integrating whole genome sequencing into clinical medicine. Trials 2014;15:85-97.

15. Zoltick ES, Linderman MD, McGinniss MA, et al. Predispositional genome sequencing in healthy adults: design, participant characteristics, and early outcomes of the PeopleSeq Consortium. Genome Med 2019;11:10.

All of Us RESEARCH PROGRAM

JANE DOE
DOB: May 25, 1977
ID: 123456

Specimen: Blood
Barcode: 223 234234 2343
Collected: September 15, 2018
Report date: October 2, 2018

RESEARCH RESULT - Your doctor will need to confirm this result with a clinical test before using it in your care.

Your Result:
We did not find anything significant for your health in the genes we looked at.

What does this mean?

Some people have changes in their genes that increase their risk of developing certain diseases. **We did not find any of these kinds of changes in your genes.**

That's good, right?

That's generally good news. But we can't say too much more. Here's why:

- **A lot of genes can impact your health or cause disease, and we did not look at all of them.**
- *All of Us* is a research program. The way we check DNA might not be the same as a doctor-ordered test. There could even be something we couldn't see or can't understand in the genes that we did look at.

IMPORTANT!

- This report comes from a research program so **it is a research result**. Your doctor will need to confirm these results with a clinical genetics test before using them in your care.
- **Do not change your medical care** based on this result.
- **Results provided are from an investigational device**. An "investigational device" is a device that is the subject of a clinical study.

Your Result (continued)

What does this mean for my health and what should I do next?

This information should not change anything about how you think about your health.

Keep taking care of yourself. Eat well. Get enough sleep. Exercise when you can. If you smoke, think about quitting. See a doctor regularly. Tell your doctor about your family history. We know that these things work to help keep people healthy.

Understanding this report

This test looked at 59 genes in your DNA that can be related to serious diseases like cancer and heart disease.

How did *All of Us* look at my DNA?

You gave a **blood** sample to the *All of Us* Research Program. We processed that blood to get some of your **DNA**. An *All of Us* genetics lab gave a readout of that DNA.

Because you said "Yes" to getting health-related DNA results, a specially trained scientist looked closely at some of the genes in your DNA. We wrote this report for you, based on what they found.

What if I have more questions?

You can talk to an *All of Us* genetic counselor by calling XXX-XXX-XXXX. They can answer questions about your result or help you find a local genetic counselor to talk to.

Common Questions

What was done to get this result?

Actually, quite a lot! DNA is in your blood and other samples. You gave a sample to *All of Us*. We processed your sample to extract the DNA. We sent some of your DNA to a special lab. The lab gave a readout of your DNA. A specially trained scientist checked some of the genes in your DNA for disease-causing changes. We wrote this report based on what they found.

What does this mean for my family?

Your DNA is a lot like your family member's DNA, but everyone is different. This result does not say anything about their health or their own DNA.

What was the point of looking at my DNA?

Everyone has the same set of genes, but different people can have slightly different *versions* of those genes.

Some people have a version of a gene that increases their chance of developing a serious disease. In some cases, knowing that can be life-saving. Because the information can be so important, it's worth looking at a lot of people's DNA to find these rare people.

We couldn't know what your result would be before checking your DNA.

Could my result change?

Yes. *All of Us* could look at more genes or look again at these genes as science improves. Check your *All of Us* account to make sure this is the most up-to-date version of this report.

RESEARCH RESULT - Your doctor will need to confirm this result with a clinical test before using it in your care.

We Looked at These Genes

This is the list of genes we looked at, and some of the diseases they can be related to.

There are limitations to the analysis we did. There are a lot of genes that can cause disease, and we didn't look at all of them. There could even be something we couldn't see or can't understand in the genes that we did look at.

All of Us might look at other genes in the future or look again at these genes as science advances.

Gene	Technical name and number of a disease associated with this gene
ACTA2	Aortic aneurysm, familial thoracic 6 (MIM 611788)
ACTC1	Familial hypertrophic cardiomyopathy 11 (MIM 612098)
APC	Adenomatous polyposis coli (MIM 175100)
APOB	Familial hypercholesterolemia (MIM 143890)
ATP7B	Wilson disease (MIM 277900)
BMPR1A	Juvenile polyposis syndrome, (MIM 174900)
BRCA1	Breast-ovarian cancer, familial 1 (MIM 604370)
BRCA2	Breast-ovarian cancer, familial 2 (MIM 612555)
CACNA1S	Malignant hyperthermia (MIM 145600)
COL3A1	Ehlers-Danlos syndrome, type 4 (MIM 130050)
DSC2	Arrhythmogenic right ventricular cardiomyopathy, type 11 (MIM 610476)
DSG2	Arrhythmogenic right ventricular cardiomyopathy, type 10 (MIM 610193)
DSP	Arrhythmogenic right ventricular cardiomyopathy, type 8 (MIM 607450)
FBN1	Marfan syndrome (MIM 154700)
GLA	Fabry disease (MIM 301500)
KCNH2	Long QT syndrome 2 (MIM 613688)
KCNQ1	Long QT syndrome 1 (MIM 192500)
LDLR	Familial hypercholesterolemia (MIM 143890)
LMNA	Dilated cardiomyopathy 1A (MIM 115200)

We looked at these genes (continued)

Gene	Technical name and number of a disease associated with this gene
MEN1	Multiple endocrine neoplasia, type 1 (MIM 131100)
MLH1	Lynch syndrome (MIM 120435)
MSH2	Lynch syndrome (MIM 120435)
MSH6	Lynch syndrome (MIM 120435)
MUTYH	MYH-associated polyposis (MIM 608456)
MYBPC3	Dilated cardiomyopathy 1A (MIM 115200) & Familial hypertrophic cardiomyopathy 4 (MIM 115197)
MYH11	Aortic aneurysm, familial thoracic 4 (MIM 132900)
MYH7	Familial hypertrophic cardiomyopathy 1 (MIM 192600)
MYL2	Familial hypertrophic cardiomyopathy 10 (MIM 608758)
MYL3	Familial hypertrophic cardiomyopathy 8 (MIM 608751)
NF2	Neurofibromatosis, type 2 (MIM 101000)
OTC	Ornithine carbamoyltransferase deficiency (MIM 311250)
PCSK9	Hypercholesterolemia, autosomal dominant, 3 (MIM 603776)
PKP2	Arrhythmogenic right ventricular cardiomyopathy, type 9 (MIM 609040)
PMS2	Lynch syndrome (MIM 120435)
PRKAG2	Familial hypertrophic cardiomyopathy 6 (MIM 600858)
PTEN	PTEN hamartoma tumor syndrome (MIM 153480)
RB1	Retinoblastoma (MIM 180200)
RET	Familial medullary thyroid carcinoma (MIM 155240)
RYR1	Malignant hyperthermia (MIM 145600)
RYR2	Catecholaminergic polymorphic ventricular tachycardia (MIM 604772)
SCN5A	Long QT syndrome 3 (MIM 603830)
SDHAF2	Paragangliomas 2 (MIM 601650)
SDHB	Paragangliomas 4 (MIM 115310)
SDHC	Paragangliomas 3 (MIM 605373)
SDHD	Paragangliomas 1 (MIM 168000)

We looked at these genes (continued)

Gene	Technical name and number of a disease associated with this gene
SMAD3	Loeys-Dietz syndrome type 3 (MIM 613795)
SMAD4	Juvenile polyposis syndrome, (MIM 174900)
STK11	Peutz-Jeghers syndrome (MIM 175200)
TGFBR1	Marfan syndrome (MIM 154700)
TGFBR2	Loeys-Dietz syndrome type 1B (MIM 610168)
TMEM43	Arrhythmogenic right ventricular cardiomyopathy, type 5 (MIM 604400)
TNNI3	Familial hypertrophic cardiomyopathy 7 (MIM 613690)
TNNT2	Left ventricular noncompaction 6 (MIM 601494)
TP53	Li-Fraumeni syndrome 1 (MIM 151623)
TPM1	Familial hypertrophic cardiomyopathy 3 (MIM 115196)
TSC1	Tuberous sclerosis 1 (MIM 191100)
TSC2	Tuberous sclerosis 2 (MIM 613254)
VHL	Von Hippel-Lindau syndrome (MIM 193300)
WT1	Wilms tumor (MIM 194070)

Methods and Limitations

Methods

This section has some technical information about the test that was performed.

This report represents the analysis of a sample submitted as a part of the *All of Us* Research Program. The sample was collected at <COLLECTION_SITE>. The sample was stored and the DNA was extracted at <BIOBANK_SITE>. Genetic data was generated at <GENOMECENTER_SITE> and interpreted at <CVL_SITE>.

Genomic DNA was extracted from the submitted sample and sequenced using Illumina Next Generation Sequencing. Sequence data was aligned to a reference genome, and variants were identified using a suite of bioinformatic tools designed to detect single nucleotide variants and small insertions/deletions.

This test was developed and its performance characteristics determined by the *All of Us* Research Program, with clinical laboratories accredited by the College of American Pathologists (CAP) and certified under the Clinical Laboratory Improvement Amendments (CLIA) to perform high-complexity testing.

Genes & Transcripts

ACTA2 (NM_001613), ACTC1 (NM_005159), APC (NM_000038), APOB (NM_000384), ATP7B (NM_000053), BMPR1A (NM_004329), BRCA1 (NM_007294), BRCA2 (NM_000059), CACNA1S (NM_000069), COL3A1 (NM_000090), DSC2 (NM_024422), DSG2 (NM_001943), DSP (NM_004415), FBN1 (NM_000138), GLA (NM_000169), KCNH2 (NM_000238), KCNQ1 (NM_000218), LDLR (NM_000527), LMNA (NM_005572; NM_170707), MEN1 (NM_130799), MLH1 (NM_000249), MSH2 (NM_000251), MSH6 (NM_000179), MUTYH (NM_001128425), MYBPC3 (NM_000256), MYH11 (NM_001040113), MYH7 (NM_000257), MYL2 (NM_000432), MYL3 (NM_000258), NF2 (NM_000268), OTC (NM_000531), PCSK9 (NM_174936), PKP2 (NM_004572), PMS2 (NM_000535), PRKAG2 (NM_016203), PTEN (NM_000314), RB1 (NM_000321), RET (NM_020975), RYR1 (NM_000540), RYR2 (NM_001035), SCN5A (NM_198056), SDHAF2 (NM_017841), SDHB (NM_003000), SDHC (NM_003001), SDHD (NM_003002), SMAD3 (NM_005902), SMAD4 (NM_005359), STK11 (NM_000455), TGFBR1 (NM_004612), TGFBR2 (NM_003242), TMEM43 (NM_024334), TNNI3 (NM_000363), TNNT2 (NM_001001430), TP53 (NM_000546), TPM1 (NM_001018005), TSC1 (NM_000368), TSC2 (NM_000548), VHL (NM_000551), WT1 (NM_000378)

Limitations

- **Results provided are from an investigational device.**
- **Because this report is based on data derived from a research study, this information cannot be used to diagnose, cure, mitigate, treat, or prevent disease.**
- **The interpretation of these results could be incorrect.**
- This test may not detect all variants in the analyzed genes. The *All of Us* Research Program only reports findings within the genes that are on the panel; variants in other genes are not reported. Larger chromosomal events will also not be reported.
- In very rare cases, such as allogeneic bone marrow transplant, or recent blood transfusion (within 7 days of providing the sample), the results of this analysis may reflect the DNA of the donor. DNA quality may be affected if a participant has received chemotherapy within 120 days of providing the sample. In addition, certain organ transplants or diseases (liver, kidney, heart) may limit the relevance of the results.

All of Us RESEARCH PROGRAM

JANE DOE
DOB: May 25, 1977
ID: 123456

Specimen: Blood
Barcode: 223 234234 2343
Collected: September 15, 2018
Report date: October 2, 2018

RESEARCH RESULT - Your doctor will need to confirm this result with a clinical test before using it in your care.

Your Result: Something very important for your health was found in your *BRCA1* gene.

What does this mean?

- This result means that you are more likely to get certain types of cancer than other people.
- It does **not** mean that you have certain types of cancer.
- It does **not** mean that you will definitely get certain types of cancer.
- **This result is important** and should not be ignored.

IMPORTANT!

- **Share this report with your doctor.**
- This report comes from a research program so **it is a research result**. Your doctor will need to confirm these results with a clinical genetics test before using them in your care.
- **Do not change your medical care** before this result is confirmed by your doctor.
- **Results provided are from an investigational device**. An "investigational device" is a device that is the subject of a clinical study.

The *BRCA1* gene

Women and men who have this result in the *BRCA1* gene have a higher chance to develop certain cancers in their lifetime compared to someone without this result. Women are at higher risk for breast cancer and ovarian cancer. They may also have a higher risk of pancreatic cancer. Men are at higher risk for male breast cancer and pancreatic cancer. They may also have a higher risk of prostate cancer.

Your Result (continued)

What should I do next?

Share this result with your doctor. They can confirm this result using a clinical test.

Your doctor may send you to a specialist. They will ask you about your family's health history. They may make a plan with you to reduce your risk of disease.

If you have questions right now, you can **talk to an *All of Us* genetic counselor** by calling XXX-XXX-XXXX.

Should I share this with my family?

Yes! You should explain to your family that this is a research result that has not been confirmed, and that this information is not actionable without clinical confirmation. You are welcome to invite your family to join you on a call with an All of Us Genetic Counselor if they have questions.

It can be valuable to do this because you share the same DNA with many of your relatives, which means that your relatives could have the same result. Your parents, children, and brothers and sisters would each have a 50/50 chance of having this same result. Cousins, aunts, uncles and grandparents could have it too. Men and women have the same chance of having this result and have the same chance of passing it on to their children.

Sharing your result with your family can help them think about if they want to get tested themselves. **This could help them prevent disease or detect it early.**

Some people feel nervous talking about health issues with their family. That's normal. But it is important to share your results.

Additional information about your Hereditary Disease Risk Report

Understanding this report

This test looked at 59 genes in your DNA that can be related to serious diseases like cancer and heart disease.

This report has information that could be **very important** for you.

How did *All of Us* look at my DNA?

You gave a **blood** sample to the *All of Us* Research Program. We processed that blood to get some of your **DNA**. An *All of Us* genetics lab gave a readout of that DNA.

Because you said "Yes" to getting health-related DNA results, a specially trained scientist looked closely at some of the genes in your DNA. We wrote this report for you, based on what they found.

Common Questions

Who can I talk to if I have questions about this?

You can talk to an *All of Us* genetic counselor for free by calling XXX-XXX-XXXX. You can also talk to your doctor. And if your doctor has questions, they can call us too.

What did you actually find?

Everyone has the same set of genes, but different people can have slightly different *versions* of those genes.

We looked closely to see which versions of the genes you have, and we found that you have a version of a gene that can increase your chance of developing a disease.

The technical term for what we found is a "pathogenic DNA variant." It is described in detail on the page titled "Technical Report."

How did I get this version of this gene?

The analysis we did doesn't tell us how you got this version of this gene. Most of our DNA features are inherited from one of our parents. Rarely, people have "new" DNA changes that were not inherited from either parent.

Parents do not choose which parts of their DNA they pass to their kids. It's random. The significance of this result for your health does not depend on where it came from.

What is the "Technical Report" at the end of this document?

The Technical Report has the same information you've already read. It uses more technical language and includes details that might be useful in ordering clinical testing. Share it with your doctor.

Common Questions (continued)

Why does this result need to be confirmed?

All of Us is a research program. Everything possible has been done to make sure that this information is correct, but to be absolutely sure, the test should be repeated using a new sample taken in your doctor's office.

Could my result change?

Yes. *All of Us* could look at more genes or look again at these genes or DNA changes as science improves. Check your *All of Us* account to make sure this is the most up-to-date version of this report.

RESEARCH RESULT - Your doctor will need to confirm this result with a clinical test before using it in your care.

We Looked at These Genes

This is the list of genes we looked at, and some of the diseases they can be related to. Except for the result described above, we did not find anything else significant for your health in the genes we looked at.

There are limitations to the analysis we did. There are a lot of genes that can cause disease, and we didn't look at all of them. There could even be something we couldn't see or can't understand in the genes that we did look at.

All of Us might look at other genes in the future, or look again at these genes as science advances.

Gene	Technical name and number of a disease associated with this gene
ACTA2	Aortic aneurysm, familial thoracic 6 (MIM 611788)
ACTC1	Familial hypertrophic cardiomyopathy 11 (MIM 612098)
APC	Adenomatous polyposis coli (MIM 175100)
APOB	Familial hypercholesterolemia (MIM 143890)
ATP7B	Wilson disease (MIM 277900)
BMPR1A	Juvenile polyposis syndrome, (MIM 174900)
BRCA1	Breast-ovarian cancer, familial 1 (MIM 604370)
BRCA2	Breast-ovarian cancer, familial 2 (MIM 612555)
CACNA1S	Malignant hyperthermia (MIM 145600)
COL3A1	Ehlers-Danlos syndrome, type 4 (MIM 130050)
DSC2	Arrhythmogenic right ventricular cardiomyopathy, type 11 (MIM 610476)
DSG2	Arrhythmogenic right ventricular cardiomyopathy, type 10 (MIM 610193)
DSP	Arrhythmogenic right ventricular cardiomyopathy, type 8 (MIM 607450)
FBN1	Marfan syndrome (MIM 154700)
GLA	Fabry disease (MIM 301500)
KCNH2	Long QT syndrome 2 (MIM 613688)
KCNQ1	Long QT syndrome 1 (MIM 192500)
LDLR	Familial hypercholesterolemia (MIM 143890)

We Looked at These Genes (continued)

Gene	Technical name and number of a disease associated with this gene
LMNA	Dilated cardiomyopathy 1A (MIM 115200)
MEN1	Multiple endocrine neoplasia, type 1 (MIM 131100)
MLH1	Lynch syndrome (MIM 120435)
MSH2	Lynch syndrome (MIM 120435)
MSH6	Lynch syndrome (MIM 120435)
MUTYH	MYH-associated polyposis (MIM 608456)
MYBPC3	Dilated cardiomyopathy 1A (MIM 115200) & Familial hypertrophic cardiomyopathy 4 (MIM 115197)
MYH11	Aortic aneurysm, familial thoracic 4 (MIM 132900)
MYH7	Familial hypertrophic cardiomyopathy 1 (MIM 192600)
MYL2	Familial hypertrophic cardiomyopathy 10 (MIM 608758)
MYL3	Familial hypertrophic cardiomyopathy 8 (MIM 608751)
NF2	Neurofibromatosis, type 2 (MIM 101000)
OTC	Ornithine carbamoyltransferase deficiency (MIM 311250)
PCSK9	Hypercholesterolemia, autosomal dominant, 3 (MIM 603776)
PKP2	Arrhythmogenic right ventricular cardiomyopathy, type 9 (MIM 609040)
PMS2	Lynch syndrome (MIM 120435)
PRKAG2	Familial hypertrophic cardiomyopathy 6 (MIM 600858)
PTEN	PTEN hamartoma tumor syndrome (MIM 153480)
RB1	Retinoblastoma (MIM 180200)
RET	Familial medullary thyroid carcinoma (MIM 155240)
RYR1	Malignant hyperthermia (MIM 145600)
RYR2	Catecholaminergic polymorphic ventricular tachycardia (MIM 604772)
SCN5A	Long QT syndrome 3 (MIM 603830)
SDHAF2	Paragangliomas 2 (MIM 601650)
SDHB	Paragangliomas 4 (MIM 115310)
SDHC	Paragangliomas 3 (MIM 605373)

We Looked at These Genes (continued)

Gene	Technical name and number of a disease associated with this gene
SDHD	Paragangliomas 1 (MIM 168000)
SMAD3	Loeys-Dietz syndrome type 3 (MIM 613795)
SMAD4	Juvenile polyposis syndrome, (MIM 174900)
STK11	Peutz-Jeghers syndrome (MIM 175200)
TGFBR1	Marfan syndrome (MIM 154700)
TGFBR2	Loeys-Dietz syndrome type 1B (MIM 610168)
TMEM43	Arrhythmogenic right ventricular cardiomyopathy, type 5 (MIM 604400)
TNNI3	Familial hypertrophic cardiomyopathy 7 (MIM 613690)
TNNT2	Left ventricular noncompaction 6 (MIM 601494)
TP53	Li-Fraumeni syndrome 1 (MIM 151623)
TPM1	Familial hypertrophic cardiomyopathy 3 (MIM 115196)
TSC1	Tuberous sclerosis 1 (MIM 191100)
TSC2	Tuberous sclerosis 2 (MIM 613254)
VHL	Von Hippel-Lindau syndrome (MIM 193300)
WT1	Wilms tumor (MIM 194070)

Technical Report

IMPORTANT!

This is a research result. This result was generated from a sample submitted as part of the *All of Us* Research Program. This result should be repeated on a second sample as a part of a clinical test before any care decisions are made.

Details

Gene	Variant	Classification
BRCA1	c.2035A>T(p.Lys679*) *Gene transcript:* NM_007294.3 *Genomic coordinates:* chr17.GRCh37:g.41245513T>A *Variant zygosity:* Heterozygous	Pathogenic

Supporting evidence

This variant changes 1 nucleotide in exon 10 of the *BRCA1* gene, creating a premature translation stop signal. This variant is expected to result in an absent or non-functional protein product. This variant has not been identified in the general population by the Genome Aggregation Database (gnomAD). Loss of BRCA1 function is a known mechanism of disease (clinicalgenome.org). Based on the available evidence, this variant is classified as Pathogenic.

About this result

The presence of a heterozygous variant in the *BRCA1* gene has been associated with breast-ovarian cancer, familial 1. This hereditary disorder is associated with an increased lifetime risk of breast, ovarian, pancreatic, and prostate cancers.

Clinical confirmation of this result and follow up with a doctor are recommended.

Questions?

Healthcare providers can call the *All of Us* Genetic Counseling Resource with questions about these results. Speak to an *All of Us* genetic counselor for free by calling XXX-XXX-XXXX.

Technical Report (continued)

Test performed

DNA was extracted, and genetic data was generated using next-generation sequencing of the entire genome. Observed variants in the 59 genes listed below were clinically interpreteted following ACMG variant interpretation guidelines.

Reviewed by

Sarah Genetics, PhD, FACMG Date

Methods and Limitations

Methods

This section has some technical information about the test that was performed.

This report represents the analysis of a sample submitted as a part of the *All of Us* Research Program. The sample was collected at <COLLECTION_SITE>. The sample was stored and the DNA was extracted at <BIOBANK_SITE>. Genetic data was generated at <GENOMECENTER_SITE> and interpreted at <CVL_SITE>.

