The GenoVA study: Equitable implementation of a pragmatic randomized trial of polygenic-risk scoring in primary care

Jason L. Vassy,[1,2,3,4,5,*] Charles A. Brunette,[1,3] Matthew S. Lebo,[3,6,7] Katharine MacIsaac,[1] Thomas Yi,[1] Morgan E. Danowski,[1] Nicholas V.J. Alexander,[1,8,9] Mark P. Cardellino,[1] Kurt D. Christensen,[3,10] Manish Gala,[3,11] Robert C. Green,[3,4,5,12] Elizabeth Harris,[1] Natalie E. Jones,[1,3] Benjamin J. Kerman,[2,3] Peter Kraft,[13] Preetika Kulkarni,[1] Anna C.F. Lewis,[12,14] Steven A. Lubitz,[15,16] Pradeep Natarajan,[3,4,17,18] and Ashley A. Antwi[1]

Summary

Polygenic risk scores (PRSs) hold promise for disease risk assessment and prevention. The Genomic Medicine at Veterans Affairs (GenoVA) Study is addressing three main challenges to the clinical implementation of PRSs in preventive care: defining and determining their clinical utility, implementing them in time-constrained primary care settings, and countering their potential to exacerbate healthcare disparities. The study processes used to test patients, report their PRS results to them and their primary care providers (PCPs), and promote the use of those results in clinical decision-making are modeled on common practices in primary care. The following diseases were chosen for their prevalence and familiarity to PCPs: coronary artery disease; type 2 diabetes; atrial fibrillation; and breast, colorectal, and prostate cancers. A randomized clinical trial (RCT) design and primary outcome of time-to-new-diagnosis of a target disease bring methodological rigor to the question of the clinical utility of PRS implementation. The study's pragmatic RCT design enhances its relevance to how PRS might reasonably be implemented in primary care. Steps the study has taken to promote health equity include the thoughtful handling of genetic ancestry in PRS construction and reporting and enhanced recruitment strategies to address underrepresentation in research participation. To date, enhanced recruitment efforts have been both necessary and successful: participants of underrepresented race and ethnicity groups have been less likely to enroll in the study than expected but ultimately achieved proportional representation through targeted efforts. The GenoVA Study experience to date offers insights for evaluating the clinical utility of equitable PRS implementation in adult primary care.

Introduction

A pressing question in genomics today is the clinical utility of polygenic risk scores (PRSs).[1–4] PRSs combine information from hundreds to millions of genetic loci, each with a very small association with the risk of common complex disease. The result is a continuous and quantitative measure of an individual's genetic susceptibility to conditions such as coronary artery disease and type 2 diabetes. Compared to rarer monogenic disease variants, PRSs might have greater transformative potential for public health and preventive medicine in their ability to identify larger proportions of the population at significantly elevated risk for disease, potentially facilitating evidence-based prevention and management.

Although the associations between PRSs and dozens of common diseases have been firmly established, at least three primary challenges impede the ability of PRSs to improve healthcare and health outcomes. First, how to define and determine the clinical utility of PRSs remains uncertain, although there is some consensus that prospectively collected patient outcomes data are needed to demonstrate their clinical utility and yet are lacking.[2,5,6] Second, because most preventive care is discussed and delivered in primary-care settings, PRS-based prevention strategies will need to be implemented within this time- and resource-constrained context. Third, despite increasingly large and more diverse discovery and validation cohorts and methodological improvements in trans-ancestry analysis,[7–10] concerns remain that PRS-based prediction models are less valid in underrepresented populations and that their clinical implementation might exacerbate existing healthcare disparities.[3,11]

Addressing these three overarching challenges to the evidence-based, equitable implementation of PRSs in

[1]VA Boston Healthcare System, Boston, MA, USA; [2]Division of General Internal Medicine and Primary Care, Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA; [3]Harvard Medical School, Boston, MA, USA; [4]Broad Institute of Massachusetts Institute of Technology and Harvard University, Cambridge, MA, USA; [5]Ariadne Labs, Boston, MA, USA; [6]Laboratory for Molecular Medicine, Mass General Brigham, Boston, MA, USA; [7]Department of Pathology, Brigham and Women's Hospital, Boston, MA, USA; [8]Bucharest University Emergency Hospital, Bucharest, Romania; [9]Bucharest University of Economic Studies, Bucharest, Romania; [10]Department of Population Medicine, Harvard Pilgrim Health Care Institute, Boston, MA, USA; [11]Division of Gastroenterology and Clinical and Translational Epidemiology Unit, Massachusetts General Hospital, Boston, MA, USA; [12]Department of Medicine (Genetics), Mass General Brigham, Boston, MA, USA; [13]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD, USA; [14]Edmond and Lily Safra Center for Ethics, Harvard University, Boston, MA, USA; [15]Demoulas Center for Cardiac Arrhythmias, Massachusetts General Hospital, Boston, MA, USA; [16]Novartis Institutes for BioMedical Research, Novartis, Basel, Basel-Stadt, Switzerland; [17]Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA; [18]Cardiovascular Research Center, Massachusetts General Hospital, Boston, MA, USA
*Correspondence: jvassy@bwh.harvard.edu
https://doi.org/10.1016/j.ajhg.2023.10.001.