Genomic DNA was extracted from the submitted sample and sequenced using Illumina Next Generation Sequencing. Sequence data was aligned to a reference genome, and variants were identified using a suite of bioinformatic tools designed to detect single nucleotide variants and small insertions/deletions.

This test was developed and its performance characteristics determined by the *All of Us* Research Program, with clinical laboratories accredited by the College of American Pathologists (CAP) and certified under the Clinical Laboratory Improvement Amendments (CLIA) to perform high-complexity testing.

Genes & Transcripts

ACTA2 (NM_001613), ACTC1 (NM_005159), APC (NM_000038), APOB (NM_000384), ATP7B (NM_000053), BMPR1A (NM_004329), BRCA1 (NM_007294), BRCA2 (NM_000059), CACNA1S (NM_000069), COL3A1 (NM_000090), DSC2 (NM_024422), DSG2 (NM_001943), DSP (NM_004415), FBN1 (NM_000138), GLA (NM_000169), KCNH2 (NM_000238), KCNQ1 (NM_000218), LDLR (NM_000527), LMNA (NM_005572; NM_170707), MEN1 (NM_130799), MLH1 (NM_000249), MSH2 (NM_000251), MSH6 (NM_000179), MUTYH (NM_001128425), MYBPC3 (NM_000256), MYH11 (NM_001040113), MYH7 (NM_000257), MYL2 (NM_000432), MYL3 (NM_000258), NF2 (NM_000268), OTC (NM_000531), PCSK9 (NM_174936), PKP2 (NM_004572), PMS2 (NM_000535), PRKAG2 (NM_016203), PTEN (NM_000314), RB1 (NM_000321), RET (NM_020975), RYR1 (NM_000540), RYR2 (NM_001035), SCN5A (NM_198056), SDHAF2 (NM_017841), SDHB (NM_003000), SDHC (NM_003001), SDHD (NM_003002), SMAD3 (NM_005902), SMAD4 (NM_005359), STK11 (NM_000455), TGFBR1 (NM_004612), TGFBR2 (NM_003242), TMEM43 (NM_024334), TNNI3 (NM_000363), TNNT2 (NM_001001430), TP53 (NM_000546), TPM1 (NM_001018005), TSC1 (NM_000368), TSC2 (NM_000548), VHL (NM_000551), WT1 (NM_000378)

Limitations

- **Results provided are from an investigational device.**
- **Because this report is based on data derived from a research study, this information cannot be used to diagnose, cure, mitigate, treat, or prevent disease.**
- **The interpretation of these results could be incorrect.**
- This test may not detect all variants in the analyzed genes. The *All of Us* Research Program only reports findings within the genes that are on the panel; variants in other genes are not reported. Larger chromosomal events will also not be reported.
- In very rare cases, such as allogeneic bone marrow transplant, or recent blood transfusion (within 7 days of providing the sample), the results of this analysis may reflect the DNA of the donor. DNA quality may be affected if a participant has received chemotherapy within 120 days of providing the sample. In addition, certain organ transplants or diseases (liver, kidney, heart) may limit the relevance of the results.

All of Us RESEARCH PROGRAM

JANE DOE
DOB: May 25, 1977
ID: 123456

Specimen: Blood
Barcode: 223 234234 2343
Collected: September 15, 2018
Report date: October 2, 2018

RESEARCH RESULT - Your doctor will need to confirm this result with a clinical test before using it in your care.

Your Result: Something very important for your health was found in your *MSH2* gene.

What does this mean?

- This result means that you are more likely to get certain types of cancer than other people.
- It does **not** mean that you have certain types of cancer.
- It does **not** mean that you will definitely get certain types of cancer.
- **This result is important** and should not be ignored.

IMPORTANT!

- **Share this report with your doctor.**
- This report comes from a research program so **it is a research result**. Your doctor will need to confirm these results with a clinical genetics test before using them in your care.
- **Do not change your medical care** before this result is confirmed by your doctor.
- **Results provided are from an investigational device**. An "investigational device" is a device that is the subject of a clinical study.

The *MSH2* gene

Women and men who have this result in the *MSH2* gene have a higher chance to develop certain cancers in their lifetime compared to someone without this result. Women are at higher risk for colorectal and uterine cancers. They may also have a higher risk of other cancers, like ovarian and stomach cancer. Men are at higher risk for colorectal cancer. They may also have a higher risk of other cancers, like stomach, small bowel, and pancreatic cancer.

Your Result (continued)

What should I do next?

Share this result with your doctor. They can confirm this result using a clinical test.

Your doctor may send you to a specialist. They will ask you about your family's health history. They may make a plan with you to reduce your risk of disease.

If you have questions right now, you can **talk to an *All of Us* genetic counselor** by calling XXX-XXX-XXXX.

Should I share this with my family?

Yes! You should explain to your family that this is a research result that has not been confirmed, and that this information is not actionable without clinical confirmation. You are welcome to invite your family to join you on a call with an All of Us Genetic Counselor if they have questions.

It can be valuable to do this because you share the same DNA with many of your relatives, which means that your relatives could have the same result. Your parents, children, and brothers and sisters would each have a 50/50 chance of having this same result. Cousins, aunts, uncles and grandparents could have it too. Men and women have the same chance of having this result and have the same chance of passing it on to their children.

Sharing your result with your family can help them think about if they want to get tested themselves. **This could help them prevent disease or detect it early.**

Some people feel nervous talking about health issues with their family. That's normal. But it is important to share your results.

Additional information about your Hereditary Disease Risk Report

Understanding this report

This test looked at 59 genes in your DNA that can be related to serious diseases like cancer and heart disease.

This report has information that could be **very important** for you.

How did *All of Us* look at my DNA?

You gave a **blood** sample to the *All of Us* Research Program. We processed that blood to get some of your **DNA**. An *All of Us* genetics lab gave a readout of that DNA.

Because you said "Yes" to getting health-related DNA results, a specially trained scientist looked closely at some of the genes in your DNA. We wrote this report for you, based on what they found.

Common Questions

Who can I talk to if I have questions about this?

You can talk to an *All of Us* genetic counselor for free by calling XXX-XXX-XXXX. You can also talk to your doctor. And if your doctor has questions, they can call us too.

What did you actually find?

Everyone has the same set of genes, but different people can have slightly different *versions* of those genes.

We looked closely to see which versions of the genes you have, and we found that you have a version of a gene that can increase your chance of developing a disease.

The technical term for what we found is a "pathogenic DNA variant." It is described in detail on the page titled "Technical Report."

How did I get this version of this gene?

The analysis we did doesn't tell us how you got this version of this gene. Most of our DNA features are inherited from one of our parents. Rarely, people have "new" DNA changes that were not inherited from either parent.

Parents do not choose which parts of their DNA they pass to their kids. It's random. The significance of this result for your health does not depend on where it came from.

What is the "Technical Report" at the end of this document?

The Technical Report has the same information you've already read. It uses more technical language and includes details that might be useful in ordering clinical testing. Share it with your doctor.

Common Questions (continued)

Why does this result need to be confirmed?

All of Us is a research program. Everything possible has been done to make sure that this information is correct, but to be absolutely sure, the test should be repeated using a new sample taken in your doctor's office.

Could my result change?

Yes. *All of Us* could look at more genes or look again at these genes or DNA changes as science improves. Check your *All of Us* account to make sure this is the most up-to-date version of this report.

RESEARCH RESULT - Your doctor will need to confirm this result with a clinical test before using it in your care.

We Looked at These Genes

This is the list of genes we looked at, and some of the diseases they can be related to. Except for the result described above, we did not find anything else significant for your health in the genes we looked at.

There are limitations to the analysis we did. There are a lot of genes that can cause disease, and we didn't look at all of them. There could even be something we couldn't see or can't understand in the genes that we did look at.

All of Us might look at other genes in the future, or look again at these genes as science advances.

Gene	Technical name and number of a disease associated with this gene
ACTA2	Aortic aneurysm, familial thoracic 6 (MIM 611788)
ACTC1	Familial hypertrophic cardiomyopathy 11 (MIM 612098)
APC	Adenomatous polyposis coli (MIM 175100)
APOB	Familial hypercholesterolemia (MIM 143890)
ATP7B	Wilson disease (MIM 277900)
BMPR1A	Juvenile polyposis syndrome, (MIM 174900)
BRCA1	Breast-ovarian cancer, familial 1 (MIM 604370)
BRCA2	Breast-ovarian cancer, familial 2 (MIM 612555)
CACNA1S	Malignant hyperthermia (MIM 145600)
COL3A1	Ehlers-Danlos syndrome, type 4 (MIM 130050)
DSC2	Arrhythmogenic right ventricular cardiomyopathy, type 11 (MIM 610476)
DSG2	Arrhythmogenic right ventricular cardiomyopathy, type 10 (MIM 610193)
DSP	Arrhythmogenic right ventricular cardiomyopathy, type 8 (MIM 607450)
FBN1	Marfan syndrome (MIM 154700)
GLA	Fabry disease (MIM 301500)
KCNH2	Long QT syndrome 2 (MIM 613688)
KCNQ1	Long QT syndrome 1 (MIM 192500)
LDLR	Familial hypercholesterolemia (MIM 143890)

We Looked at These Genes (continued)

Gene	Technical name and number of a disease associated with this gene
LMNA	Dilated cardiomyopathy 1A (MIM 115200)
MEN1	Multiple endocrine neoplasia, type 1 (MIM 131100)
MLH1	Lynch syndrome (MIM 120435)
MSH2	Lynch syndrome (MIM 120435)
MSH6	Lynch syndrome (MIM 120435)
MUTYH	MYH-associated polyposis (MIM 608456)
MYBPC3	Dilated cardiomyopathy 1A (MIM 115200) & Familial hypertrophic cardiomyopathy 4 (MIM 115197)
MYH11	Aortic aneurysm, familial thoracic 4 (MIM 132900)
MYH7	Familial hypertrophic cardiomyopathy 1 (MIM 192600)
MYL2	Familial hypertrophic cardiomyopathy 10 (MIM 608758)
MYL3	Familial hypertrophic cardiomyopathy 8 (MIM 608751)
NF2	Neurofibromatosis, type 2 (MIM 101000)
OTC	Ornithine carbamoyltransferase deficiency (MIM 311250)
PCSK9	Hypercholesterolemia, autosomal dominant, 3 (MIM 603776)
PKP2	Arrhythmogenic right ventricular cardiomyopathy, type 9 (MIM 609040)
PMS2	Lynch syndrome (MIM 120435)
PRKAG2	Familial hypertrophic cardiomyopathy 6 (MIM 600858)
PTEN	PTEN hamartoma tumor syndrome (MIM 153480)
RB1	Retinoblastoma (MIM 180200)
RET	Familial medullary thyroid carcinoma (MIM 155240)
RYR1	Malignant hyperthermia (MIM 145600)
RYR2	Catecholaminergic polymorphic ventricular tachycardia (MIM 604772)
SCN5A	Long QT syndrome 3 (MIM 603830)
SDHAF2	Paragangliomas 2 (MIM 601650)
SDHB	Paragangliomas 4 (MIM 115310)
SDHC	Paragangliomas 3 (MIM 605373)

We Looked at These Genes (continued)

Gene	Technical name and number of a disease associated with this gene
SDHD	Paragangliomas 1 (MIM 168000)
SMAD3	Loeys-Dietz syndrome type 3 (MIM 613795)
SMAD4	Juvenile polyposis syndrome, (MIM 174900)
STK11	Peutz-Jeghers syndrome (MIM 175200)
TGFBR1	Marfan syndrome (MIM 154700)
TGFBR2	Loeys-Dietz syndrome type 1B (MIM 610168)
TMEM43	Arrhythmogenic right ventricular cardiomyopathy, type 5 (MIM 604400)
TNNI3	Familial hypertrophic cardiomyopathy 7 (MIM 613690)
TNNT2	Left ventricular noncompaction 6 (MIM 601494)
TP53	Li-Fraumeni syndrome 1 (MIM 151623)
TPM1	Familial hypertrophic cardiomyopathy 3 (MIM 115196)
TSC1	Tuberous sclerosis 1 (MIM 191100)
TSC2	Tuberous sclerosis 2 (MIM 613254)
VHL	Von Hippel-Lindau syndrome (MIM 193300)
WT1	Wilms tumor (MIM 194070)

Technical Report

IMPORTANT!

This is a research result. This result was generated from a sample submitted as part of the *All of Us* Research Program. This result should be repeated on a second sample as a part of a clinical test before any care decisions are made.

Details

Gene	Variant	Classification
MSH2	c.942+3A>T *Gene transcript:* NM_000251.2 *Genomic coordinates:* chr2.GRCh37:g.47641560A>T *Variant zygosity:* Heterozygous	Pathogenic

Supporting evidence

This variant causes an A>T nucleotide substitution at the +3 position of intron 5 of the *MSH2* gene. Functional RNA studies have shown this variant caused skipping of exon 5 (r.793_942del) and the in-frame deletion of 50 amino acids in the DNA binding domain (PMID: 8062247, 16395668, 19267393). This variant has been reported as a recurrent de novo mutation in individuals affected with Lynch syndrome-associated cancer in different ethnicities (PMID: 10978353) and this variant also been reported in multiple suspected Lynch syndrome cases (PMID: 8062247, 10446963, 12112654, 12362047, 15222003, 16395668, 18625694, 19419416, 20682701). This variant has been reported to segregate with Lynch syndrome cancer in families with likelihood ratio of 27.66:1 (PMID: 19267393). This variant has been identified in 1/30582 chromosomes in the general population by the Genome Aggregation Database (gnomAD). Loss of MSH2 function is a known mechanism of disease. Based on the available evidence, this variant is classified as Pathogenic.

Technical Report (continued)

About this result

The presence of a heterozygous variant in the *MSH2* gene has been associated with Lynch syndrome. This hereditary disorder is associated with a higher lifetime risk of colorectal, endometrial, ovarian, stomach, small bowel, and other cancers.

Clinical confirmation of this result and follow up with a doctor are recommended.

Questions?

Healthcare providers can call the *All of Us* Genetic Counseling Resource with questions about these results. Speak to an *All of Us* genetic counselor for free by calling XXX-XXX-XXXX.

Test performed

DNA was extracted, and genetic data was generated using next-generation sequencing of the entire genome. Observed variants in the 59 genes listed below were clinically interpreteted following ACMG variant interpretation guidelines.

Reviewed by

Sarah Genetics, PhD, FACMG Date

Methods and Limitations

Methods

This section has some technical information about the test that was performed.

This report represents the analysis of a sample submitted as a part of the *All of Us* Research Program. The sample was collected at <COLLECTION_SITE>. The sample was stored and the DNA was extracted at <BIOBANK_SITE>. Genetic data was generated at <GENOMECENTER_SITE> and interpreted at <CVL_SITE>.

Genomic DNA was extracted from the submitted sample and sequenced using Illumina Next Generation Sequencing. Sequence data was aligned to a reference genome, and variants were identified using a suite of bioinformatic tools designed to detect single nucleotide variants and small insertions/deletions.

This test was developed and its performance characteristics determined by the *All of Us* Research Program, with clinical laboratories accredited by the College of American Pathologists (CAP) and certified under the Clinical Laboratory Improvement Amendments (CLIA) to perform high-complexity testing.

Genes & Transcripts

ACTA2 (NM_001613), ACTC1 (NM_005159), APC (NM_000038), APOB (NM_000384), ATP7B (NM_000053), BMPR1A (NM_004329), BRCA1 (NM_007294), BRCA2 (NM_000059), CACNA1S (NM_000069), COL3A1 (NM_000090), DSC2 (NM_024422), DSG2 (NM_001943), DSP (NM_004415), FBN1 (NM_000138), GLA (NM_000169), KCNH2 (NM_000238), KCNQ1 (NM_000218), LDLR (NM_000527), LMNA (NM_005572; NM_170707), MEN1 (NM_130799), MLH1 (NM_000249), MSH2 (NM_000251), MSH6 (NM_000179), MUTYH (NM_001128425), MYBPC3 (NM_000256), MYH11 (NM_001040113), MYH7 (NM_000257), MYL2 (NM_000432), MYL3 (NM_000258), NF2 (NM_000268), OTC (NM_000531), PCSK9 (NM_174936), PKP2 (NM_004572), PMS2 (NM_000535), PRKAG2 (NM_016203), PTEN (NM_000314), RB1 (NM_000321), RET (NM_020975), RYR1 (NM_000540), RYR2 (NM_001035), SCN5A (NM_198056), SDHAF2 (NM_017841), SDHB (NM_003000), SDHC (NM_003001), SDHD (NM_003002), SMAD3 (NM_005902), SMAD4 (NM_005359), STK11 (NM_000455), TGFBR1 (NM_004612), TGFBR2 (NM_003242), TMEM43 (NM_024334), TNNI3 (NM_000363), TNNT2 (NM_001001430), TP53 (NM_000546), TPM1 (NM_001018005), TSC1 (NM_000368), TSC2 (NM_000548), VHL (NM_000551), WT1 (NM_000378)

Limitations

- **Results provided are from an investigational device.**
- **Because this report is based on data derived from a research study, this information cannot be used to diagnose, cure, mitigate, treat, or prevent disease.**
- **The interpretation of these results could be incorrect.**
- This test may not detect all variants in the analyzed genes. The *All of Us* Research Program only reports findings within the genes that are on the panel; variants in other genes are not reported. Larger chromosomal events will also not be reported.
- In very rare cases, such as allogeneic bone marrow transplant, or recent blood transfusion (within 7 days of providing the sample), the results of this analysis may reflect the DNA of the donor. DNA quality may be affected if a participant has received chemotherapy within 120 days of providing the sample. In addition, certain organ transplants or diseases (liver, kidney, heart) may limit the relevance of the results.

All of Us RESEARCH PROGRAM

JANE DOE
DOB: May 25, 1977
ID: 123456

Specimen: Blood
Barcode: 223 234234 2343
Collected: September 15, 2018
Report date: October 2, 2018

RESEARCH RESULT - Your doctor will need to confirm this result with a clinical test before using it in your care.

Your Result: Something very important for your health was found in your *LDLR* gene.

What does this mean?

- This result means that you are more likely to get very high cholesterol than other people.
- It does **not** mean that you have very high cholesterol.
- It does **not** mean that you will definitely get very high cholesterol.
- **This result is important** and should not be ignored.

IMPORTANT!

- **Share this report with your doctor.**
- This report comes from a research program so **it is a research result**. Your doctor will need to confirm these results with a clinical genetics test before using them in your care.
- **Do not change your medical care** before this result is confirmed by your doctor.
- **Results provided are from an investigational device**. An "investigational device" is a device that is the subject of a clinical study.

The *LDLR* gene

Women and men who have this result in the *LDLR* gene can have a large build up of bad cholesterol (LDL-C) in their blood vessels. This can lead to a heart attack or stroke.

Your Result (continued)

What should I do next?

Share this result with your doctor. They can confirm this result using a clinical test.

Your doctor may send you to a specialist. They will ask you about your family's health history. They may make a plan with you to reduce your risk of disease.

If you have questions right now, you can **talk to an *All of Us* genetic counselor** by calling XXX-XXX-XXXX.

Should I share this with my family?

Yes! You should explain to your family that this is a research result that has not been confirmed, and that this information is not actionable without clinical confirmation. You are welcome to invite your family to join you on a call with an All of Us Genetic Counselor if they have questions.

It can be valuable to do this because you share the same DNA with many of your relatives, which means that your relatives could have the same result. Your parents, children, and brothers and sisters would each have a 50/50 chance of having this same result. Cousins, aunts, uncles and grandparents could have it too. Men and women have the same chance of having this result and have the same chance of passing it on to their children.

Sharing your result with your family can help them think about if they want to get tested themselves. **This could help them prevent disease or detect it early.**

Some people feel nervous talking about health issues with their family. That's normal. But it is important to share your results.

Additional information about your Hereditary Disease Risk Report

Understanding this report

This test looked at 59 genes in your DNA that can be related to serious diseases like cancer and heart disease.

This report has information that could be **very important** for you.

How did *All of Us* look at my DNA?

You gave a **blood** sample to the *All of Us* Research Program. We processed that blood to get some of your **DNA**. An *All of Us* genetics lab gave a readout of that DNA.

Because you said "Yes" to getting health-related DNA results, a specially trained scientist looked closely at some of the genes in your DNA. We wrote this report for you, based on what they found.

Common Questions

Who can I talk to if I have questions about this?

You can talk to an *All of Us* genetic counselor for free by calling XXX-XXX-XXXX. You can also talk to your doctor. And if your doctor has questions, they can call us too.

What did you actually find?

Everyone has the same set of genes, but different people can have slightly different *versions* of those genes.

We looked closely to see which versions of the genes you have, and we found that you have a version of a gene that can increase your chance of developing a disease.

The technical term for what we found is a "pathogenic DNA variant." It is described in detail on the page titled "Technical Report."

How did I get this version of this gene?

The analysis we did doesn't tell us how you got this version of this gene. Most of our DNA features are inherited from one of our parents. Rarely, people have "new" DNA changes that were not inherited from either parent.

Parents do not choose which parts of their DNA they pass to their kids. It's random. The significance of this result for your health does not depend on where it came from.

What is the "Technical Report" at the end of this document?

The Technical Report has the same information you've already read. It uses more technical language and includes details that might be useful in ordering clinical testing. Share it with your doctor.

Common Questions (continued)

Why does this result need to be confirmed?

All of Us is a research program. Everything possible has been done to make sure that this information is correct, but to be absolutely sure, the test should be repeated using a new sample taken in your doctor's office.

Could my result change?

Yes. *All of Us* could look at more genes or look again at these genes or DNA changes as science improves. Check your *All of Us* account to make sure this is the most up-to-date version of this report.

RESEARCH RESULT - Your doctor will need to confirm this result with a clinical test before using it in your care.

We Looked at These Genes

This is the list of genes we looked at, and some of the diseases they can be related to. Except for the result described above, we did not find anything else significant for your health in the genes we looked at.

There are limitations to the analysis we did. There are a lot of genes that can cause disease, and we didn't look at all of them. There could even be something we couldn't see or can't understand in the genes that we did look at.

All of Us might look at other genes in the future, or look again at these genes as science advances.

Gene	Technical name and number of a disease associated with this gene
ACTA2	Aortic aneurysm, familial thoracic 6 (MIM 611788)
ACTC1	Familial hypertrophic cardiomyopathy 11 (MIM 612098)
APC	Adenomatous polyposis coli (MIM 175100)
APOB	Familial hypercholesterolemia (MIM 143890)
ATP7B	Wilson disease (MIM 277900)
BMPR1A	Juvenile polyposis syndrome, (MIM 174900)
BRCA1	Breast-ovarian cancer, familial 1 (MIM 604370)
BRCA2	Breast-ovarian cancer, familial 2 (MIM 612555)
CACNA1S	Malignant hyperthermia (MIM 145600)
COL3A1	Ehlers-Danlos syndrome, type 4 (MIM 130050)
DSC2	Arrhythmogenic right ventricular cardiomyopathy, type 11 (MIM 610476)
DSG2	Arrhythmogenic right ventricular cardiomyopathy, type 10 (MIM 610193)
DSP	Arrhythmogenic right ventricular cardiomyopathy, type 8 (MIM 607450)
FBN1	Marfan syndrome (MIM 154700)
GLA	Fabry disease (MIM 301500)
KCNH2	Long QT syndrome 2 (MIM 613688)
KCNQ1	Long QT syndrome 1 (MIM 192500)
LDLR	Familial hypercholesterolemia (MIM 143890)

We Looked at These Genes (continued)

Gene	Technical name and number of a disease associated with this gene
LMNA	Dilated cardiomyopathy 1A (MIM 115200)
MEN1	Multiple endocrine neoplasia, type 1 (MIM 131100)
MLH1	Lynch syndrome (MIM 120435)
MSH2	Lynch syndrome (MIM 120435)
MSH6	Lynch syndrome (MIM 120435)
MUTYH	MYH-associated polyposis (MIM 608456)
MYBPC3	Dilated cardiomyopathy 1A (MIM 115200) & Familial hypertrophic cardiomyopathy 4 (MIM 115197)
MYH11	Aortic aneurysm, familial thoracic 4 (MIM 132900)
MYH7	Familial hypertrophic cardiomyopathy 1 (MIM 192600)
MYL2	Familial hypertrophic cardiomyopathy 10 (MIM 608758)
MYL3	Familial hypertrophic cardiomyopathy 8 (MIM 608751)
NF2	Neurofibromatosis, type 2 (MIM 101000)
OTC	Ornithine carbamoyltransferase deficiency (MIM 311250)
PCSK9	Hypercholesterolemia, autosomal dominant, 3 (MIM 603776)
PKP2	Arrhythmogenic right ventricular cardiomyopathy, type 9 (MIM 609040)
PMS2	Lynch syndrome (MIM 120435)
PRKAG2	Familial hypertrophic cardiomyopathy 6 (MIM 600858)
PTEN	PTEN hamartoma tumor syndrome (MIM 153480)
RB1	Retinoblastoma (MIM 180200)
RET	Familial medullary thyroid carcinoma (MIM 155240)
RYR1	Malignant hyperthermia (MIM 145600)
RYR2	Catecholaminergic polymorphic ventricular tachycardia (MIM 604772)
SCN5A	Long QT syndrome 3 (MIM 603830)
SDHAF2	Paragangliomas 2 (MIM 601650)
SDHB	Paragangliomas 4 (MIM 115310)
SDHC	Paragangliomas 3 (MIM 605373)

We Looked at These Genes (continued)

Gene	Technical name and number of a disease associated with this gene
SDHD	Paragangliomas 1 (MIM 168000)
SMAD3	Loeys-Dietz syndrome type 3 (MIM 613795)
SMAD4	Juvenile polyposis syndrome, (MIM 174900)
STK11	Peutz-Jeghers syndrome (MIM 175200)
TGFBR1	Marfan syndrome (MIM 154700)
TGFBR2	Loeys-Dietz syndrome type 1B (MIM 610168)
TMEM43	Arrhythmogenic right ventricular cardiomyopathy, type 5 (MIM 604400)
TNNI3	Familial hypertrophic cardiomyopathy 7 (MIM 613690)
TNNT2	Left ventricular noncompaction 6 (MIM 601494)
TP53	Li-Fraumeni syndrome 1 (MIM 151623)
TPM1	Familial hypertrophic cardiomyopathy 3 (MIM 115196)
TSC1	Tuberous sclerosis 1 (MIM 191100)
TSC2	Tuberous sclerosis 2 (MIM 613254)
VHL	Von Hippel-Lindau syndrome (MIM 193300)
WT1	Wilms tumor (MIM 194070)

Technical Report

IMPORTANT!

This is a research result. This result was generated from a sample submitted as part of the *All of Us* Research Program. This result should be repeated on a second sample as a part of a clinical test before any care decisions are made.

Details

Gene	Variant	Classification
LDLR	c.530C>T (p. Ser177Leu) *Gene transcript:* NM_000527.4 *Genomic coordinates:* chr19.GRCh37:g.11216112C>T *Variant zygosity:* Heterozygous	Pathogenic

Supporting evidence

This missense variant (also known as p.Ser156Leu in the mature protein and as FH Puerto Rico) is located in the fourth *LDLR* type A repeat of the ligand binding domain of the LDLR protein. Computational prediction tools and conservation analyses suggest that this variant may have deleterious impact on the protein function. Computational splicing tools suggest that this variant may not impact RNA splicing. Experimental functional studies have shown that this variant significantly reduces the ability of the LDLR protein to bind LDL (PMID: 2760205, 25647241). This variant has been reported in over 30 individuals affected with familial hypercholesterolemia in a heterozygous, compound heterozygous or homozygous state (PMID: 15241806, 17765246, 18263977, 22698793, 25461735 , 25487149, 25647241, 2760205, 27816806, 28235710). This variant has shown a strong segregation with autosomal dominant hypercholesterolemia in a large family (PMID: 2760205). This variant has been identified in 4/246154 chromosomes in the general population by the Genome Aggregation Database (gnomAD). Based on available evidence, this variant is classified as Pathogenic.

Technical Report (continued)

About this result

The presence of a heterozygous variant in the *LDLR* gene has been associated with familial hypercholesterolemia. This is a hereditary disorder associated with elevated LDL cholesterol levels that lead to atherosclerotic plaque deposition in the coronary arteries and proximal aorta at an early age, leading to an increased risk for coronary artery disease and cardiovascular disease.

Clinical confirmation of this result and follow up with a doctor are recommended.

Questions?

Healthcare providers can call the *All of Us* Genetic Counseling Resource with questions about these results. Speak to an *All of Us* genetic counselor for free by calling XXX-XXX-XXXX.

Test performed

DNA was extracted, and genetic data was generated using next-generation sequencing of the entire genome. Observed variants in the 59 genes listed below were clinically interpreteted following ACMG variant interpretation guidelines.

Reviewed by

Sarah Genetics, PhD, FACMG Date

Methods and Limitations

Methods

This section has some technical information about the test that was performed.

This report represents the analysis of a sample submitted as a part of the *All of Us* Research Program. The sample was collected at <COLLECTION_SITE>. The sample was stored and the DNA was extracted at <BIOBANK_SITE>. Genetic data was generated at <GENOMECENTER_SITE> and interpreted at <CVL_SITE>.

Genomic DNA was extracted from the submitted sample and sequenced using Illumina Next Generation Sequencing. Sequence data was aligned to a reference genome, and variants were identified using a suite of bioinformatic tools designed to detect single nucleotide variants and small insertions/deletions.

This test was developed and its performance characteristics determined by the *All of Us* Research Program, with clinical laboratories accredited by the College of American Pathologists (CAP) and certified under the Clinical Laboratory Improvement Amendments (CLIA) to perform high-complexity testing.

Genes & Transcripts

ACTA2 (NM_001613), ACTC1 (NM_005159), APC (NM_000038), APOB (NM_000384), ATP7B (NM_000053), BMPR1A (NM_004329), BRCA1 (NM_007294), BRCA2 (NM_000059), CACNA1S (NM_000069), COL3A1 (NM_000090), DSC2 (NM_024422), DSG2 (NM_001943), DSP (NM_004415), FBN1 (NM_000138), GLA (NM_000169), KCNH2 (NM_000238), KCNQ1 (NM_000218), LDLR (NM_000527), LMNA (NM_005572; NM_170707), MEN1 (NM_130799), MLH1 (NM_000249), MSH2 (NM_000251), MSH6 (NM_000179), MUTYH (NM_001128425), MYBPC3 (NM_000256), MYH11 (NM_001040113), MYH7 (NM_000257), MYL2 (NM_000432), MYL3 (NM_000258), NF2 (NM_000268), OTC (NM_000531), PCSK9 (NM_174936), PKP2 (NM_004572), PMS2 (NM_000535), PRKAG2 (NM_016203), PTEN (NM_000314), RB1 (NM_000321), RET (NM_020975), RYR1 (NM_000540), RYR2 (NM_001035), SCN5A (NM_198056), SDHAF2 (NM_017841), SDHB (NM_003000), SDHC (NM_003001), SDHD (NM_003002), SMAD3 (NM_005902), SMAD4 (NM_005359), STK11 (NM_000455), TGFBR1 (NM_004612), TGFBR2 (NM_003242), TMEM43 (NM_024334), TNNI3 (NM_000363), TNNT2 (NM_001001430), TP53 (NM_000546), TPM1 (NM_001018005), TSC1 (NM_000368), TSC2 (NM_000548), VHL (NM_000551), WT1 (NM_000378)

Limitations

- **Results provided are from an investigational device.**
- **Because this report is based on data derived from a research study, this information cannot be used to diagnose, cure, mitigate, treat, or prevent disease.**
- **The interpretation of these results could be incorrect.**
- This test may not detect all variants in the analyzed genes. The *All of Us* Research Program only reports findings within the genes that are on the panel; variants in other genes are not reported. Larger chromosomal events will also not be reported.
- In very rare cases, such as allogeneic bone marrow transplant, or recent blood transfusion (within 7 days of providing the sample), the results of this analysis may reflect the DNA of the donor. DNA quality may be affected if a participant has received chemotherapy within 120 days of providing the sample. In addition, certain organ transplants or diseases (liver, kidney, heart) may limit the relevance of the results.

JANE DOE
DOB: May 25, 1977
ID: 123456

Specimen: Blood
Barcode: 223 234234 2343
Collected: September 15, 2018
Report date: October 2, 2018

RESEARCH RESULT - Do NOT use this result to make any changes to your medicines.

Medicine and your DNA

Our **genes** affect how we respond to **medicine**.

They do that in many different ways. Some genes help move medicines to the right part of the body. Some genes help break down medicines and clear them from your body. Some genes even change medicines into a form that makes them work properly.

This test looked at a few of the genes in your DNA that can affect how medicines are used. The technical term for this kind of information is "pharmacogenetics".

What is this kind of information used for?

Doctors and pharmacists use this kind of information when they consider why medicines work differently for different people.

But doctors and pharmacists don't make decisions based on just DNA. Some other important considerations can be age, weight, health, diet, and other medications you are taking at the same time.

IMPORTANT!

- **If your doctor has prescribed medicine for you, keep taking it**. It can be dangerous to stop taking a medicine, or to change the dose or timing of it, without first asking your doctor.
- This report comes from a research program so **it is a research result**. That means that neither you nor your doctor should use it to make any changes to your medicines. Your doctor would need a separate clinical test if they wanted to use the information.
- **Share this report with your doctor** so they can decide if they should order that clinical test for you.
- **Results provided are from an investigational device.** An "investigational device" is a device that is the subject of a clinical study.

Understanding Your Report

There are three parts to this report:

1. **Your Genetic Results** (page 3) shows the genes we checked and which versions of the genes we saw in your DNA.
2. **DNA and Medicine** (page 4) indicates some medicines that *may* be impacted by your genetics. Please remember: the only way to know for sure is by talking to a doctor or pharmacist.
3. **Next Steps** (page 6) talks about why you might share this report with your doctor.

We're going to repeat this a few times because it's so important:

- **If your doctor has prescribed medicine for you, keep taking it**. It can be dangerous to stop taking a medicine, or to change the dose or timing of it, without first asking your doctor.
- This report comes from a research program so **it is a research result**. That means that neither you nor your doctor should use it to make any changes to your medicines. Your doctor would need a separate clinical test if they wanted to use the information.
- **Share this report with your doctor** so they can decide if they should order that clinical test for you.
- **Results provided are from an investigational device.** An "investigational device" is a device that is the subject of a clinical study.

Your Genetic Results

This table shows three things:

1. The **genes** we checked. Gene names are usually a string of letters and numbers. They are often pronounced just by spelling them out.
2. The **versions** of the genes you have. Everyone has the same genes, but some people can have different *versions* that can work slightly differently. For genes that affect medicine, the versions of the genes are named things like *1, *2 or *3. Sometimes they're named after the place in the world where they were first observed.
3. **What it means.** These terms describe how quickly or slowly your versions of these genes will do their work, or 'metabolize'.

Gene	Version	What it Means
CYP2C19	*2/*2	Poor metabolizer
DPYD	*1/*1	Normal metabolizer
G6PD	B/B	Normal
NUDT15	*1/*1	Normal metabolizer
SLCO1B1	*1/*5	Decreased function
TPMT	*1/*1	Normal metabolizer
UGT1A1	*1/*1	Normal metabolizer

Note: Definitions of these terms are in the Methods and Limitations section on the last page of this report.

How could this impact my medications?

In the "DNA and Medicine" section on the next page, you'll learn which medications could be impacted by these genetic results.

DNA and Medicine

These medicines MAY BE impacted by your genetics

In some cases, pharmacogenetic information may help doctors and pharmacists choose medicines and doses.

The table below points out some medicines that may be affected by your genetic results. If you are taking one of these medicines, talk with your doctor or pharmacist about whether ordering a clinical pharmacogenetic test is right for you.

Medicine	Gene
simvastatin (Zocor®)	*SLCO1B1*
amitriptyline (Elavil®)	*CYP2C19*
citalopram (Celexa®)	*CYP2C19*
clobazam (Onfi®)	*CYP2C19*
clomipramine (Anafranil®)	*CYP2C19*
clopidogrel (Plavix®)	*CYP2C19*
doxepin (Sinequan®)	*CYP2C19*
escitalopram (Lexapro®)	*CYP2C19*
imipramine (Tofranil®)	*CYP2C19*
setraline (Zoloft®)	*CYP2C19*
trimipramine (Surmontil®)	*CYP2C19*
voriconazole (Vfend®)	*CYP2C19*

Just because a medicine is listed here doesn't mean that you should or should not be taking it. Some people with these genetic results still process these medications normally.

DNA and Medicine (continued)

IMPORTANT!

Genetic information is really just one piece of the puzzle.

- It won't tell us if a medicine will definitely work.
- It wont tell us if a medicine will definitely cause side effects or won't work at all.
- It won't tell us exactly how much medicine someone should take.
- It only applies to medicines that you eat, drink, or inject. It doesn't apply to medicines that are rubbed on your skin or used in your eyes or ears.

If your doctor has prescribed medicine for you, keep taking it. It can be dangerous to stop taking a medicine, or to change the dose or timing of it, without first asking your doctor.

Next Steps

What's next?

- **Share this report with your doctors** so they can determine if they should order a clinical pharmacogenetics test.
- Ordering a clinical pharmacogenetics test could be helpful if you see a medication you're currently taking on the table titled "These medicines MAY BE impacted by your genetics."
- Do **not** use this report to make changes to any medicine you take. If your doctor has prescribed medicine for you, keep taking it. It can be dangerous to stop taking a medicine, or to change the dose or timing of it, without first asking your doctor.

What if I have questions?

Ask your doctor or pharmacist.

Because *All of Us* is a research program, we cannot give advice about your medications specifically.

Common Questions

How do I know if this matters for me?

Ask your doctor or pharmacist. Because *All of Us* is a research program, we cannot give advice about your medications specifically.

How did *All of Us* look at my DNA?

You gave a **blood** sample to the *All of Us* Research Program. We processed that blood to get some of your **DNA**. An *All of Us* genetics lab gave a readout of that DNA.

Because you said "Yes" to getting health-related DNA results, a specially trained scientist looked closely at some of the genes in your DNA. We wrote this report for you, based on what they found.

What was done to get this result?

Actually, quite a lot! DNA is in your blood and other samples. You gave a sample to *All of Us*. We processed your sample to extract the DNA. We sent some of your DNA to a special lab. The lab gave a readout of your DNA. A specially trained scientist checked some of the genes in your DNA and wrote this report based on what they found.

What does this mean for my family?

Your DNA is a lot like your family member's DNA, but everyone is different. This result doesn't say anything about their health or their own DNA.

Could my result change?

Yes. *All of Us* could look at more genes or look again at these genes as science improves. Check your *All of Us* account to make sure this is the most up-to-date version of this report.

Methods and Limitations

Methods

This section has some technical information about the test that was performed.

This report represents the analysis of a sample submitted as a part of the *All of Us* Research Program. The sample was collected at <COLLECTION_SITE>. The sample was stored and the DNA was extracted at <BIOBANK_SITE>. Genetic data was generated at <GENOMECENTER_SITE> and interpreted at <CVL_SITE>.

Genomic DNA was extracted from the submitted sample and sequenced using Illumina Next Generation Sequencing. Sequence data was aligned to a reference genome, and variants were identified using a suite of bioinformatic tools designed to detect single nucleotide variants and small insertions/deletions.

This test was developed and its performance characteristics determined by the *All of Us* Research Program, with clinical laboratories accredited by the College of American Pathologists (CAP) and certified under the Clinical Laboratory Improvement Amendments (CLIA) to perform high-complexity testing.

Genes & Alleles

This analysis aims to detect the presence or absence of any of the following alleles, or genotypes at the specified positions: *CYP2C19*: *2, *3, *4, *6, *8, *9, *10, *16, *17, *22, *24, *35; *DPYD*: c.1905+1G>A (*2), c.1129-5923C>G, c.1679T>G (*13), c.2846A>T; *G6PD*:A-202A_376G; A-968C_376G; Asahi; Aures; Canton, Taiwan-Hakka, Gifu-like, Agrigento-like; Chinese-5; Ilesha; Kaiping, Anant, Dhon, Sapporo-like, Wosera; Kambos; Kalyan-Kerala, Jamnaga, Rohini; Mediterranean, Dallas, Panama, Sassari, Cagliari, Birmingham; Quing Yuan, Chinese-4; Seattle, Lodi, Modena, Ferrara II, Athens-like; Sibari; Ube Konan; Union, Maewo, Chinese-2, Kalo; Viangchan, Jammu; *NUDT15*: *2, *3; *SLCO1B1*: *5, *15, *17; *TPMT*: *2, *3A, *3B, *3C; *UGT1A1*: *6, *27, *28, *36, *37

Phenotypes

Term	Definition
Normal Function / Normal Metabolizer / Normal	The gene may act at a rate that is considered average.
Intermediate Metabolizer / Likely Intermediate Metabolizer	The gene may act at a rate that is considered slower than average.
Variable (*G6PD*)	The gene may act at a rate that is considered slower than average, but can be different for different people.
Poor Function / Poor Metabolizer / Likely Poor Metabolizer / Deficient / Deficient with CNSHA (chronic nonspherocytic hemolytic anemia)	The gene may act at a rate that is considered much slower than average.
Rapid Metabolizer	The gene may act at a rate that is considered faster than average.
Ultra-rapid Metabolizer	The gene may act at a rate that is considered much faster than average.

Methods and Limitations (continued)

Phenotypes (continued)

Term	Definition
Indeterminate	This result could not be reported for technical reasons.

Medications

Only the following gene/drug interactions were considered: *CYP2C19* [amitriptyline (Elavil®), citalopram (Celexa®), clobazam (Onfi®), clomipramine (Anafranil®), clopidogrel (Plavix®), doxepin (Sinequan®), escitalopram (Lexapro®), imipramine (Tofranil®), sertraline (Zoloft®), trimipramine (Surmontil®), voriconazole (Vfend®)]; *DPYD* [capecitabine (Xeloda®), fluorouracil (Adrucil®)]; *G6PD* [chloramphenicol, dabrafenib (Tafinlar®), dapsone, hydroxychloroquine (Plaquenil®), local anesthetic containing drugs (e.g.articaine, chloroprocaine, lidocaine, mepivacaine, ropivacaine, tetracaine), mafenide (Sulfamylon®), methylene blue, nalidixic acid (NegGram®), nitrofurantoin (Macrobid®, Macrodantin®, Furadentin®), pegloticase (Krystexxa®), phenazopyridine, primaquine, probenecid (Col-Benemid®), rasburicase (Elitek®), sodium nitrite, sulfacetamide, sulfamethoxazole/trimethoprim (Bactrim®, Septra®), sulfanilamide, sulfasalazine (Azulfidine®), Sulfonylurea drugs[chlorpropamide (Diabinese®), glimepiride (Amaryl®), glipizide (Glucotrol®), glyburide (Diabeta®), tolazamide (Tolinase®), tolbutamide (Orinase®)], tafenoquine (Krintafel®)]; *NUDT15* [azathioprine (Imuran®), mercaptopurine (Purinethol®), thioguanine]; *SLCO1B1* [simvastatin (Zocor®)]; *TPMT* [azathioprine (Imuran®), mercaptopurine (Purinethol®), thioguanine]; *UGT1A1* [atazanavir (Reyataz®), belinostat (Beleodaq®), Irinotecan (Camptosar®)]

Limitations

- **Results provided are from an investigational device.**
- **Because this report is based on data derived from a research study, this information cannot be used to diagnose, cure, mitigate, treat, or prevent disease.**
- **These results could be incorrect.** Based on validation data, results were incorrect < 0.12% of the time.
- This analysis does not detect all possible variants in the tested genes. When *1 (or B in the case of *G6PD*) is reported, it indicates that none of the alleles listed above were identified; it does not rule out the presence of an allele not analyzed by this test and does not rule out the possibility that a non-normal allele is present. This analysis cannot phase variants.
- The reported result may be refined as new alleles are added to the analysis.
- In some cases, observed data can be consistent with more than one possible diplotype, and in these cases the diplotype may be reported as “indeterminate”.
- This analysis cannot distinguish between the more common *1/*3A and the more rare *3B/*3C diplotypes in *TPMT*; clinical phenotypic testing can distinguish between these alleles.
- In very rare cases, such as allogeneic bone marrow transplant, or recent blood transfusion (within 7 days of providing the sample), the results of this analysis may reflect the DNA of the donor. DNA quality may be affected if a participant has received chemotherapy within 120 days of providing the sample. In addition, certain organ transplants or diseases (liver, kidney, heart) may limit the relevance of the results.

Appendix 6: Genotype to Phenotype Translations for the AoU PGx Report

Reference: CPIC Diplotype-Phenotype translation tables for planned alleles/variants to be evaluated as of 1/2020 (CPICPGx.org)

Gene	Diplotype	Reported Predicted Phenotype
CYP2C19	*1/*1	CYP2C19 Normal Metabolizer
CYP2C19	*1/*2	CYP2C19 Intermediate Metabolizer
CYP2C19	*1/*3	CYP2C19 Intermediate Metabolizer
CYP2C19	*1/*4	CYP2C19 Intermediate Metabolizer
CYP2C19	*1/*6	CYP2C19 Intermediate Metabolizer
CYP2C19	*1/*8	CYP2C19 Intermediate Metabolizer
CYP2C19	*1/*9	CYP2C19 Likely Intermediate Metabolizer
CYP2C19	*1/*10	CYP2C19 Likely Intermediate Metabolizer
CYP2C19	*1/*16	CYP2C19 Likely Intermediate Metabolizer
CYP2C19	*1/*17	CYP2C19 Rapid Metabolizer
CYP2C19	*1/*22	CYP2C19 Intermediate Metabolizer
CYP2C19	*1/*24	CYP2C19 Intermediate Metabolizer
CYP2C19	*1/*35	CYP2C19 Intermediate Metabolizer
CYP2C19	*2/*2	CYP2C19 Poor Metabolizer
CYP2C19	*2/*3	CYP2C19 Poor Metabolizer
CYP2C19	*2/*4	CYP2C19 Poor Metabolizer
CYP2C19	*2/*6	CYP2C19 Poor Metabolizer
CYP2C19	*2/*8	CYP2C19 Poor Metabolizer
CYP2C19	*2/*9	CYP2C19 Likely Poor Metabolizer
CYP2C19	*2/*10	CYP2C19 Likely Poor Metabolizer
CYP2C19	*2/*16	CYP2C19 Likely Poor Metabolizer
CYP2C19	*2/*17	CYP2C19 Intermediate Metabolizer
CYP2C19	*2/*22	CYP2C19 Poor Metabolizer
CYP2C19	*2/*24	CYP2C19 Poor Metabolizer
CYP2C19	*2/*35	CYP2C19 Poor Metabolizer
CYP2C19	*3/*3	CYP2C19 Poor Metabolizer
CYP2C19	*3/*4	CYP2C19 Poor Metabolizer
CYP2C19	*3/*6	CYP2C19 Poor Metabolizer
CYP2C19	*3/*8	CYP2C19 Poor Metabolizer

Appendix 6: Genotype to Phenotype Translations for the AoU PGx Report		
Reference: CPIC Diplotype-Phenotype translation tables for planned alleles/variants to be evaluated as of 1/2020 (CPICPGx.org)		
Gene	**Diplotype**	**Reported Predicted Phenotype**
CYP2C19	*3/*9	CYP2C19 Likely Poor Metabolizer
CYP2C19	*3/*10	CYP2C19 Likely Poor Metabolizer
CYP2C19	*3/*16	CYP2C19 Likely Poor Metabolizer
CYP2C19	*3/*17	CYP2C19 Intermediate Metabolizer
CYP2C19	*3/*22	CYP2C19 Poor Metabolizer
CYP2C19	*3/*24	CYP2C19 Poor Metabolizer
CYP2C19	*3/*35	CYP2C19 Poor Metabolizer
CYP2C19	*4/*4	CYP2C19 Poor Metabolizer
CYP2C19	*4/*6	CYP2C19 Poor Metabolizer
CYP2C19	*4/*8	CYP2C19 Poor Metabolizer
CYP2C19	*4/*9	CYP2C19 Likely Poor Metabolizer
CYP2C19	*4/*10	CYP2C19 Likely Poor Metabolizer
CYP2C19	*4/*16	CYP2C19 Likely Poor Metabolizer
CYP2C19	*4/*17	CYP2C19 Intermediate Metabolizer
CYP2C19	*4/*22	CYP2C19 Poor Metabolizer
CYP2C19	*4/*24	CYP2C19 Poor Metabolizer
CYP2C19	*4/*35	CYP2C19 Poor Metabolizer
CYP2C19	*6/*6	CYP2C19 Poor Metabolizer
CYP2C19	*6/*8	CYP2C19 Poor Metabolizer
CYP2C19	*6/*9	CYP2C19 Likely Poor Metabolizer
CYP2C19	*6/*10	CYP2C19 Likely Poor Metabolizer
CYP2C19	*6/*16	CYP2C19 Likely Poor Metabolizer
CYP2C19	*6/*17	CYP2C19 Intermediate Metabolizer
CYP2C19	*6/*22	CYP2C19 Poor Metabolizer
CYP2C19	*6/*24	CYP2C19 Poor Metabolizer
CYP2C19	*6/*35	CYP2C19 Poor Metabolizer
CYP2C19	*8/*8	CYP2C19 Poor Metabolizer
CYP2C19	*8/*9	CYP2C19 Likely Poor Metabolizer
CYP2C19	*8/*10	CYP2C19 Likely Poor Metabolizer

Appendix 6: Genotype to Phenotype Translations for the AoU PGx Report		
Reference: CPIC Diplotype-Phenotype translation tables for planned alleles/variants to be evaluated as of 1/2020 (CPICPGx.org)		
Gene	**Diplotype**	**Reported Predicted Phenotype**
CYP2C19	*8/*16	CYP2C19 Likely Poor Metabolizer
CYP2C19	*8/*17	CYP2C19 Intermediate Metabolizer
CYP2C19	*8/*22	CYP2C19 Poor Metabolizer
CYP2C19	*8/*24	CYP2C19 Poor Metabolizer
CYP2C19	*8/*35	CYP2C19 Poor Metabolizer
CYP2C19	*9/*9	CYP2C19 Likely Intermediate Metabolizer
CYP2C19	*9/*10	CYP2C19 Likely Intermediate Metabolizer
CYP2C19	*9/*16	CYP2C19 Likely Intermediate Metabolizer
CYP2C19	*9/*17	CYP2C19 Likely Intermediate Metabolizer
CYP2C19	*9/*22	CYP2C19 Likely Poor Metabolizer
CYP2C19	*9/*24	CYP2C19 Likely Poor Metabolizer
CYP2C19	*9/*35	CYP2C19 Likely Poor Metabolizer
CYP2C19	*10/*10	CYP2C19 Likely Intermediate Metabolizer
CYP2C19	*10/*16	CYP2C19 Likely Intermediate Metabolizer
CYP2C19	*10/*17	CYP2C19 Likely Intermediate Metabolizer
CYP2C19	*10/*22	CYP2C19 Likely Poor Metabolizer
CYP2C19	*10/*24	CYP2C19 Likely Poor Metabolizer
CYP2C19	*10/*35	CYP2C19 Likely Poor Metabolizer
CYP2C19	*16/*16	CYP2C19 Likely Intermediate Metabolizer
CYP2C19	*16/*17	CYP2C19 Likely Intermediate Metabolizer
CYP2C19	*16/*22	CYP2C19 Likely Poor Metabolizer
CYP2C19	*16/*24	CYP2C19 Likely Poor Metabolizer
CYP2C19	*16/*35	CYP2C19 Likely Poor Metabolizer
CYP2C19	*17/*17	CYP2C19 Ultrarapid Metabolizer
CYP2C19	*17/*22	CYP2C19 Intermediate Metabolizer
CYP2C19	*17/*24	CYP2C19 Intermediate Metabolizer
CYP2C19	*17/*35	CYP2C19 Intermediate Metabolizer
CYP2C19	*22/*22	CYP2C19 Poor Metabolizer
CYP2C19	*22/*24	CYP2C19 Poor Metabolizer

Appendix 6: Genotype to Phenotype Translations for the AoU PGx Report		
Reference: CPIC Diplotype-Phenotype translation tables for planned alleles/variants to be evaluated as of 1/2020 (CPICPGx.org)		
Gene	**Diplotype**	**Reported Predicted Phenotype**
CYP2C19	*22/*35	CYP2C19 Poor Metabolizer
CYP2C19	*24/*24	CYP2C19 Poor Metabolizer
CYP2C19	*24/*35	CYP2C19 Poor Metabolizer
CYP2C19	*35/*35	CYP2C19 Poor Metabolizer
DPYD	*1/*1	DPYD Normal Metabolizer
DPYD	*1/*2	DPYD Intermediate Metabolizer
DPYD	*1/*13	DPYD Intermediate Metabolizer
DPYD	*1/c.1129-5923C>G	DPYD Intermediate Metabolizer
DPYD	*1/2846A>T	DPYD Intermediate Metabolizer
DPYD	*2/*2	DPYD Poor Metabolizer
DPYD	*2/*13	DPYD Poor Metabolizer
DPYD	*2/c.1129-5923C>G	DPYD Poor Metabolizer
DPYD	*2/2846A>T	DPYD Poor Metabolizer
DPYD	*13/*13	DPYD Poor Metabolizer
DPYD	*13/c.1129-5923C>G	DPYD Poor Metabolizer
DPYD	c.1129-5923C>G/c.1129-5923C>G	DPYD Intermediate Metabolizer
DPYD	2846A>T/*13	DPYD Poor Metabolizer
DPYD	2846A>T/c.1129-5923C>G	DPYD Intermediate Metabolizer
DPYD	2846A>T/2846A>T	DPYD Intermediate Metabolizer
G6PD	Male: carrying a nondeficient (class IV) allele. Female: carrying two nondeficient (class IV) alleles.	G6PD Normal
G6PD	Male: carrying a deficient (class II-III) allele. Female: carrying two deficient (class II-III) alleles.	G6PD Deficient
G6PD	Male: carrying a deficient (class I) allele. Female: carrying two deficient (class I) alleles.	G6PD Deficient with CNSHA
G6PD	Female: carrying one nondeficient (class IV) and one deficient (class I-III) alleles.	G6PD Variable
NUDT15	*1/*1	NUDT15 Normal metabolizer
NUDT15	*1/*2	NUDT15 Intermediate Metabolizer

Appendix 6: Genotype to Phenotype Translations for the AoU PGx Report		
Reference: CPIC Diplotype-Phenotype translation tables for planned alleles/variants to be evaluated as of 1/2020 (CPICPGx.org)		
Gene	**Diplotype**	**Reported Predicted Phenotype**
NUDT15	*1/*3	NUDT15 Intermediate Metabolizer
NUDT15	*2/*2	NUDT15 Poor Metabolizer
NUDT15	*2/*3	NUDT15 Poor Metabolizer
NUDT15	*3/*3	NUDT15 Poor Metabolizer
SLCO1B1	*1/*1	SLCO1B1 Normal Metabolizer
SLCO1B1	*5/*5	SLCO1B1 Poor Function
SLCO1B1	*5/*15	SLCO1B1 Poor Function
SLCO1B1	*5/*17	SLCO1B1 Poor Function
SLCO1B1	*15/*15	SLCO1B1 Poor Function
SLCO1B1	*15/*17	SLCO1B1 Poor Function
SLCO1B1	*17/*17	SLCO1B1 Poor Function
SLCO1B1	*1a/*5	SLCO1B1 Decreased Function
SLCO1B1	*1a/*15	SLCO1B1 Decreased Function
SLCO1B1	*1a/*17	SLCO1B1 Decreased Function
SLCO1B1	*1b/*5	SLCO1B1 Decreased Function
SLCO1B1	*1b/*15	SLCO1B1 Decreased Function
SLCO1B1	*1b/*17	SLCO1B1 Decreased Function
TPMT	*1/*1	TPMT Normal Metabolizer
TPMT	*1/*2	TPMT Intermediate Metabolizer
TPMT	*1/*3A	TPMT Intermediate Metabolizer
TPMT	*1/*3B	TPMT Intermediate Metabolizer
TPMT	*1/*3C	TPMT Intermediate Metabolizer
TPMT	*2/*2	TPMT Poor Metabolizer
TPMT	*2/*3A	TPMT Poor Metabolizer
TPMT	*2/*3B	TPMT Poor Metabolizer
TPMT	*2/*3C	TPMT Poor Metabolizer
TPMT	*3A/*3A	TPMT Poor Metabolizer
TPMT	*3A/*3B	TPMT Poor Metabolizer
TPMT	*3A/*3C	TPMT Poor Metabolizer

Appendix 6: Genotype to Phenotype Translations for the AoU PGx Report

Reference: CPIC Diplotype-Phenotype translation tables for planned alleles/variants to be evaluated as of 1/2020 (CPICPGx.org)

Gene	Diplotype	Reported Predicted Phenotype
TPMT	*3B/*3B	TPMT Poor Metabolizer
TPMT	*3B/*3C	TPMT Poor Metabolizer
TPMT	*3C/*3C	TPMT Poor Metabolizer
UGT1A1	*1/*1	UGT1A1 Normal Metabolizer
UGT1A1	*1/*6	UGT1A1 Intermediate Metabolizer
UGT1A1	*1/*27	UGT1A1 Intermediate Metabolizer
UGT1A1	*1/*28	UGT1A1 Intermediate Metabolizer
UGT1A1	*1/*36	UGT1A1 Normal Metabolizer
UGT1A1	*1/*37	UGT1A1 Intermediate Metabolizer
UGT1A1	*6/*6	UGT1A1 Poor Metabolizer
UGT1A1	*6/*27	UGT1A1 Poor Metabolizer
UGT1A1	*6/*28	UGT1A1 Poor Metabolizer
UGT1A1	*6/*36	UGT1A1 Intermediate Metabolizer
UGT1A1	*6/*37	UGT1A1 Poor Metabolizer
UGT1A1	*27/*27	UGT1A1 Poor Metabolizer
UGT1A1	*27/*28	UGT1A1 Poor Metabolizer
UGT1A1	*27/*36	UGT1A1 Intermediate Metabolizer
UGT1A1	*27/*37	UGT1A1 Poor Metabolizer
UGT1A1	*28/*28	UGT1A1 Poor Metabolizer
UGT1A1	*28/*36	UGT1A1 Intermediate Metabolizer
UGT1A1	*28/*37	UGT1A1 Poor Metabolizer
UGT1A1	*36/*36	UGT1A1 Normal Metabolizer
UGT1A1	*36/*37	UGT1A1 Intermediate Metabolizer
UGT1A1	*37/*37	UGT1A1 Poor Metabolizer

CNSHA = chronic nonspherocytic hemolytic anemia

Appendix 7: Evidence for listing drugs on the AoURP PGx Report

Medications are listed on the AoU PGx report if there is evidence of an actionable phenotype-drug assocation. The AoU PGx report does not provide medication dose change or alternative therapy recommendations to participants. The medications names only are listed on the report to provide participants some context on which to have discussions with their heathcare providers to determine whether ordering a clinical test may be desired. The report repeatly indicates this is an investigational/research result and that results hould not be used for clinical decision making.

To be listed; a gene-phenotype-drug combination must appear in FDA-approved drug product labeling (minimally in the Boxed Warning, Dosage and Administration, Contraindications, or Indications section), appear in the FDA Table of Pharmacogenetic Associations (Table of Pharmacogenetic associations for which the data support therapeutic management recommendations specifically) or have a recommendation for alternative medication or dosing modificaiton within a CPIC guideline. For G6PD associations, all FDA-approved drug product

Rubric/semantic logic:

"For this gene" -> ***"if this phenotype is reported" ->*** ***"based on these genotypes" ->*** ***"then, this drug gets listed" ->*** ***"based on this evidence."***

Gene	Predicted Phenotype	Genotypes that are interpreted as this phenotype	Medication [generic (brand)]	Supporting evidence (Drug product labeling, FDA Table of PGx associatiosn, CPIC guideline, or primary literature)	Gene-drug CPIC Level*	Gene-drug PharmGKB Level of Evidence**
TPMT	*TPMT* Poor Metabolizer	Per consensus translation tables (CPIC); See Table 2.	azathioprine (Imuran®)	FDA-approved drug product label (Dosage and Administration, Warnings, Precautions, Drug Interactions, Adverse Reactions, Clinical Pharmacology sections). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations) PMID: 30447069 (CPIC guideline; strong recommendation to consider alternative medications)	A	1A
			mercaptopurine (Purinethol®)	FDA-approved drug product label (Dosage and Administration, Warnings, Precautions, Clinical Pharmacology sections). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations) PMID: 30447069 (CPIC guideline;strong recommendation to reduce dose)	A	1A
			thioguanine	FDA-approved drug product label (Dosage and Administration, Warnings, Precautions, Clinical Pharmacology sections). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations) PMID: 30447069 (CPIC guideline;strong recommendation to reduce dose)	A	1A
	TPMT Intermediate Metabolizer	Per consensus translation tables (CPIC); See Table 2.	azathioprine (Imuran®)	FDA-approved drug product label (Dosage and Administration, Warnings, Precautions, Drug Interactions, Adverse Reactions, Clinical Pharmacology sections). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations) PMID: 30447069 (CPIC guidelines; strong recommendation to reduce dose)	A	1A
			mercaptopurine (Purinethol®)	FDA-approved drug product label (Dosage and Administration, Warnings and Precautions, Adverse Reactions, Clinical Pharmacology sections). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations) PMID: 30447069 (CPIC guideline; strong recommendation to reduce dose)	A	1A
			thioguanine	FDA-approved drug product label (Dosage and Administration, Warnings, Precautions, Clinical Pharmacology sections). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations) PMID: 30447069 (CPIC guideline; moderate recommendation to reduce dose)	A	1A

NUDT15	*NUDT15* Poor Metabolizer	Per consensus translation tables (CPIC); See Table 2.	azathioprine (Imuran®)	FDA-approved drug product label (Dosage and Administration, Warnings, Precautions, Drug Interactions, Adverse Reactions, Clinical Pharmacology sections). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations) PMID: 30447069 (CPIC guidelines; strong recommendation to consider alternative medications)	A	1A
			mercaptopurine (Purinethol®)	FDA-approved drug product label (Dosage and Administration, Warnings and Precautions, Clinical Pharmacology sections). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations) PMID: 30447069 (CPIC guidelines; strong recommendation to reduce dose)	A	1A
			thioguanine	FDA-approved drug product label (Dosage and Administration, Warnings, Precautions, Clinical Pharmacology sections). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations) PMID: 30447069 (CPIC guidelines; strong recommendation to reduce dose)	A	1A
	NUDT15 Intermediate Metabolizer	Per consensus translation tables (CPIC); See Table 2.	azathioprine (Imuran®)	FDA-approved drug product label (Dosage and Administration, Warnings, Precautions, Drug Interactions, Adverse Reactions, Clinical Pharmacology sections). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations) PMID: 30447069 (CPIC guidelines; strong recommendation to reduce dose)	A	1A
			mercaptopurine (Purinethol®)	FDA-approved drug product label (Dosage and Administration, Warnings and Precautions, Clinical Pharmacology sections). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations) PMID: 30447069 (CPIC guidelines; strong recommendation to reduce dose)	A	1A
			thioguanine	FDA-approved drug product label (Dosage and Administration, Warnings, Precautions, Clinical Pharmacology sections). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations) PMID: 30447069 (CPIC guidelines; moderate recommendation to reduce dose)	A	1A
DPYD	*DPYD* poor metabolizer	Per consensus translation tables (CPIC); See Table 2.	capecitabine (Xeloda®)	FDA-approved drug product label (Warnings and Precautions, Patient Counseling sections). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations) PMID: 29152729 (CPIC guideline; strong recommendation to reduce dose)	A	1A
			fluorouracil (Adrucil®)	FDA-approved drug product label (Contraindications, Warnings and Precautions, Patient Counseling sections). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations) PMID: 29152729 (CPIC guideline; strong recommendation to reduce dose)	A	1A
	DPYD intermediate metabolizer	Per consensus translation tables (CPIC); See Table 2.	capecitabine (Xeloda®)	FDA-approved drug product label (Warnings and Precautions, Patient Counseling sections). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations) PMID: 29152729 (CPIC guideline; strong/moderate recommendation to reduce dose)	A	1A
			fluorouracil (Adrucil®)	FDA-approved drug product label (Contraindications, Warnings and Precautions, Patient Counseling sections). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations) PMID: 29152729 (CPIC guideline; strong/moderate recommendation to reduce dose)	A	1A

UGT1A1	*UGT1A1* poor metabolizer	Per consensus translation tables (CPIC); See Table 2.	atazanavir (Reyataz®)	PMID: 26417955 (CPIC guideline; strong recommendation to consider alternative therapy)	A	1A
			belinostat (Beleodaq®)	FDA-approved drug product label (Dosage and Administration, Clinical Pharmacology sections). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations)	B	3
			Irinotecan (Camptosar®)	FDA-approved drug product label (Dosage and Administration, Warning and Precautions, Clinical Pharmacology sections). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations)	A	2A
SLCO1B1	*SLCO1B1* poor function	Per consensus translation tables (CPIC); See Table 2.	simvastatin (Zocor®)	PMID: 24918167 (CPIC guideline; strong recommendation to consider alternative therapy or reduce the dose). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations)	A	1A
	SLCO1B1 decreased function	Per consensus translation tables (CPIC); See Table 2.	simvastatin (Zocor®)	PMID: 24918167 (CPIC guideline; strong recommendation to consider alternative therapy or reduce the dose). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations)	A	1A
CYP2C19	*CYP2C19* poor metabolizer; *CYP2C19* likely poor metabolizer	Per consensus translation tables (CPIC); See Table 2.	amitriptyline (Elaviil®)	PMID: 27997040 (CPIC guideline; moderate recommendation to consider alternative therapy)	A	1A
			citalopram (Celexa®)	FDA-approved drug product label (Dosage and Administration, Warnings, Clinical Pharmacology sections). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations). PMID: 25974703 (CPIC guideline; moderate recommendation to reduce dose)	A	1A
			clobazam (Onfi®)	FDA-approved drug product label (Dosage and Administration, Use in Special Populations, Clinical Pharmacology sections). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations)	C	2A
			clomipramine (Anafranil®)	PMID: 27997040 (CPIC guideline; optional recommendation to consider alternative therapy)	B	2A
			clopidogrel (Plavix®)	FDA-approved drug product label (Boxed Warnings, Warnings and Precautions, Clinical Pharmacology sections). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations) PMID: 23698643 (CPIC guideline; strong recommendation to consider alternative therapy). New Primary Literature: PMIDs; 31479209, 29102571; 29540324; 29280137.	A	1A
			doxepin (Sinequan®)	PMID: 27997040 (CPIC guideline; optional recommendation to consider alternative therapy)	B	3
			escitalopram (Lexapro®)	FDA-approved drug product label (*Adverese Reaction* section). PMID: 25974703 (CPIC guideline; moderate recommendation to reduce dose)	A	1A
			imipramine (Tofranil®)	PMID: 27997040 (CPIC guideline; optional recommendation to consider alternative therapy)	B	2A
			setraline (Zoloft®)	PMID: 25974703 (CPIC guideline; optional recommendation to reduce dose)	B	1A

			trimipramine (Surmontil®)	PMID: 27997040 (CPIC guideline; optional recommendation to consider alternative therapy)	B	2A
			voriconazole (Vfend®)	FDA-approved drug product label (*Clincial Pharmacology* section). PMID: 27981572 (CPIC guideline; moderate recommendation to consider alternative therapy)	A	1A
			flibanserin (Addyi®)	FDA-approved drug product label (*Adverse Reactions, Use in Specific Populations, Clinical Pharmacology* sections). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations)	C	
			pantoprazole (Protonix®)	FDA-approved drug product label (Clincial Pharmacology section). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations)	B	3
			brivaracetam (Briviact®)	FDA-approved drug product label (Clincial Pharmacology section). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations)	B/C	4
	CYP2C19 intermediate metabolizer; *CYP2C19* likely intermediate metabolizer	Per consensus translation tables (CPIC); See Table 2.	clopidogrel (Plavix®)	PMID: 23698643 (CPIC guideline; moderate recommendation to consider alternative therapy). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations) New Primary Literature: PMIDs; 31479209, 29102571; 29540324; 29280137.	A	1A
			brivaracetam (Briviact®)	FDA-approved drug product label (only mentioned in Clincial Pharmacology section). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations)	B/C	4
			clobazam (Onfi®)	"FDA-approved drug product label (Dosage and Administration, Use in Special Populations, Clinical Pharmacology sections). FDA Table of Pharmacogenetic Associations*** (data supports therapeutic management recommendations) "	C	2A
	CYP2C19 rapid metabolizer	Per consensus translation tables (CPIC); See Table 2.	voriconazole (Vfend®)	PMID: 27981572 (CPIC guideline; moderate recommendation to consider alternative therapy)	A	1A
	CYP2C19 ultra-rapid metabolizer	Per consensus translation tables (CPIC); See Table 2.	amitriptyline (Elaviil®)	PMID: 27997040 (CPIC guideline; optional recommendation to consider alternative therapy)	A	1A
			citalopram (Celexa®)	PMID: 27997040 (CPIC guideline; moderate recommendation to consider alternative therapy)	A	1A
			clomipramine (Anafranil®)	PMID: 27997040 (CPIC guideline; optional recommendation to consider alternative therapy)	B	2A
			doxepin (Sinequan®)	PMID: 27997040 (CPIC guideline; optional recommendation to consider alternative therapy)	B	3
			escitalopram (Lexapro®)	PMID: 27997040 (CPIC guideline; moderate recommendation to consider alternative therapy)	A	1A
			imipramine (Tofranil®)	PMID: 27997040 (CPIC guideline; optional recommendation to consider alternative therapy)	B	2A
			trimipramine (Surmontil®)	PMID: 27997040 (CPIC guideline; optional recommendation to consider alternative therapy)	B	2A
			voriconazole (Vfend®)	PMID: 27981572 (CPIC guideline; moderate recommendation to consider alternative therapy)	A	1A

G6PD	*G6PD variable, G6PD d*eficient, *G6PD* deficient with CNSHA****	Per consensus translation tables (CPIC); See Table 2.	chloramphenicol (note: this doesn't apply to medicines that are rubbed on your skin or used in your eyes or ears)	PMID: 24787449 (CPIC guideline and supplement; Patients with G6PD deficiency should be advised that they are at an increased risk of hemolysis after exposure to high-risk drugs and that it is recommended to avoid such compounds).	B	3
			dabrafenib (Tafinlar®)	FDA-approved drug product label (*Warnings and Precautions, Adverse Reactions, Patient Counseling Information* sections).	B/C	
			dapsone	FDA-approved drug product label (*Warnings and Precautions, Precautions, Adverse Reactions, Overdosage, Use in Specific Populations, Patient Counseling Informationn* sections). PMID: 24787449 (CPIC guideline and supplement; Patients with G6PD deficiency should be advised that they are at an increased risk of hemolysis after exposure to high-risk drugs and that it is recommended to avoid such compounds).	B	1B
			hydroxychloroquine (Plaquenil®)	FDA-approved drug product label (*Precautions, Adverse Reactions* sections).	B	
			Local anesthetic containing drugs (e.g. articaine, chloroprocaine, lidocaine, mepivacaine, ropivacaine, tetracaine)	FDA-approved drug product label (*Clinical Pharmacology, Warnings, Warnings and Precautions* sections).	B/C	
			mafenide (Sulfamylon®)	FDA-approved drug product label (*Warning, Adverse Reactions* sections).	B	
			methylene blue	FDA-approved drug product label (C*ontraindications, Warnings and Precautions* sections).	B	3
			nalidixic acid (NegGram®)	FDA-approved drug product label (*Precautions, Adverse Reactions* sections). PMID: 24787449 (CPIC guideline and supplement; Patients with G6PD deficiency should be advised that they are at an increased risk of hemolysis after exposure to high-risk drugs and that it is recommended to avoid such compounds).	B	
			nitrofurantoin (Macrobid®, Macrodantin®, Furadentin®)	FDA-approved drug product label (*Warnings, Adverse Reactions* sections). PMID: 24787449 (CPIC guideline and supplement; Patients with G6PD deficiency should be advised that they are at an increased risk of hemolysis after exposure to high-risk drugs and that it is recommended to avoid such compounds).	B	3
			pegloticase (Krystexxa®)	FDA-approved drug product label (*Boxed Warning, Contraindications, Warnings and Precautions, Patient Counseling* sections).	B	3
			phenazopyridine	PMID: 24787449 (CPIC guideline and supplement; Patients with G6PD deficiency should be advised that they are at an increased risk of hemolysis after exposure to high-risk drugs and that it is recommended to avoid such compounds).	B	3
			primaquine	FDA-approved drug product label (*Contraindications, Warnings, Precautions, Adverse Reactions, Overdosage* sections).	B	3
			probenecid (Col-Benemid®)	FDA-approved drug product label (*Adverse Reactions* sections). PMID: 24787449 (CPIC guideline and supplement; Patients with G6PD deficiency should be advised that they are at an increased risk of hemolysis after exposure to high-risk drugs and that it is recommended to avoid such compounds).	B	
			rasburicase (Elitek®)	FDA-approved drug product label (*Boxed Warning, Contraindications, Warnings and Precautions* sections). PMID: 24787449 (CPIC guideline; strong recommendation that drug is contraindicated)	A	1A
			sodium nitrite	FDA-approved drug product label (*Warnings and Precautions* sections).	B	

			sulfacetamide (note: this doesn't apply to medicines that are rubbed on your skin or used in your eyes or ears)	PMID: 24787449 (CPIC guideline and supplement; Patients with G6PD deficiency should be advised that they are at an increased risk of hemolysis after exposure to high-risk drugs and that it is recommended to avoid such compounds).	B	
			sulfamethoxazole/trimethoprim (Bactrim®, Septra®)	FDA-approved drug product label (*Precautions* sections). PMID: 24787449 (CPIC guideline and supplement; Patients with G6PD deficiency should be advised that they are at an increased risk of hemolysis after exposure to high-risk drugs and that it is recommended to avoid such compounds).	B	3
			sulfanilamide	PMID: 24787449 (CPIC guideline and supplement; Patients with G6PD deficiency should be advised that they are at an increased risk of hemolysis after exposure to high-risk drugs and that it is recommended to avoid such compounds).		
			sulfasalazine (Azulfidine®)	FDA-approved drug product label (*Precautions* sections). PMID: 24787449 (CPIC guideline and supplement; Patients with G6PD deficiency should be advised that they are at an increased risk of hemolysis after exposure to high-risk drugs and that it is recommended to avoid such compounds).	B	4
			Sulfonylurea drugs [chlorpropamide (Diabinese®), glimepiride (Amaryl®), glipizide (Glucotrol®), glyburide (Diabeta®), tolazamide (Tolinase®), tolbutamide (Orinase®)]	FDA-approved drug product label (*Adverse Reactions, Precautions, and/or Warnings and Precautions* sections). [for glyburide only]]PMID: 24787449 (CPIC guideline and supplement; Patients with G6PD deficiency should be advised that they are at an increased risk of hemolysis after exposure to high-risk drugs and that it is recommended to avoid such compounds).	Most B	
			tafenoquine (Krintafel®)	FDA-approved drug product label (*Dosage and Administration, Contraindications, Warnings and Precautions, Use in Specific Populaitons, Patient Counseling Information* sections).	A	

* CPIC level A and B gene/drug pairs are defined as having sufficient evidence for at least one prescribing action to be recommended. See https://cpicpgx.org/prioritization/#leveldef

** PharmGKB creates Clinical Annotation Levels of Evidence based on primary literature annoations. See: https://www.pharmgkb.org/page/clinAnnLevels

*** FDA Table of Pharmacogenetic Associations, 2/25/2020; See https://www.fda.gov/medical-devices/precision-medicine/table-pharmacogenetic-associations

**** CNSHA = chronic nonspherocytic hemolytic anemia

Informing loop for Medicine and your DNA report

Participants who complete the Consent to Get DNA Results have the option to respond “Yes,” “No,” or “I’m not sure right now” to receiving some or all of their genomic results. While the necessary information for informed decision-making is provided in the Consent to Get DNA Results, the program will employ just-in-time informing loops to remind participants of salient risk and benefit information, as well as topical educational content, for those who consent to the return of their genomic results.

These informing loops will roll out as the program develops readiness to support responsible return for each type of genetic result. Participants will be able to opt-in to each type of result return through these informing loops, allowing for greater granular control to meet their individual needs.

Below is the informing loop to receive the *All of Us* Medicine and your DNA report. At the time the program is ready to return pharmacogenetics results, individuals who responded “Yes” to the Consent to Get DNA Results will be provided with this informing loop. After reviewing the informing loop, participants can choose to opt-in or out of receiving their personalized Medicine and your DNA report. If they opt-in, their Medicine and your DNA report will be generated by the program and made available to the participant.

Screen 1: How do I get my results about medicine and my DNA?

There are both benefits and risks to getting your results about medicine and your DNA. Before you decide, please review the following information.

If you decide you want your DNA results for hereditary disease risk, you can get them by answering “Yes, I want results about medicine and my DNA” at the end.

Questions?
1-844-842-2855
help@joinallofus.org
Chat Live

[Buttons]
Cancel/Next

Screen 2: What will my results tell me about medicine and my DNA?

Certain genes in our DNA affect how we respond to some medicines. Some genes help move medicines to the right part of your body. Some genes help break down medicines and clear them from your body. Some genes even change medicines into a form that makes them work properly.

Small differences in your DNA can change how these genes do their work. These differences can be one reason some medicines work differently for different people.

These results will look at a few of the genes in your DNA that can affect how your body responds to some medicines. For example, some genes, like the gene *CYP2C19*, affect the way people's bodies break down certain medicines. People with a different version of the *CYP2C19* gene may break down medicines more slowly than other people. This includes medicines such as some antidepressants and a blood thinner. [add Tooltip; see below. *Note: A Tooltip is extra information a participant can view by clicking on an "information" icon that is present on the e-screen.*]

[Tool-tip –
The technical word for this kind of information is "pharmacogenetics". In some cases, pharmacogenetic information may help doctors and other healthcare providers choose what medicines to give you and how much.

There are a lot of genes that can affect how your body responds to medicines. We won't look at all of them. We will only look at a few of these genes in your DNA.]

Questions?
1-844-842-2855
help@joinallofus.org
Chat Live

[Buttons]
Previous/Next

Screen 3: What are the risks of getting my results about medicine and my DNA?

When you signed the Consent to Get DNA Results, you learned that there are different kinds of risks to learning about your DNA. Here are some that are important to remember for these results.

- **If your doctor or healthcare provider has prescribed medicine for you, keep taking it as prescribed**. It can be dangerous to stop taking a medicine, or to change the dose or timing of it, without first talking to your doctor or healthcare provider.
- Your results could show that a medicine you are taking may be affected by your DNA. If this happens, your doctor or healthcare provider may decide to order a clinical test for you. This is because *All of Us* is a research program and results from a research program cannot be used directly in clinical care.
    - If your doctor or healthcare provider orders a clinical test, you or your insurance may be billed for it.
    - You can decide what care is right for you. Changing your care may cost more.

- You could get information you weren't expecting in your results. For example:
    - Your results may make you wonder if you are related to a family member in the way you thought. Keep in mind that DNA results are not always the same for all family members.

You can contact the *All of Us* Support Center at any time. We can help answer your questions. We can also help you find more resources.

Questions?
1-844-842-2855
help@joinallofus.org
Chat Live

[Buttons]
Previous/Next

Screen 4: What are the benefits of getting my results about medicine and my DNA?

You results could show you have a version of a gene that may affect how your body responds to some medicines. If you do, your results will also contain a list of specific medicines that may be affected.

Doctors and other healthcare providers may use this information to determine if you should get a clinical test. They can use the results of that clinical test when they consider what medicines to give you and how much.

To learn more about this type of DNA result and others, visit the Learning Center. Learning more may help you decide if you would like to see this information.

Questions?
1-844-842-2855
help@joinallofus.org
Chat Live

[Buttons]
Previous/Next

Screen 5: What are the limits of getting my results about medicine and my DNA?

- **These results are research results**. That means that neither you nor your doctor or healthcare provider should use them to make any changes to your medicines. Your doctor will need to confirm these results with a separate clinical test if they want to use them in your care.
- **Doctors and other healthcare providers don't make decisions based only on DNA.** Some other important considerations can be age, weight, health, diet, and other medicines you are taking at the same time.
- These results are based on current scientific understanding. There is a chance they could be incorrect. As we learn more information, *All of Us* could look at more genes or look at these genes again to provide new results. [add Tooltip; see below]

[Tooltip -
What we know about DNA and health comes from many scientific studies. Most of these studies were about people with European genetic ancestry. For people with different genetic ancestry, what we know about DNA and health may not be as complete. Some DNA changes may be found more often or only among people from specific ancestry groups. Scientists may not yet know about them if these groups were not included in previous studies.

All of Us is working to dramatically increase who gets included in scientific studies. Including more people from different genetic ancestry groups in scientific studies will help increase our understanding about DNA and health for people of all backgrounds.]

Questions?
1-844-842-2855
help@joinallofus.org
Chat Live

[Buttons]
Previous/Next

Screen 6: What are my choices?

It is your choice whether you want to get your results about medicine and your DNA.

If you say 'Yes, I want my results about medicine and my DNA.':

- A specially trained scientist will look closely at some of the genes in your DNA related to how your body responds to some medicines. They will generate results for you based on what they find.
- The results will list the genes they looked at and the medicines that may be affected. They will point out if your DNA may affect how your body processes these medicines.

If you say 'No, I do not want results about medicine and my DNA':

- No one will look at your DNA in this way.

If you say 'I'm not sure right now':

- You can come back and make a decision later.
- We won't generate your personalized results until you tell us you want us to.

You will be able to decide separately if you want other types of DNA results, like your risk for certain health conditions. We will send you messages when new results are available.

Questions?
1-844-842-2855
help@joinallofus.org
Chat Live

[Buttons]
Previous/Next

Screen 7: Would you like your results about medicine and your DNA?

- Yes, I want my results about medicine and my DNA.
    - I understand I will be able to see my results once they are ready. These results will take some time to generate. *All of Us* will contact me when my results are ready.
- No, I do not want my results about medicine and my DNA.
    - I know this means I can change my mind later.
- I'm not sure right now.
    - I know this means I can change my mind later.

This decision does not affect your ability to get other types of DNA results, like your risk for certain health conditions. *All of Us* will contact you when other results are available.

Questions?
1-844-842-2855
help@joinallofus.org
Chat Live

[Buttons]
Previous/Finish

Screen 8a: Decision Confirmation - Yes

Thanks for taking the time to consider the benefits and risks of getting your results about medicine and your DNA.

Your decision:
Yes, I want my results about medicine and my DNA.

These results will take some time to generate. We will notify you when your results are ready. In the meantime, you can learn more about DNA and medicine in the Learning Center.

Questions?
1-844-842-2855
help@joinallofus.org
Chat Live

[Button]
Done

Screen 8b: Decision Confirmation - No

Thanks for taking the time to consider the benefits and risks of getting your results about medicine and your DNA.

Your decision:
No, I do not want my results about medicine and my DNA.

If you change your mind and would like to get your results about medicine and your DNA, you can update your decision in the “Agreements” section of your account at any time.

Questions?
1-844-842-2855
help@joinallofus.org
Chat Live

[Button]
Done

Screen 8c: Decision Confirmation - I’m not sure right now

Thanks for taking the time to consider the benefits and risks of getting your results about medicine and your DNA.

Your decision:
I’m not sure right now.

That’s ok. Take your time so you can make a decision that’s right for you. To learn more about DNA and medicine, visit the Learning Center. Learning more may help you decide if you want to see this information.

If you change your mind and would like to get your results about medicine and your DNA, you can update your decision in the “Agreements” section of your account at any time.

Questions?
1-844-842-2855
help@joinallofus.org
Chat Live

[Button]
Done

Table of Contents

1 Executive Summary

Introduction

In compliance with the FDA requirement to assess participant comprehension of the *All of Us* Research Program (AoURP, the Program) Reports, a mixed methods research approach was employed to assess the content validity of the survey items and participant comprehension of report-specific concepts in a broader participant cohort through a computer-administered survey. Presented here are the research methodology and findings from the quantitative arm of the study.

Methods

Participants were recruited through Respondent, a user experience participant recruitment platform. Study objectives and procedures were shared with participants through a virtual study information session. All participants who attended the information session were invited to take the survey through SurveyMonkey. Internal quality control measures were implemented within the survey to ensure that participants reviewed the reports and to allow for exclusion of data from participants who did not review the reports. Demographic characteristics were calculated using descriptive statistics (e.g., mean, frequency), and total percent pass rate for each question, domain, and general concept were calculated.

Results

Survey response rates were 31.5% for the Positive Hereditary Disease Risk (HDR) Report (e.g., *BRCA1*), 31.0% for the Pharmacogenetics Report (hereafter referred to as 'Medicine and Your DNA Report'), and 23.8% for the Uninformative HDR Report. Participants were 45 years of age or older (n=401, 48.9%), female (n=526, 64.1%), non-white (n=366, n=44.6%), Latino/a (n=125, 15.2%), had an associate degree or less education (n=417, 50.9%), and earned $74,999 annually or less (n=498, 60.7%). Participant comprehension rates for the Positive HDR Report (n=347) were 96.9% (96.7% genetic knowledge, 97.5% self-efficacy concepts), 96.6% for the Uninformative HDR Report (n=287; 94.6% genetic knowledge, 98.6% self-efficacy concepts), and 98.1% for the Medicine and Your DNA Report (n=205; 97.6% genetic knowledge, 98.4% self-efficacy concepts).

Conclusion

Participants were able to understand the AoURP Positive HDR, Uninformative HDR, and the Medicine and Your DNA Reports.

2 Introduction

In compliance with the FDA requirement to assess participant comprehension of the *All of Us* Research Program (AoURP, the Program) Reports, a mixed methods research approach was employed to assess the content validity of the survey items and comprehension in a broader participant cohort through a computer administered survey. Presented here are the research methodology and results from the quantitative arm of the study.

3 Objectives

- To quantitatively assess participant comprehension of health-related genomic information presented on the AoURP Reports.

4 Quantitative Research Methods

This study is a quantitative research study to assess the comprehension of the AoURP Hereditary Disease Risk (HDR) and Pharmacogenetics (hereafter referred to as 'Medicine and Your DNA') Reports through a structured, web-administered survey. Surveys were administered after participants had reviewed the mock reports with the specific aims of assessing participant understanding of the implication of the test results.

4.1 Objectives

- Assess participant comprehension of the implications of a Positive AoURP HDR Report (e.g., *BRCA1*), the Uninformative HDR Report, Medicine and Your DNA Report.

4.2 Study Design

This is a web-administered survey study to evaluate participant understanding of the implication of a AoURP Positive HDR Report (Appendix 7.1), the Uninformative HDR Report (Appendix 7.2), and Medicine and Your DNA Report (Appendix 7.3).

4.3 Sampling Strategy and Participant Population

Participants were recruited through Respondent, a user experience participant recruitment platform after meeting eligibility criteria. Eligibility criteria was established based on the inclusion/exclusion criteria (Table 1) and recorded through the completion of a screening questionnaire. Eligible participants were invited to attend the study information session, after which they were invited to take the survey. During the information session, study objectives and procedures were shared with participants.

Table 1. Participant inclusion/exclusion criteria

The study inclusion/exclusion criteria are detailed in Table 1. These criteria, the associated screening, and enrollment forms were established by the AoURP. This process ensured that recruited participants reflected (or as closely as feasible) that of the AoURP study demographics (50% non-European, 80% underrepresented individuals in biomedical research).

Inclusion Criteria	Exclusion Criteria
1. Participant has never undergone genetic testing. 2. Participant has never reviewed genetic testing reports or materials. 3. Participant is not enrolled in the AoURP. 4. Participant resides in the United States.	1. Participant is enrolled in the AoURP. 2. Participant has no access to a computer or internet connection. 3. Participant did not read the report.

4.4 Recruitment Strategy

To ensure a participant cohort that closely reflected the AoURP participant demographics (50% non-European, 80% underrepresented individuals in biomedical research), three participant recruitment pools (race, education, and age) were created for recruitment. Recruitment pool 1 (race) gathered responses from only Black/African American, American Indian/Alaskan Native (AIAN), Asian or Pacific Islander, or Latino. Pool 2 (education) only gathered responses from individuals with “some college; no degree” or lower. Pool 3 (age) only gathered responses from participants 45 years or older (Appendix 7.4).

Recruitment screeners (Appendix 7.5) were created for each specific recruitment pool to recruit participants reflecting (or as closely as feasible) the AoURP participant demographics.

4.5 Procedures

In addition to demographic characteristics, two report-specific domains -- genetic knowledge, self-efficacy -- were used to assess participant comprehension of each report. Participants were recruited through Respondent. Participants were invited to attend a study information session after meeting study eligibility criteria (Table 1). All participants having attended the information sessions were given the link to take the survey on SurveyMonkey. Figures 1-3 outline the inclusion/exclusion of participants per recruitment pool.

Figure 1. Inclusion/exclusion of Positive HDR Report-specific participants

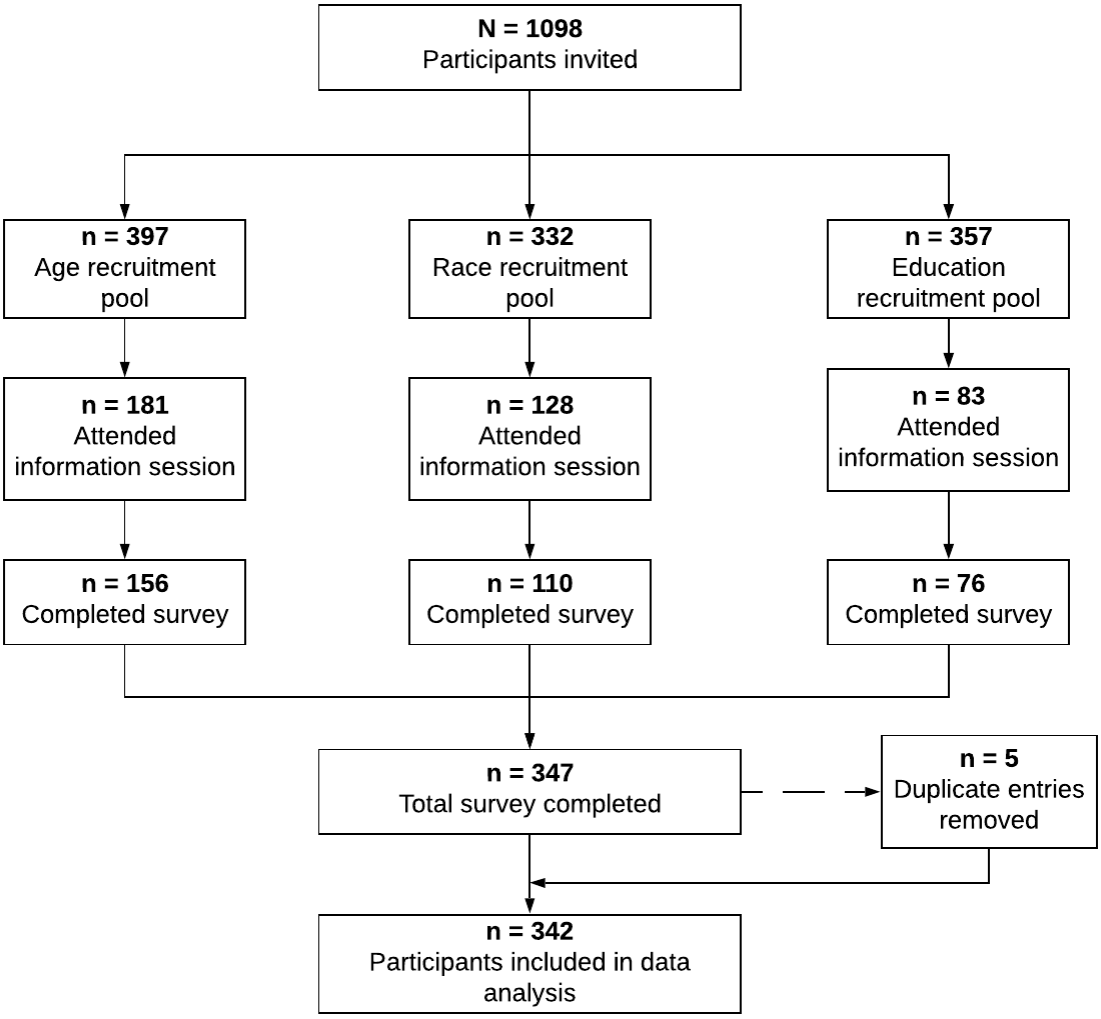


Figure 2. Inclusion/exclusion of Uninformative HDR Report-specific participants

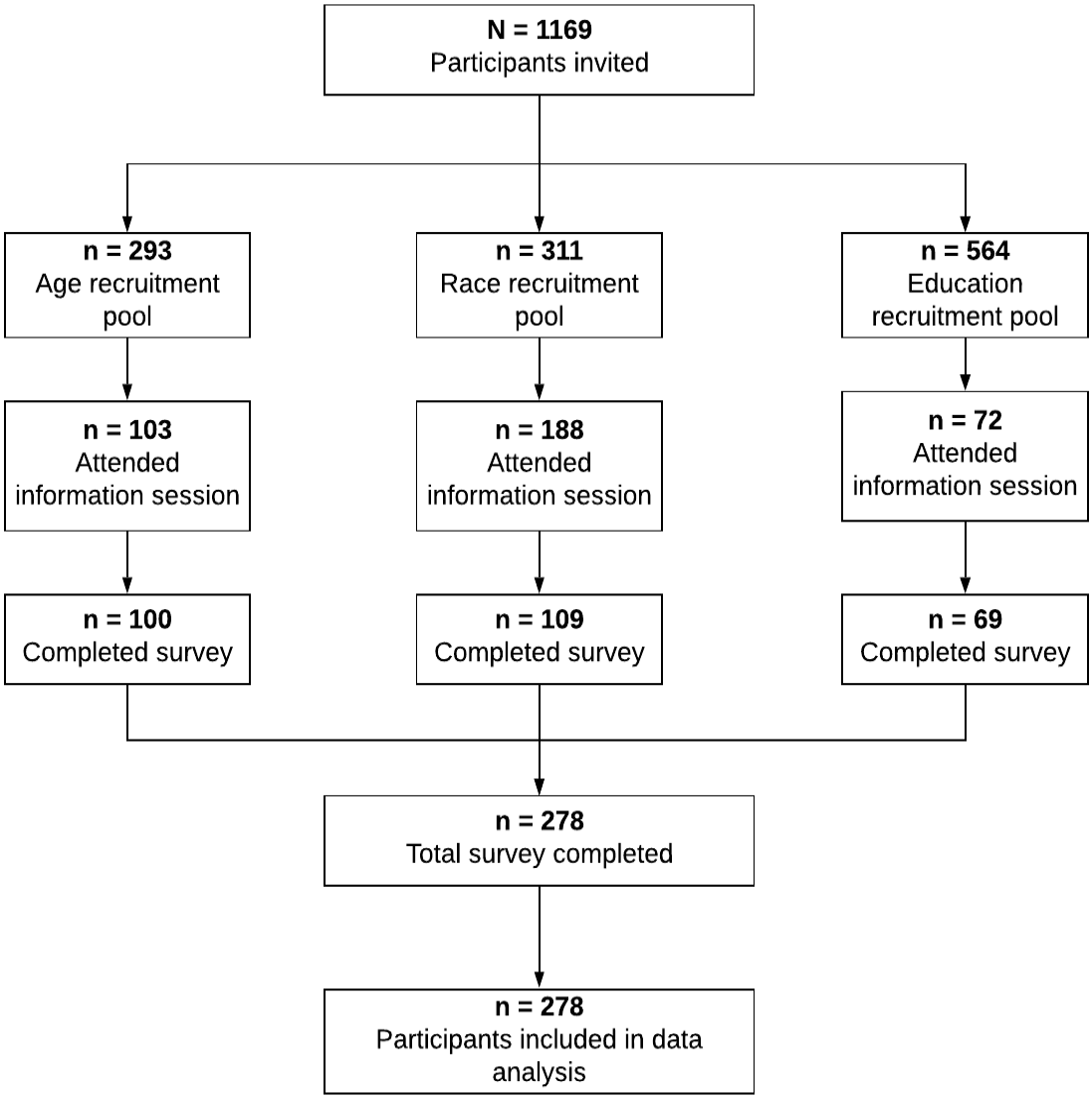
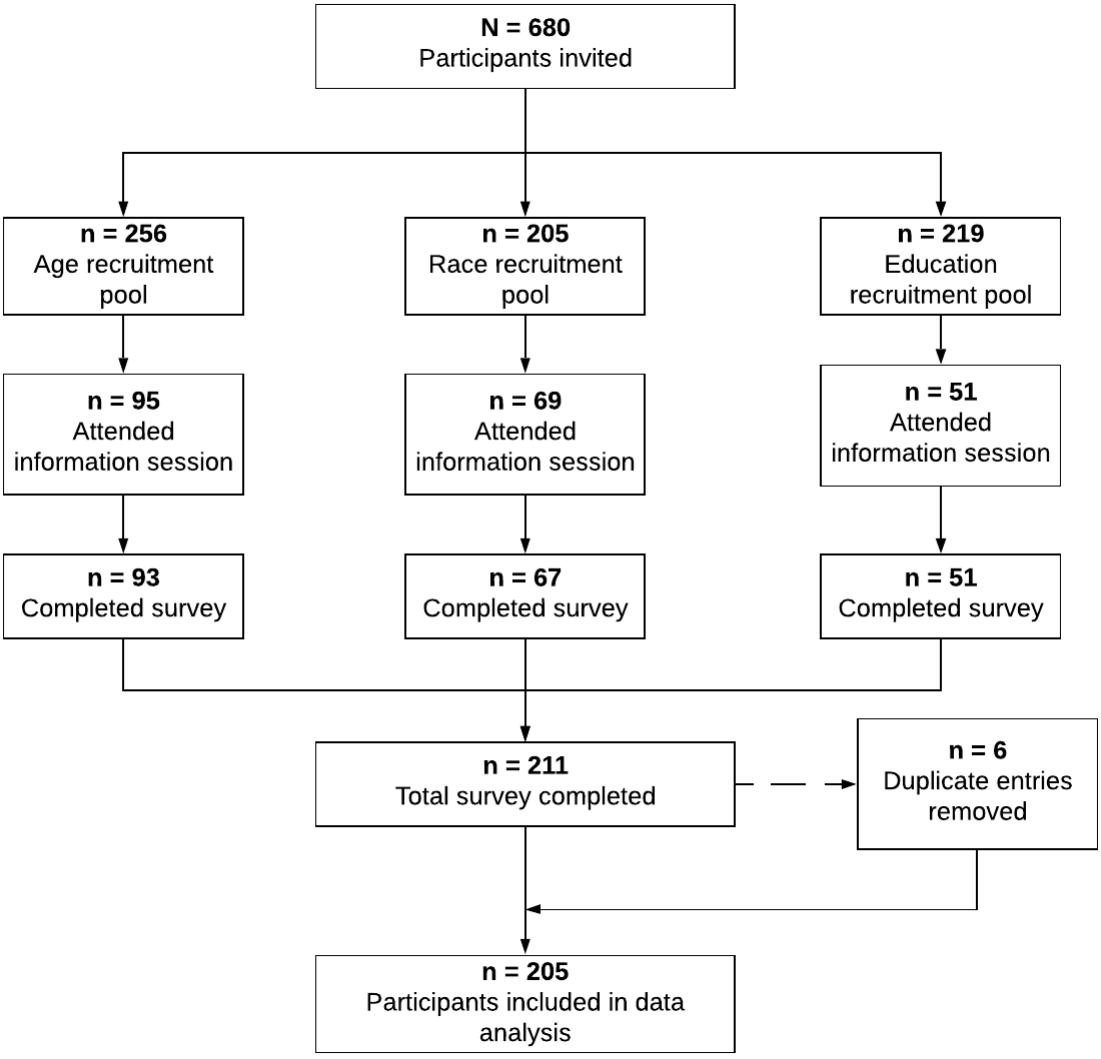


Figure 3. Inclusion/exclusion of the Medicine and Your DNA Report-specific participants



During the information session, study objectives and procedures were shared with participants. Participants were reminded that their involvement in the study was optional and that they could opt-out of the study at any time. Only participants that attended the information session were invited to take the survey.

Reports were shared with participants 12-hours prior to them taking the survey. Participants were reminded to review the report prior to taking the survey. In addition, reports were embedded within the survey platform, and participants were reminded that they could refer to the report any time while completing the survey. Internal quality control (QC) measures were embedded into each survey to ensure that participants actually read/reviewed the report before taking the survey. This also allowed us to disregard data from participants who failed to correctly

answer the QC questions. To avoid duplicate responses, participants were restricted from taking the survey twice.

4.6 Analysis

Survey response rates were calculated for each recruitment pool. Response rates were calculated using the following formula:

$$\text{Response rate} = \frac{\text{(Number of surveys completed)}}{\text{(Total number of participants invited)}} \times 100$$

Demographic characteristics, as well as domain-specific comprehension rate of all surveyed participants, was summarized using descriptive statistics (e.g., means and frequencies). In addition to an average comprehension score per report, the percentage of participants who correctly passed each survey question was calculated for each domain.

For both genetic knowledge and self-efficacy domains, the arithmetic mean (average) per domain was calculated using the following formula:

$$\text{Average} = \frac{\text{(Comprehension rate of items 1, 2, 3…)}}{\text{(Total number of items in that domain)}}$$

5 Results

Total survey response rate was 31.5% for the Positive HDR Report, 31.0% for the Medicine and Your DNA Report, and 23.8 % for the Uninformative HDR Report.

5.1 Summary of demographic characteristics

For the Positive HDR Report, participants (n=342) were female (n=206, 60.2%), 45 years of age or older (n=225, 65.8%), non-white (n= 156, 45.6%), had an associate degree or less education (n=150, 43.9%), employed for wages (n=187, 54.7%), and earned $74,999 annually or less (n=192, 56.1%).

Participants who reviewed the Uninformative HDR Report (n=278) were female (n=193, 69.4%), 45 years old of age or older (n=118, 42.4%), non-white (n=114, 41.0%), had an associate degree or less education (n=162, 58.3%), employed for wages (n=130, 46.8%), and earned $74,999 annually or less (n=184, 66.2%).

Participants who reviewed the Medicine and Your DNA Report (n=205) were mostly 45 years of age or older (n=119, 59.5%), female (n=127, 63.5%), non-white (n=96, 48.0%), had an associate degree or less education (n=105, 52.5%), employed for wages (n=92, 46.0%), and earned $74,999 annually or less (n=122, 61.0%).

Table 2. Demographic characteristics of participants who reviewed the Positive HDR Reports, Uninformative HDR Reports, and the Medicine and Your DNA Reports

Report types		**All Reports n (%)**	**Positive HDR Report n (%)**	**Uninformative HDR Report n (%)**	**Medicine and Your DNA Report n (%)**
Total		**825 (100)**	**342(100)**	**287(100)**	**205(100)**
Age (Years)	**18 -24**	175 (21.3)	58 (17.0)	62 (22.3)	55 (26.8)
	25 - 34	142 (17.3)	59 (17.3)	52 (18.7)	31 (15.1)
	35 - 44	107 (13.1)	39 (11.4)	46 (16.5)	22 (10.7)
	45 - 54	260 (31.1)	121 (35.4)	75 (27.0)	64 (31.2)
	55 - 64	121 (14.8)	51 (14.9)	39 (14.0)	31 (15.1)
	65+	20 (2.4)	14 (4.1)	4 (1.4)	2 (1.0)
Gender	**Male**	299 (36.3)	136 (39.8)	85 (30.6)	78 (38.0)
	Female	526 (63.7)	206 (60.2)	193 (69.4)	127 (62.0)
Race	**Non-white**	366 (44.4)	156 (45.6)	114 (41.0)	96 (46.8)
Ethnicity	**Hispanic/Latino**	125 (15.0)	49 (14.3)	45 (16.2)	31 (15.1)
Education	**Some high school (no degree)**	4 (0.5)	2 (0.6)	2 (0.7)	0 (0)
	High school graduate or equivalent	70 (8.4)	16 (4.7)	35 (12.6)	19 (9.3)

	Some college, no degree	275 (33.5)	102 (29.8)	102 (36.7)	71 (34.6)
	Associate degree	68 (8.3)	30 (8.8)	23 (8.3)	15 (7.3)
	Bachelor's degree	278 (33.5)	133 (38.9)	79 (28.4)	66 (32.2)
	Graduate or professional degree	130 (15.7)	59 (17.3)	37 (13.3)	34 (16.6)
Employment status	**Employed for wages full time**	309 (37.6)	147 (43.0)	92 (33.1)	70 (34.1)
	Employed for wages part time	100 (12.2)	40 (11.7)	38 (13.7)	22 (10.7)
	Self-employed	156 (18.8)	68 (19.9)	46 (16.5)	42 (20.5)
	Out of work for less than 1 year	72 (8.8)	26 (7.6)	31 (11.2)	15 (7.3)
	Out of work for 1 year or more	4 (0.5)	1 (0.3)	2 (0.7)	1 (0.5)
	Retired	29 (3.5)	14 (4.1)	7 (2.5)	8 (3.9)
	A student	108 (13.2)	32 (9.4)	46 (16.5)	30 (14.6)
	A homemaker	27 (3.3)	8 (2.3)	9 (3.2)	10 (4.9)
	Unable to work	19 (2.2)	6 (1.8)	7 (2.5)	6 (2.9)
Household income	**USD 24,999 or less**	161 (19.6)	55 (16.1)	62 (22.3)	44 (21.5)
	USD 25,000 - 34,999	91 (11.0)	32 (9.4)	37 (13.3)	22 (10.7)
	USD 35,000 - 49,999	86 (10.5)	33 (9.6)	33 (11.9)	20 (9.8)
	USD 50,000 - 74,999	160 (19.4)	72 (21.1)	52 (18.7)	26 (12.7)
	USD 75,000 - 99,999	116 (14.2)	55 (16.1)	29 (10.4)	32 (15.6)

	USD 100,000 - 199,999	175 (21.0)	77 (22.5)	54 (19.4)	44 (21.5)
	USD 200,000 or more	36 (4.4)	18 (5.3)	11 (4.0)	7 (3.4)

5.2 Comprehension rates per report type

5.2.1 Participant comprehension of the positive hereditary disease risk report (n=342)

Average participant comprehension of the Positive HDR Report (n=342) was 97.1%, with 96.7% for genetic knowledge (n=331) and 97.7% for self-efficacy (n=334) concepts.

General concept	Survey tem	Comprehension rate, n (%)
Genetic knowledge	Something significant for my health was found in the genes that were checked.	326 (95.3)
Genetic knowledge	My DNA test results tell me I definitely have cancer.	335 (98.0)
Genetic knowledge	My DNA test results tell me I will definitely get cancer in the future.	335 (98.0)
Genetic knowledge	If something significant for my health was found in my DNA, my relatives could have the same results.	337 (98.5)
Genetic knowledge	Doctors only consider DNA when making treatment decisions.	320 (93.6)
Self-efficacy	I understand I could share my DNA test results with my doctor.	341 (99.7)
Self-efficacy	I understand my DNA test result is a research result.	333 (97.4)
Self-efficacy	I understand I should not change my medical care based on my DNA test results.	323 (94.4)
Self-efficacy	I understand I could share my DNA test results with my family.	340 (99.4)
	Average comprehension rate, n (%)	332 (97.1)

For both genetic knowledge and self-efficacy domains, the arithmetic mean (average) per domain was calculated using the following formula:

$$\text{Average} = \frac{(\text{Comprehension rate of items 1, 2, 3…})}{(\text{Total number of items in that domain})}$$

Such that the average comprehension rate per domain was 96.7%, and 97.7% for genetic knowledge and self-efficacy concepts, respectively.

5.2.2 Participant comprehension rate of the Uninformative hereditary disease risk report (n=278)

Average participant comprehension of the Uninformative HDR Report (N=278) was 96.6%; 94.6% for genetic knowledge (n=263) and 98.6% for self-efficacy (n=274) concepts.

General concept	Survey item	Comprehension rate, n (%)
Genetic knowledge	My DNA test results do not eliminate my risk of developing a hereditary condition.	270 (97.1)
Genetic knowledge	Doctors only consider DNA when making treatment decisions.	257 (92.4)
Genetic knowledge	Aside from my DNA, other factors such as age and weight may affect my risk of developing diseases.	262 (94.2)
Self-efficacy	I understand that nothing significant for my health was found in the genes that were checked.	269 (96.8)
Self-efficacy	I understand my DNA test result is a research result.	276 (99.3)
Self-efficacy	I understand I should not change my medical care based on my DNA test results.	277 (99.6)
	Average comprehension rate, n (%)	269 (96.6)

For both genetic knowledge and self-efficacy domains, the arithmetic mean (average) per domain was calculated using the following formula:

$$\text{Average} = \frac{\text{(Comprehension rate of items 1, 2, 3…)}}{\text{(Total number of items in that domain)}}$$

Such that the average comprehension rate per domain was 94.6%, and 98.6% for genetic knowledge and self-efficacy concepts, respectively.

5.2.3 Participant comprehension rate of the Medicine and your DNA report (n=205)

Average participant comprehension of the Medicine and Your DNA Report (N=205) was 98.1%; 97.6% for genetic knowledge (n=195) and 98.4% for self-efficacy (n=197) concepts.

General concept	Survey items	Comprehension rate, n (%)
Genetic knowledge	Doctors and pharmacists only consider DNA when making treatment decisions.	199 (97.1)
Genetic knowledge	Aside from my DNA, other factors such as age and weight may impact my response to medicines.	205 (100.0)
Genetic knowledge	My DNA may impact how I respond to certain medicines.	203 (99.0)
Genetic knowledge	My DNA test results tell me which medicines will definitely work for me.	193 (94.1)
Self-efficacy	Based on my DNA test results, I can stop taking my prescribed medicines before talking to my doctor.	205 (100.0)
Self-efficacy	Based on my DNA test results, I can change how often I take my prescribed medicines before talking to my doctor.	204 (99.5)
Self-efficacy	I understand I could share my DNA test results with my doctor.	205 (100.0)
Self-efficacy	I understand my DNA test result is a research result.	198 (96.6)
Self-efficacy	I understand I should not change my medical care based on my DNA test results.	196 (95.6)
	Average comprehension rate, n (%)	**201 (98.0)**

For both genetic knowledge and self-efficacy domains, the arithmetic mean (average) per domain was calculated using the following formula:

$$\text{Average} = \frac{\text{(Comprehension rate of items 1, 2, 3…)}}{\text{(Total number of items in that domain)}}$$

Such that the average comprehension rate per domain was 97.6%, and 98.4% for genetic knowledge and self-efficacy concepts, respectively.

6 Summary

Average participant comprehension of the Positive HDR Report (n=342) was 97.1%; 96.7% for genetic knowledge (n=331) and 97.7% for self-efficacy (n=334) concepts. Average participant comprehension of the Uninformative HDR Report (n=278) was 96.6%; 94.6% for genetic knowledge (n=263) and 98.6% for self-efficacy (n=274) concepts. Average participant comprehension of the Medicine and Your DNA Report (N=205) was 98.1%; 97.6% for genetic knowledge (n=195) and 98.4% self-efficacy (n=197) concepts.

Results from this study indicate that participants were able to understand the important concepts presented in the Positive HDR Report, Uninformative HDR Report, and the Medicine and Your DNA Report

7 Questionnaires

7.1 Demographic questionnaire

1. **What was your sex at birth?**
    - ❑ Male
    - ❑ Female

2. **What is your age?**
    - ❑ 17 years old or less ← Exclude
    - ❑ 18 - 24 years old
    - ❑ 25 - 34 years old
    - ❑ 45 - 54 years old
    - ❑ 55 - 64 years old
    - ❑ 65 years or older

3. **Are you Hispanic, Latino/a or Spanish origin?**
    - ❑ Yes
    - ❑ No
    - ❑ I don't know

4. **Which one or more of the following would you say is your race?**
    - ❑ White
    - ❑ Black or African American
    - ❑ American Indian or Alaska Native
    - ❑ Asian
    - ❑ Pacific Islander
    - ❑ Other

5. **Are you…**
    - ❑ Married
    - ❑ Divorced
    - ❑ Widowed
    - ❑ Separated
    - ❑ Never married

6. **What is the highest grade or year of school you completed?**
    - ❑ Never attended school or only attended kindergarten
    - ❑ Grades 1-8 (Elementary)
    - ❑ Grades 9-11 (Some high school)
    - ❑ Grade 12 or GED (High school graduate)
    - ❑ Some College, no degree
    - ❑ Associate degree

- ❑ Bachelor's degree
- ❑ Graduate or professional degree

7. **Are you currently...?**
   - ❑ Employed for wages full time
   - ❑ Employed for wages part-time
   - ❑ Self-employed
   - ❑ Out of work for 1 year or more
   - ❑ Out of work for less than 1 year
   - ❑ A Homemaker
   - ❑ A Student
   - ❑ Retired
   - ❑ Unable to work

8. **Is your annual household income from all sources…**
   - ❑ < $24,999
   - ❑ $25,000 to $34,999
   - ❑ $35,000 to $49,999
   - ❑ $50,000 to $74,999
   - ❑ $75,000 to $99,999
   - ❑ $100,000 to $199,999
   - ❑ $200,000 or more

7.2 Recruitment and sampling

7.2.1 Participant screening questionnaire - Recruitment Pool 1

1. **Which of the following describes your knowledge or use of genetic testing?**
   - ❑ I have used a genetic testing company (eg. 23andMe, Color, Ancestry, Invitae, Helix, etc.). ← Disqualify
   - ❑ I have taken or undergone genetic testing through my healthcare provider. ← Disqualify
   - ❑ I have taken part in a study where I reviewed genetic test reports. ← Disqualify
   - ❑ None of the above. ← Qualify

2. **Which of the following would you say is your race/ethnicity?**
   - ❑ White or Caucasian **= Disqualify**
   - ❑ Black or African American **= Qualify**
   - ❑ American Indian or Alaskan Native **= Qualify**
   - ❑ Asian **= Qualify**

- ❑ Pacific Islander **= Qualify**
- ❑ Hispanic, Latino/a or Spanish **= Qualify**
- ❑ Mixed race **= Qualify**

7.2.2 Participant screening questionnaire - Recruitment Pool 2

1. **Select the following qualifiers within respondent**
   - ❑ Some high school
   - ❑ High school graduate
   - ❑ Trade/Technical/Vocational training
   - ❑ Some college, no degree

2. **Which of the following describes your knowledge or use of genetic testing?**
   - ❑ I have used a genetic testing company (eg. 23andMe, Color, Ancestry, Invitae, Helix, etc.). ← Disqualify
   - ❑ I have taken or undergone genetic testing through my healthcare provider. ← Disqualify
   - ❑ I have taken part in a study where I reviewed genetic test reports. ← Disqualify
   - ❑ None of the above. ← Qualify

7.2.3 Participant screening questionnaire - Recruitment Pool 3

1. **Which of the following describes your knowledge or use of genetic testing?**
   - ❑ I have used a genetic testing company (eg. 23andMe, Color, Ancestry, Invitae, Helix, etc.). ← Disqualify
   - ❑ I have taken or undergone genetic testing through my healthcare provider. ← Disqualify
   - ❑ I have taken part in a study where I reviewed genetic test reports. ← Disqualify
   - ❑ None of the above. ← Qualify

2. **What is your age?**
   - ❑ 17 years or less **= Disqualify**
   - ❑ 18-24 years old **= Disqualify**
   - ❑ 25-34 years old **= Disqualify**
   - ❑ 35-44 years old **= Disqualify**
   - ❑ 45-54 years old **= Qualify**
   - ❑ 55-64 years old **= Qualify**
   - ❑ 65 years or older **= Qualify**

8 Survey items

8.1 Positive Hereditary Disease Risk Report

Instructions: The following statements are about the Hereditary Disease Risk Report you reviewed. Pick the best answer that reflects your understanding of the report.

1. Something significant for my health was found in the genes that were checked.
    - ❑ True
    - ❑ False
    - ❑ I don't know

2. My DNA test results tell me I definitely have cancer.
    - ❑ True
    - ❑ False
    - ❑ I don't know

3. My DNA test results tell me I will definitely get cancer in the future.
    - ❑ True
    - ❑ False
    - ❑ I don't know

4. If something significant for my health was found in my DNA, my relatives could have the same results.
    - ❑ True
    - ❑ False
    - ❑ I don't know

5. Doctors **only** consider DNA when making treatment decisions.
    - ❑ True
    - ❑ False
    - ❑ I don't know

6. I understand I could share my DNA test results with my doctor.
    - ❑ True
    - ❑ False
    - ❑ I don't know

7. I understand my DNA test result is a research result.
    - ❑ True
    - ❑ False
    - ❑ I don't know

8. I understand I should **not** change my medical care based on my DNA test results.
    - ❑ True
    - ❑ False
    - ❑ I don't know

9. I understand I could share my DNA test results with my family.
    - ❑ True
    - ❑ False
    - ❑ I don't know

8.2 Uninformative Hereditary Disease Risk Report

Instructions: The following statements are about the Hereditary Disease Risk Report you reviewed. Pick the best answer that reflects your understanding of the report.

1. My DNA test results do not eliminate my risk of developing a hereditary condition.
    - ❑ True
    - ❑ False
    - ❑ I don't know

2. Doctors **only** consider DNA when making treatment decisions.
    - ❑ True
    - ❑ False
    - ❑ I don't know

3. Aside from my DNA, other factors such as age and weight **may** affect my risk of developing diseases.
    - ❑ True
    - ❑ False
    - ❑ I don't know

4. I understand that nothing significant for my health was found in the genes that were checked.
    - ❑ True
    - ❑ False
    - ❑ I don't know

5. I understand my DNA test result is a research result.
    - ❑ True
    - ❑ False
    - ❑ I don't know

6. I understand I should **not** change my medical care based on my DNA test results.
    - ❑ True
    - ❑ False
    - ❑ I don't know

8.3 Medicine and Your DNA Report

Instructions: The following statements are about the Medicine and Your DNA Report you reviewed. Pick the best answer that reflects your understanding of the report.

1. Doctors and pharmacists **only** consider DNA when making treatment decisions.
    - ❑ True
    - ❑ False
    - ❑ I don't know

2. Aside from my DNA, other factors such as age and weight **may** impact my response to medicines.
    - ❑ True
    - ❑ False
    - ❑ I don't know

3. My DNA **may** impact how I respond to certain medicines.
    a. True
    b. False
    c. I don't know

4. My DNA test results tell me which medicines **will definitely** work for me.
    - ❑ True
    - ❑ False
    - ❑ I don't know

5. Based on my DNA test results, I can stop taking my prescribed medicines before talking to my doctor.
    - ❑ True
    - ❑ False
    - ❑ I don't know

6. Based on my DNA test results, I can change how often I take my prescribed medicines before talking to my doctor.
    - ❑ True
    - ❑ False
    - ❑ I don't know

7. I understand I could share my DNA test results with my doctor.
    - ❑ True
    - ❑ False
    - ❑ I don't know

8. I understand my DNA test result is a research result.
    - ❑ True
    - ❑ False
    - ❑ I don't know

9. I understand I should **not** change my medical care based on my DNA test results.
    - ❑ True
    - ❑ False
    - ❑ I don't know

		Education Level					
		Total N = 205	Grade 12 or GED (High school graduate) N = 19	Some College, no degree N = 71	Associate degree N = 15	Bachelor's degree N = 66	Graduate or professional degree N = 34
Survey Question	Comprehension rate	**N (%)**	N (%)	N (%)	N (%)	N (%)	N (%)
1. Doctors and pharmacists only consider DNA when making treatment decisions.	Pass	**199 (97.1)**	17 (89.5)	69 (97.2)	15 (100.0)	65 (98.5)	33 (97.1)
	Fail	6 (2.9)	2 (10.5)	2 (10.5)	0 (0.0%)	1 (1.5)	1 (2.9)
2. Aside from my DNA, other factors such as age and weight may impact my response to medicines.	Pass	**205 (100)**	19 (100)	71 (100)	15 (100)	66 (100)	34 (100)
	Fail	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
3. My DNA may impact how I respond to certain medicines.	Pass	**203 (99.0)**	19 (100)	70 (98.6)	15 (100)	65 (98.5)	34 (100)
	Fail	2 (1.0)	0 (0.0)	1 (1.4)	0 (0.0)	1 (1.5)	0 (0.0)
4. My DNA test results tell me which medicines will definitely work for me.	Pass	**193 (94.1)**	15 (78.9)	68 (95.8)	15 (100)	63 (95.5)	32 (94.1)
	Fail	12 (5.9)	4 (21.1)	3 (4.2)	0 (0.0)	3 (4.5)	2 (5.9)
5. Based on my DNA test results, I can stop taking my prescribed medicines before talking to my doctor.	Pass	**205 (100)**	19 (100)	71 (100)	15 (100)	66 (100)	34 (100)
	Fail	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
6. Based on my DNA test results, I can change how often I take my prescribed medicines before talking to my doctor.	Pass	**205 (100)**	19 (100)	70 (98.6)	15 (100)	66 (100)	34 (100)
	Fail	0 (0.0)	0 (0.0)	1 (1.4)	0 (0.0)	0 (0.0)	0 (0.0)
7. I understand I could share my DNA test results with my doctor.	Pass	**205 (100)**	19 (100)	71 (100)	15 (100)	66 (100)	34 (100)
	Fail	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
8. I understand my DNA test result is a research result.	Pass	**198 (96.6)**	19 (100)	67 (94.4)	14 (93.3)	65 (98.5)	33 (97.1)
	Fail	7 (3.4)	0 (0.0)	4 (5.6)	1 (6.7)	1 (1.5)	1 (2.9)
9. I understand I should not change my medical care based on my DNA test results.	Pass	**196 (95.6)**	17 (89.5)	68 (95.8)	15 (100)	62 (93.9)	34 (100)
	Fail	7 (3.4)	2 (10.5)	3 (4.22)	0 (0.0)	4 (6.1)	0 (0.0)

		Education Level					
		Total N = 342	Grade 12 or GED (High school graduate) N = 18	Some College, no degree N = 102	Associate degree N = 30	Bachelor's degree N = 133	Graduate or professional degree N = 59
Survey Question	Comprehension rate	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)
1. Something significant for my health was found in the genes that were checked.	Pass	**326 (95.3)**	17 (94.4)	95 (93.1)	28 (93.3)	129 (97.0)	57 (96.6)
	Fail	16 (4.7)	1 (5.6)	7 (6.9)	2 (6.7)	4 (3.0)	2 (3.4)
2. My DNA test results tell me I definitely have cancer.	Pass	**335 (98.0)**	18 (100.0)	99 (97.1)	29 (96.7)	130 (97.0)	59 (100.0)
	Fail	7 (2.0)	0 (0.0)	3 (2.9)	1 (3.3)	3 (2.3)	0 (0.0)
3. My DNA test results tell me I will definitely get cancer in the future.	Pass	**335 (98.0)**	17 (94.4)	100 (98.0)	29 (96.7)	130 (97.0)	59 (100.0)
	Fail	7 (2.0)	1 (5.6)	2 (2.0)	1 (3.3)	3 (2.3)	0 (0.0)
4. If something significant for my health was found in my DNA, my relatives could have the same results.	Pass	**337 (98.5)**	17 (94.4)	101 (99.0)	30 (100.0)	130 (97.0)	59 (100.0)
	Fail	5 (1.5)	1 (5.6)	1 (1.0)	0 (0.0)	3 (2.3)	0 (0.0)
5. Doctors only consider DNA when making treatment decisions.	Pass	**320 (93.6)**	16 (88.9)	94 (92.2)	26 (86.7)	126 (97.0)	58 (98.3)
	Fail	22 (6.4)	2 (11.1)	8 (7.8)	4 (13.3)	7 (5.3)	1 (1.7)
6. I understand I could share my DNA test results with my doctor.	Pass	**341 (99.7)**	18 (100.0)	101 (99.0)	30 (100.0)	133 (97.0)	59 (100.0)
	Fail	1 (0.3)	0 (0.0)	1 (1.0)	0 (0.0)	0 (0.0)	0 (0.0)
7. I understand my DNA test result is a research result.	Pass	**333 (97.4)**	17 (94.4)	99 (97.1)	28 (93.3)	132 (97.0)	57 (96.6)
	Fail	8 (2.3)	1 (5.6)	3 (2.9)	1 (3.3)	1 (0.8)	2 (3.4)
8. I understand I should not change my medical care based on my DNA test results.	Pass	**323 (94.4)**	18(100.0)	98(96.1)	29(96.7)	122(97.0)	56(94.9)
	Fail	19 (5.6)	0(0.0)	4(3.9)	1(3.3)	11(8.3)	3(5.1)
9. I understand I could share my DNA test results with my family.	Pass	**340 (99.4)**	18(100.0)	102(100.0)	30(100.0)	131(97.0)	59(100.0)
	Fail	2 (0.6)	0(0.0)	0(0.0)	0(0.0)	2(1.5)	0(0.0)

		Education Level					
		Total N = 278	Grade 12 or GED (High school graduate) N = 37	Some College, no degree N = 102	Associate degree N = 23	Bachelor's degree N = 79	Graduate or professional degree N = 37
Survey Question	Comprehension rate	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)
1. My DNA test results do not eliminate my risk of developing a hereditary condition.	Pass	**270 (97.1)**	35 (94.6)	97 (95.1)	23 (100)	79 (100)	36 (97.3)
	Fail	8 (2.9)	2 (5.4)	5 (4.9)	0 (0.0)	0 (0.0)	1 (2.7)
2. Doctors only consider DNA when making treatment decisions.	Pass	**257 (92.4)**	29 (78.4)	94 (92.2)	22 (95.7)	76 (96.2)	36 (97.3)
	Fail	21 (7.6)	8 (21.6)	8 (7.8)	1 (4.3)	3 (3.8)	1 (2.7)
3. Aside from my DNA, other factors such as age and weight may affect my risk of developing diseases.	Pass	**262 (94.2)**	35 (94.6)	96 (94.1)	23 (100)	74 (93.7)	34 (91.9)
	Fail	16 (5.6)	2 (5.4)	6 (5.9)	0 (0.0)	5 (6.3)	3 (8.1)
4. I understand that nothing significant for my health was found in the genes that were checked.	Pass	**269 (96.8)**	35 (94.6)	99 (97.1)	23 (100)	78 (98.7)	34 (91.9)
	Fail	9 (3.2)	2 (5.4)	3 (2.9)	0 (0.0)	1 (1.3)	3 (8.1)
5. I understand my DNA test result is a research result.	Pass	**276 (99.3)**	37 (100)	102 (100)	23 (100)	77 (97.5)	37 (100)
	Fail	2 (0.7)	0 (0.0)	0 (0.0)	0 (0.0)	2 (2.5)	0 (0.0)
6. I understand I should not change my medical care based on my DNA test results.	Pass	**277 (99.6)**	37 (100)	101 (99.0)	23 (100)	79 (100)	37 (100)
	Fail	1 (0.4)	0 (0.0)	1 (1.0)	0 (0.0)	0 (0.0)	0 (0.0)

All of Us Research Program: Investigator Agreement

FDA Investigational Device Exemption Study – Return of Genetic Results in the *All of Us* Research Program

Sponsor: National Institutes of Health *All of Us* Research Program

This Investigator Agreement provides acknowledgement of the signatory of his/her responsibilities as a co-Responsible Investigator in the referenced study, per requirements specified by 21 CFR 812.43.

Instructions in italics.

_____

Name / Title / Institution

_____

a. Curriculum Vitae
NIH Biosketch may substitute for items a and b.

_____

b. Relevant experience
NIH Biosketch may substitute for items a and b.

_____

c. Information on any terminated studies you were involved in, including an explanation of the circumstances that led to termination

_____

d. Investigator statement: I certify that I will conduct the study in accordance with this agreement, the Investigational Plan (IDE study description), the IDE and other applicable FDA regulations, and conditions of approval imposed by the reviewing IRB or FDA; supervise all testing involving human subjects; and ensure requirements for informed consent are met, as applicable.

---

e. Financial disclosure: I certify that based on information obtained from the sponsor or from other participating investigators, the listed clinical investigators (list of names contained in email request) did not participate in any financial arrangement with the sponsor of a covered study whereby the value of compensation to the investigator for conducting the study could be affected by the outcome of the study (as defined in 21 CFR 54.2(a)); had no proprietary interest in this product or significant equity interest in the sponsor of the covered study (as defined in 21 CFR 54.2(b)); and was not the recipient of significant payments of other sorts (as defined in 21 CFR 54.2(f)). I affirm that I will notify the Sponsor if new financial conflict should become relevant, for a time period to extend to one year after the conclusion of the study.

Note that the information provided will not be submitted in the IDE application to the FDA. Its collection by the Sponsor is required for submission in any marketing application involving the device. AoURP has no intention to market or license the device described in the study.

Signature _____

Printed Name

Date _____

Consent to Join the *All of Us* Research Program

Principal Investigator: Paul Harris, PhD
Vanderbilt University Medical Center
2525 West End Ave, Suite 1500
Nashville, TN 37203

Sponsor: National Institutes of Health

This form is for people age 18 or older.

This form tells you about the *All of Us* Research Program (*All of Us*). It explains what we will ask you to do if you join. Please read this form carefully. If you have questions, there is a list of people you can ask at the end of this form. We will give you a copy of this form.

What is *All of Us*?

All of Us is a health research program funded by the U.S. government. If you join, we will gather data about you. We will combine it with data from other people who join. Researchers will use this data for lots of studies. By looking for patterns, researchers may learn more about what affects people's health.

We hope that 1 million people or more will join *All of Us*.

How long will *All of Us* last?

All of Us will last for at least ten years. If you decide to join, you can withdraw (quit) at any time.

What will you ask me to do?

If you decide to join *All of Us*, we will gather data about you. We will gather some of the data from you directly. We will gather some of the data from elsewhere.

Data that we will gather from you directly:

1. Basic data

   We will ask you for data like your name and date of birth. Once a year we may ask if you have moved or changed your phone number or email address. We will ask for the phone number or email address for a friend or family member as a backup in case we need to reach you. We might use social media or public listings to help us keep your contact information up to date.

2. Health data

   We will ask you questions about your health, family, home, and work. This will take about 30-60 minutes. From time to time, we will ask you follow-up questions about your health. We may ask the same question more than one time, so we can see if there are any changes. You do not have to answer any questions you do not want to answer.

3. Physical measurements

   We may ask you to go to a local clinic to be measured. If you are asked, you can decide yes or no. You can say no and still take part in *All of Us.*

   If you say yes to being measured, it will take about 15 minutes. *All of Us* trained research staff will do the measurements. We will measure your height, weight, hips, and waist. We will check your blood pressure and heart rate. We may ask you to have other measurements over time.

4. Samples

   We may ask you to go to a local clinic to give a blood sample. If you are asked, you can decide yes or no. You can say no and still take part in *All of Us.*

   If you say yes to giving a sample, we will use a needle to draw about 3 tablespoons of blood from your arm. We may ask you to give a urine sample (“pee in a cup”). We may ask for other samples, like saliva (“spit”).

   We will store your blood, urine, and saliva samples in the *All of Us* biobank. The biobank is a secure storage place for samples. We will store your samples at the biobank until they are used up by researchers

for different studies. We may ask you to give more samples in the future. You can say yes or no.

5. Fitness trackers

   If you have a fitness tracker (like one to count the steps you take in a day), you might be able to share data from it with *All of Us*. If you don't have a fitness tracker, we may ask you to use one that we give you. You can say no and still take part in *All of Us.*

Data that we will gather from elsewhere:

1. Electronic health records

   If you have electronic health records, we may ask for access. If you are asked, you can decide yes or no. You can say no and still take part in *All of Us,* but it might limit what other data we ask to collect from you. For example, if you say no, you might not be asked to give samples.

   There will be a separate form called the HIPAA Authorization for you to sign if you decide to give us access. We will see data about your health problems, test results, medical procedures, images (such as X-rays), and medicines you take. Health records can contain sensitive data. For example, they may tell us about your mental health, genetic conditions, or use of alcohol or drugs. They may contain sexual or infection data, including HIV status.

2. Data about your health from other sources

   We will add data from other sources to the data you give us. For example, environmental data and pharmacy records. This will give researchers more data about factors that might affect your health.

   There are two ways we will add data from other sources to your *All of Us* record:

   - *Based on where you live and work*

     We will add data about your area based on where you live and work. For example, we may add data about the number of people in your area. We may add pollution data. We may add data like how close you live to the nearest grocery store or park.

   - *Based on data that identifies you*

We will use data that identifies you like your name and date of birth to add data that is specific to you. For example, we may add data from pharmacy records or health insurance records. If you have had cancer, we may add data from cancer registries.

If you have a social security number, we may ask you for it to help with adding data. It is optional. Even if we ask, you do not have to give us your social security number. You can say no and still take part in *All of Us.*

These other sources can contain sensitive data. For example, they may tell us about your mental health, or use of alcohol or drugs. They may contain sexual or infection data, including HIV status. Because of this, we will ask the *All of Us* ethics committee to review and approve each data source before we add it.

What will you do with my data and samples?

We will store your data and samples securely, along with the data and samples from all the other people who take part in *All of Us*. Researchers will use the data and samples to make discoveries.

1. <u>We will study your samples, including your DNA</u>

   We may measure things that naturally occur in our bodies, like cholesterol. We may look for signs of outside factors that affect health. For example, we may look for environmental toxins, medicines, or drugs.

   We will also study your DNA. DNA is in your blood and other samples.

   All human beings share more than 99% of their DNA with each other. The tiny bit that is different is part of what makes each of us unique. Things like our hair color and eye color depend on the bits of DNA that are different between human beings. We call these our DNA changes. These DNA changes can also tell you about your health and how your body works. They can tell you about where your ancestors may be from. We are still learning about what role DNA plays in many parts of our lives.

   DNA is passed from parents to kids. Half of your DNA came from your mom and half came from your dad. If you have kids, each of them will

get half your DNA. In this way, your DNA also tells you about your family.

We will use many methods to study your samples. For example, we might study your DNA using whole genome sequencing. Whole genome sequencing is a way of studying nearly all of a person's DNA. Every person's whole genome sequence is different. It is unique to them, like a fingerprint.

Because *All of Us* will last for ten or more years, some of the methods we will use may not even be invented yet.

2. <u>We will create a public database on the *All of Us* website</u>

   The data in the public database will be about the group. It will not include data about individual people. It will not include your name or other data that directly identifies you. Everyone will be able to use the public database.

3. <u>We will create a scientific database</u>

   The scientific database will have individual-level data and samples. This includes your DNA data. Access to this database will be controlled. Researchers will have to be approved by *All of Us* to use this database. They will have to have special training before they can be approved. Their research may be on nearly any topic. They may look for patterns in DNA. This may help them discover different ways that DNA affects people. These researchers may be from anywhere in the world. They may work for commercial companies, like drug companies. They may be citizen or community scientists. Citizen and community scientists are people who do science in their spare time.

4. <u>Researchers can also ask to study your samples or DNA directly</u>

   We may send them a small amount of your samples or DNA so that they can do this. Before we send researchers your samples or DNA, they will have to take special training and sign a contract stating that they will not try to find out who you are. They will have to tell us what they want to study. *All of Us* will have to approve it.

   Researchers will use many methods to study your samples and DNA. Because *All of Us* will last for ten or more years, some of the methods

may not even be invented yet. The data researchers get from studying your samples and DNA may be added to the *All of Us* scientific database.

You can learn more about the research being done at www.joinallofus.org.

Except if you withdraw ("quit") or there are limits imposed by law, there is no limit on the length of time we will store your samples and data. Researchers will use your samples and data for research long into the future.

What else will you ask me to do?

We may ask if you want to hear about chances to take part in other studies. You can say yes or no to taking part in other studies. You can say no and still take part in *All of Us*.

What are the risks of taking part in *All of Us*?

The main risk of taking part in *All of Us* is to your privacy. A data breach is when someone sees or uses data without permission. If there is a data breach, someone could see or use the data we have about you. Even without your name, there is a chance someone could figure out who you are. They could misuse your data. We believe the chance of this is very small, but it is not zero.

We will gather data from you through the *All of Us* app and/or website. You may be asked to wear a fitness tracker. There is a risk to your privacy whenever you use an app, website, or fitness tracker. In general, there is no additional risk to your privacy if you use any of them as part of *All of Us.* That said, we will be gathering many different types of data in your *All of Us* record. If there is a data breach, there may be additional risk to your privacy because of the amount of data in your *All of Us* record.

Researchers will use basic facts like your race, ethnic group, and sex in their studies. This data helps researchers learn if the things that affect health are the same in different groups of people. These studies could one day help people of the same race, ethnic group, or sex as you. However, there is a risk that others could use this data to support harmful ideas about

groups.

If you give a blood sample, the most common risks are brief pain and bruising. Some people may become dizzy or feel faint. There is also a small risk of infection.

Taking part in *All of Us* may have risks that we don't know about yet. We will tell you if we learn anything that might change your decision to take part.

What are the risks of letting you use my DNA for research?

Your DNA is a type of private information. It is unique to you.

If there is a data breach, someone could see or use your DNA information without permission. There is a very small chance they could figure out who you are. They could try to use information about your DNA against you. It could impact your employment, insurance, or family relationships.

There are federal laws that can help protect your privacy. Some of these laws say that employers can't treat people differently because of their DNA. These laws do not apply to employers with fewer than 15 employees. These laws also say that health insurers can't use DNA information to change your coverage, drop you, or charge you more.

What will you do to protect my privacy?

Your privacy is very important to us. We will take great care to protect it. Here are a few of the steps we will take:

- Data we have about you will be stored on protected computers. We will limit and keep track of who can see this data.
- We will limit who is allowed to see information that could directly identify you, like your name or social security number.
- In order to work with your health data researchers must sign a contract stating they will not try to find out who you are.
- We will tell you if there is a data breach.

- *All of Us* has Certificates of Confidentiality from the U.S. government. These will help us fight legal demands (such as a subpoena or a request from federal, state, or local law enforcement) to give out information that could identify you.

All of Us will only use information about DNA changes for research. We will not tell insurance companies about who has DNA changes. We will not tell employers. We will not tell banks. We will not tell any school, college, or university.

Will you ever give out my name or other information that identifies me?

There are a few times when we might need to give out your name or other information that identifies you.

- We will give out information about you to protect your health or the health of others
    - If we learn or suspect that you are being abused.
    - If we learn or suspect you are abusing, neglecting, or have abandoned someone who depends on you for care, like a child or dependent adult.
    - If we learn that you plan to harm yourself or someone else.
    - If we learn that you have a disease that is a risk to public health, like measles.
- We will give out any data needed to meet U.S. laws and regulations. This may include information that identifies you. For example, there is a regulation that says the Food and Drug Administration (FDA) may ask to look at the records for the *All of Us* Research Program. The FDA checks how programs like *All of Us* give people DNA results about their health. If the FDA asks to look at these records to do their checks, we will let them.

Once your information is shared with *All of Us*, it may no longer be protected by patient privacy rules (like HIPAA). However, it will still be protected by other privacy rules. These include the rules that researchers

must follow to access the *All of Us* scientific database.

Are there any benefits?

All of Us is not medical treatment. It is a research program. You will not get direct medical benefits from taking part in *All of Us*.

That said, you may indirectly benefit from taking part in *All of Us*. For example, we will provide ways for you to get access to all the data you share with us and some of the results about you. This information may be interesting to you. You may learn about your health. You may learn about your DNA changes. You will be able to share your *All of Us* information with your healthcare provider if you choose. You will have the option to learn about additional study opportunities. Finally, you will be helping researchers make discoveries that may help future generations.

Are there any costs?

There are no direct costs to taking part in the *All of Us* Research Program.

However, we will do various medical tests as part of this study. We will give you the results. You can decide to seek follow-up care on your own because of these results. If you receive follow-up care, your doctor will bill you or your insurance company per usual practice. If you do not have insurance, or if your insurance will not pay, you will be responsible for the cost of follow-up care.

Are there any payments?

If we ask you to be physically measured and give samples and you decide to do it, we will offer you a one-time payment of $25.

Researchers will use your data to make discoveries. If any of their studies lead to new tests, drugs, or other commercial products, you will not get any profits. These inventions will be the property of the researchers who develop them.

Will I be able to see my data?

Yes, you will be able to see some of the data we collect about you. This includes:

- Any data you give us, like your health data.
- Your physical measurements.
- Some measurements from your samples. You can choose to see any of the measurements from studies *All of Us* does on your samples, like your whole genome sequence. You may not be able to see all the measurements from other studies researchers do using your samples.

You will be able to share this data if you choose. For example, you might want to share your *All of Us* data with your family or your healthcare provider.

Will I find out the results of the research?

Results explain or interpret data. *All of Us* involves two kinds of results: results about you and results about the group.

1. Results about you

   Over the many years of the *All of Us* Research Program, we will study lots of things about your data and samples. We will tell you if there are results about you from what *All of Us* studies. You will be able to choose if you want to see these results.

   Sometimes, we will ask you if you want us to check your data or samples for results that you might find interesting. For example, we may ask you to fill out another form where you can choose if you want us to check your DNA for certain kinds of DNA changes and return your results to you. This form is called the Consent to Receive DNA Results. It will tell you about the risks and benefits of having us check your DNA and about learning your results. We will not check for these kinds of DNA changes until you make a decision.

   Some of the results we give you may tell you about your health and others may not.

- Results that might tell you about your health

  These are results that could be used by a healthcare provider to take better care of you. For example, if any of your physical measurements are outside of what we would expect, we will tell you so you can follow-up with your healthcare provider. You will have to pay for the cost of follow-up care with your own healthcare provider.

- Results that would not tell you about your health

  These results might be interesting to you, but a healthcare provider probably would not use them to take better care of you. For example, these results might come from tests that are still experimental.

2. Results about the group

   These are reports of what researchers learn about health from studying data and samples from all the different people in the *All of Us* Research Program. You can get these reports, as well as general news and updates about *All of Us* at www.joinallofus.org.

   While researchers might learn results about you from studying your *All of Us* data and samples, you may not be able to see these results.

What if I get injured?

If you think you have been injured because of taking part in *All of Us*, contact us using the information at the end of this form. If we find that you were injured as a direct result of taking part in *All of Us*:

- You will not have to pay for any immediate medical care to treat your injury.
- Beyond your immediate medical care, we will not pay for your injury.
- If you need follow-up care to treat your injury, you and/or your insurance will have to pay for it.
- If you have any long-term costs to treat your injury, you and/or your insurance will have to pay them.
- You do not give up any of your legal rights if you take part in *All of Us.*

Do I have to take part?

Taking part in *All of Us* is voluntary. You can choose to join or not. No matter what you decide, now or in the future, it will not affect your medical care.

If you decide to join *All of Us*, you can change your mind at any time. If you decide you want to withdraw (quit), you need to tell us. You can tell us through the app or website, or use the contact information at the end of this form to call or write to us.

If you withdraw, your samples will be destroyed. Your data will not be used for new studies. However, if researchers already have your data or samples for their studies, we at *All of Us* cannot get it back. Also, we will let researchers check the results of past studies. If they need your old data to do this work, we will give it to them.

Even if you withdraw, we will keep your name and contact information. We keep this information so we can follow U.S. research laws and regulations.

Who can answer my questions?

If you have questions:	***Please contact the:***
About the *All of Us* Research Program	***All of Us* Support Center** **Hours**: Mon-Sun, 7am-10pm ET **Phone**: 1-844-842-2855 **Email**: help@joinallofus.org **Chat** (website or app): www.joinallofus.org **Languages**: English and Spanish
About your rights as a research participant	***All of Us* Institutional Review Board** **Phone**: 1-844-200-8990 **Email**: AoUIRBContact@emmes.com **Address**: 401 N. Washington Street, 7th Floor, Rockville, MD 20850

I know and agree that:

- My data will be stored in the *All of Us* databases.
- If I give a blood, urine, or saliva sample, it will be stored at the *All of Us* biobank. This includes my DNA. Information that researchers learn by studying my samples will be stored in the *All of Us* databases.
- Researchers will do studies using the *All of Us* databases and biobank. Their research may be on nearly any topic.
- I may be asked to give more samples in the future. I can say yes or no.
- My contact information may be used to tell me about other studies.
- I can withdraw (quit) at any time. There is no penalty if I withdraw.

Please check the box below if you agree to take part:

☐ **I have read this consent form (or someone read it to me). I understand the information in this form. All of my questions have been answered. I freely and willingly choose to take part in the *All of Us* Research Program.**

Sign Your Full Name:

_____

Date:

_____

Please check the box below if someone from *All of Us* helped you with completing the consent process:

☐ **I received help from *All of Us* to complete the consent process.**

Name of the person who helped you:

_____

All of Us Research Program Consent to Receive DNA Results

Person in charge of this study:
Paul Harris, PhD
Vanderbilt University Medical Center
2525 West End Ave, Suite 1500
Nashville, TN 37203

Sponsor: National Institutes of Health

This form is for people age 18 or older.

When you joined the program, you signed a form that says you agree to let scientists study your DNA for research. Now we want to know if you want *All of Us* to check your DNA for changes and tell you what we find. These are your DNA results. This form explains the choice you have about learning your DNA results. It explains about different kinds of DNA changes. It explains how we will check for them and how long it will take. It explains how you can learn about your results.

Some people will want to learn about their DNA results. Other people will not. We want you to make the best decision for yourself. No matter what you decide, you can still participate in *All of Us*.

Please read this form carefully. If you have questions, there is a list of people you can ask at the end of this form. We will give you a copy of this form if you want one.

What are “DNA changes”?

All human beings share more than 99% of their DNA with each other. The tiny bit that is different is part of what makes each of us unique. Things like our hair color and eye color depend on the bits of our DNA that are different between human beings. We call these our DNA changes. We know what some DNA changes mean, but we still have a lot to learn. For example, we

are still learning what role DNA plays in most health conditions. In fact, that's one of the reasons we are doing the *All of Us* Research Program. But for a small number of things we already know a lot about the role DNA plays.

We know that certain changes in our DNA can affect our health. For example:

- Certain DNA changes can increase our risk for a few specific health conditions. This could include some cancers and types of heart disease.
- Certain changes in our DNA can increase the risk of passing specific health conditions onto our children, even if we don't have those conditions.
- Certain changes in our DNA can impact how a few specific medicines work.

We also know that other changes in our DNA can tell us about things like:

- Where our ancestors may be from.
- How our bodies work.

The more we study our DNA, the more we will learn what DNA changes mean about us.

How will you check my DNA?

If you say yes, we will check your DNA for certain types of DNA changes. We will do this by having a specially trained scientist look at your DNA.

How long will it take to get my results?

It might take a few months or even a few years for *All of Us* to check your DNA. We may be able to give you results about some types of DNA changes sooner than others. You will be able to follow our progress through your *All of Us* account.

What exactly will you check for?

If you say yes, you are telling us that you want to learn about some or all of your DNA results. When we are ready to check for a specific type of DNA change, we will tell you more about what the results may mean for you. Then you can decide whether you want your results for that type of DNA change. For example, you may want to learn about any health-related DNA changes you have. You might only want to learn results about where your ancestors may be from. Or you might want to learn about all of your DNA results. You get to choose.

The list of DNA changes that we will check for may change as researchers make new discoveries. There will be a link to the most updated list of what we check for in your *All of Us* account.

Over time, we may learn new information that could change your results. We may go back and check your DNA again. We will tell you if we find anything new or if we find anything that changes your results.

What will you tell me?

Health-related DNA results

If we find a change in your DNA that increases your risk of a health condition, we will try to contact you directly. We will help you make an appointment with an *All of Us* genetic counselor. They will tell you your results. They will answer your questions. They will send you a report. If you want, they will help you find a healthcare provider in your area.

If we find other types of health-related results, we will send you a note through your *All of Us* account. We will give you a report that you can share with your healthcare provider.

If we do not find any known health-related changes in your DNA, we will also send you a note through your *All of Us* account.

Remember, ***All of Us* is a research program. It is not medical care.** Do not use *All of Us* as a substitute for medical care.

Results about your health-related DNA changes from *All of Us* are

research results. Research results are not the same as medical testing. Do not make any changes to your medicines or care until you talk to a healthcare provider. They will need to confirm your *All of Us* results with medical tests. If you don't have a healthcare provider, we can help you find one.

Other types of DNA results

DNA can also give you information that is not about your health. DNA can tell you where your ancestors may be from. It can help explain how your body works. You can decide if you want us to check for these kinds of DNA changes. If you say yes, we will send you a note through your *All of Us* account once we have your results. In the note, we will tell you how to see your DNA results using our interactive learning tool.

Do I need to pay to get my DNA results?

You do not need to pay to have us check for DNA changes. You do not need to pay to find out your results. You do not need to pay to talk to an *All of Us* genetic counselor.

How could learning my DNA results help me?

Knowing you have a health-related DNA change may help your healthcare provider take better care of you. They may be able to prevent or find a health condition early. This could help you get better treatment. They may be able to adjust the type or amount of medicine they give you. You might have fewer side effects.

If you have a health-related DNA change, your blood relatives might have it too. If you tell them about your DNA results, their healthcare providers might suggest different care for them.

You may think it is interesting to learn information from your DNA that is not about your health. For example, you may want to learn where your ancestors may be from.

Remember, *All of Us* is a research program. It is not medical care. Do not use *All of Us* as a substitute for medical care. Do not make any changes to

your medicines or care until you talk to a healthcare provider. They will need to confirm your *All of Us* results with medical tests.

What are the risks of learning my DNA results?

There are different kinds of risks to learning about your DNA. When we are ready to check for certain types of DNA changes, we will tell you more about the results you would receive. You may have questions. Genetic counselors can help answer any questions you may have. You can ask to talk to an *All of Us* genetic counselor through the *All of Us* Support Center. They can help you find support if you need it.

Healthcare and Insurance Risks

Changes in healthcare

A healthcare provider must confirm your health-related DNA results by medical testing. You or your insurance may be billed for it. Your healthcare provider may recommend new or different care based on these medical tests. You may want to make some of these changes to your care and not others. You can decide what care is right for you. These changes in your care may cost more than regular care. Some of these changes in your care may not be covered by your insurance. Your healthcare provider may recommend care that makes you need to take time off of work, like surgery.

Disability, life, and long-term care insurance

In most places, DNA information can be used by disability, life, and long-term care insurers. These insurers can ask you if you have information about your DNA and you have to tell them what you know. They can use that information to decide if they will cover you and how much they charge. If you find out that you have a health-related DNA change from *All of Us*, it could make it difficult or more expensive to get these types of insurance.

In some places, there are laws that say life, disability, and long-term care insurers can't use DNA information to decide about your coverage. To find out if you are protected by these kinds of laws, contact the attorney general for your state or territory. You can learn how to contact your attorney general at www.usa.gov/state-attorney-general.

Family Risks

Your blood relatives

Your DNA tells about you and people who are related to you by blood. If you have a certain DNA change, your blood relatives might have it too. They may or may not want to know this information. You may realize you are not related to some family members in the way you thought you were.

Emotional Risks

If you decide to have your DNA checked, you may receive news that worries or scares you. You may be afraid of passing health-related DNA changes on to your children. Remember, you can ask to talk to an *All of Us* genetic counselor at any time. They can answer your questions and help you find support.

What are the risks of sharing my DNA results?

Sharing your DNA results could be a risk to your privacy. It could make it easier for someone to find out who you are in the *All of Us* scientific database and learn other information about you. They could misuse your information.

You can help protect your privacy. Be careful about posting about your DNA results on social media or in other public areas, as you may not be able to control how the information is used. If you decide to find out about your health-related DNA changes, only share those results with people you trust, like your healthcare provider.

What are the limits of *All of Us* DNA results?

The results you get will not tell you everything about your DNA. This is especially important to remember about health-related DNA results. For example, we will not check for every health-related DNA change. Even if we do not find a health-related DNA change, you could still have one.

The DNA results you get from *All of Us* are not the same as medical test results. Because DNA results from *All of Us* are research results, there is a

chance they could be incorrect. Do not make changes to your medicines or care based on your *All of Us* DNA results.

Are there ways that DNA results cannot be used?

Employment

There are federal laws that say employers can't treat people differently because of their DNA information. These laws do not apply to employers with fewer than 15 people.

Health Insurance

There are federal laws that say health insurers can't use DNA information against people. They can't use it to change your coverage, cancel your coverage, or charge you more.

Will you ever give out my DNA results?

Protecting the privacy of your DNA results is very important to us.

Here are a few of the steps we will take to protect it:

- DNA results we have about you will be stored on protected computers. We will limit and keep track of who can see this information.
- *All of Us* has Certificates of Confidentiality from the U.S. government. These will help us fight legal demands (such as a subpoena or a request from federal, state, or local law enforcement) to give out information that could identify you.
- We will not tell insurance companies about your DNA results. We will not tell employers. We will not tell banks. We will not tell any school, college, or university.

We will give out any data needed to meet U.S. laws and regulations. This may include information that identifies you. There is a U.S. federal regulation that says the Food and Drug Administration (FDA) may ask to look at the records for the *All of Us* Research Program. The FDA checks

how programs like *All of Us* give people DNA results about their health. If the FDA asks to look at these records to do their checks, we will let them.

Do I have to learn my DNA results?

No, learning your DNA results is voluntary. Some people will want to find out. Other people will not. Some people may not be sure right now. It is up to you. No matter what you decide, you can still participate in *All of Us*.

You can change your mind about learning your DNA results at any time. If you change your mind, you need to tell us. You can tell us through the app or website or use the contact information at the end of this form to call or write to us.

When will my consent expire?

Unless you tell us that you have changed your mind, your consent lasts until December 31, 2099.

This form is only for learning about DNA changes. It cannot be used for any other purpose.

Who can answer my questions?

If you have questions:	***Please contact the:***
About the *All of Us* Research Program	***All of Us* Support Center** **Hours**: Mon-Sun, 7am-10pm ET **Phone**: 1-844-842-2855 **Email**: help@joinallofus.org **Chat** (website or app): www.joinallofus.org **Languages**: English and Spanish
About your rights as a research participant	***All of Us* Institutional Review Board** **Phone**: 1-844-200-8990 **Email**: AoUIRBContact@emmes.com **Address**: 401 N. Washington Street, 7th Floor, Rockville, MD 20850

Would you like to learn any of your DNA results?

☐ **No, I do <u>not want</u> to learn about any DNA results.**

- I know I can change my mind later.
- I know this means that researchers can still use my DNA to make discoveries unless I withdraw (quit).

☐ **I'm not sure right now.**

- I know that until I decide, I **will not** learn about any of my DNA results.
- I know I can change my mind later.
- I know this means that researchers can still use my DNA to make discoveries unless I withdraw (quit).

☐ **Yes, I <u>want</u> to learn some or all of my DNA results.**

- I know *All of Us* will ask me later what specific types of DNA results I want. I get to choose.
- I know this means *All of Us* will tell me the kinds of results I choose to learn.
- I know this means I have to keep my contact information in *All of Us* up-to-date so that you can give me my results.
- I know this means that researchers can still use my DNA to make discoveries unless I withdraw (quit).

Sign Your
Full Name:

_____

Today's date:

_____

This document prepared for: _____

Please check the box below if someone from *All of Us* helped you with completing the consent process:

☐ **I received help from *All of Us* to complete the consent process.**

Name of the person who helped you:

_____