preventive care, the Genomic Medicine at Veterans Affairs (GenoVA) Study is a pragmatic randomized controlled trial (RCT) of PRS testing and reporting for six common diseases screened for by primary-care providers (PCPs).[12] Here we describe how the design, processes, and lessons learned in the study illustrate potential solutions for the equitable implementation of PRSs and for informing their clinical utility in adult primary care.

Study overview and conceptual model

The goal of the GenoVA Study is to model how PRSs might be equitably integrated into the busy primary-care context while using a randomized trial design to rigorously compare the impact of PRS implementation versus usual care on patient outcomes. In the conceptual model for the study (adapted from Vassy 2018[13]), polygenic-risk information acts through both patients and providers to improve preventive health outcomes when linked to specific actionable recommendations, such as tailored screening strategies or targeted preventive therapy.

The trial protocol is registered on Clinicaltrials.gov (identifier NCT04331535) and is fully described in Note S1. The study is conducted at the VA Boston Healthcare System (VABHS), an integrated healthcare system within the U.S. Department of Veterans Affairs (VA) comprising eight facilities in eastern Massachusetts. Figure 1 illustrates the GenoVA Study processes from recruitment through results reporting. In brief, patients aged 50 to 70 years without known diagnoses of six target diseases (atrial fibrillation, coronary artery disease, type 2 diabetes, colorectal cancer, breast cancer, and prostate cancer; see full eligibility criteria in Table S1) are recruited and complete genotyping and a baseline survey on enrollment (Note S2). Genotyping categorizes each participant into one of three groups: participants with an actionable monogenic disease result, as defined by the American College of Medicine Genetics and Genomics (ACMG),[14] participants with at least one high-risk PRS result, and participants with no high-risk PRS results, a group we term "average risk." Participants with an ACMG finding receive their monogenic and PRS results from a genetic counselor. All other participants undergo randomization to the PRS intervention or usual-care arm and are stratified by sex and either high-risk or average-risk status. The PRS intervention consists of a PRS laboratory report, targeted genetic counseling for high-risk individuals, communication with each participant's PCP, and patient- and provider-oriented materials to support decision-making around high-risk PRS results. End-of-study data collection from the EHR and surveys (Note S3) occurs 24 months after randomization, after which participants in the usual-care arm receive their results. The primary outcome compares the time to new diagnosis of the six target diseases for high-risk participants randomized to the PRS intervention or usual-care arms. The GenoVA Study has been approved by the VABHS institutional review board (IRB #3241), and all participants provide informed consent to participate.

Modeling clinical polygenic-risk-score testing and reporting

We have previously described our development of the GenoVA Study clinical PRS assay and reporting workflows.[12] In brief, we accessed publicly available genetic loci and weights for PRSs for the six target diseases, developed an assay and bioinformatics pipeline to calculate these PRSs from Illumina Global Diversity Array genotype data in a CLIA-certified laboratory, and confirmed the disease associations of the six PRSs from this assay in an independent cohort, the Mass General Brigham Biobank. Figure S1 illustrates the format and content of the resulting PRS report.

The choices made in how to report PRSs to participants and PCPs reflect the study's conceptual model and focus on actionability. That is, the study chose to report dichotomous PRS categories ("high risk" vs. "average risk") instead of continuous scores to simplify interpretability for patients and providers.[12,15] In the GenoVA Study, we defined a high-risk PRS as one associated with a published odds ratio (OR) of >2 for the target disease, as compared to the median PRS value. Although estimating absolute risk may be considered the gold standard for risk stratification for certain diseases (e.g., the pooled cohort equations for atherosclerotic cardiovascular disease and the BOADICEA model for breast cancer),[16,17] validated absolute risk models are not available for most diseases screened for in primary care. The OR > 2 threshold approximates the effect sizes considered significant in Mendelian genetics[18] and of other risk factors, such as family history or body-mass index, considered clinically important for risk stratification for the target diseases.[19–23] We also chose to report ACMG actionable monogenic findings separately from the PRS results. Despite evidence that PRS might modulate the penetrance of monogenic variants associated with the same diseases,[24] integrated models are not yet robustly validated for clinical use in diverse populations and should not be used to lessen the significance of the monogenic findings, which have more established guidelines governing their management.[25–27] The PRS report itself models the format and content of a traditional laboratory report (Figure S1). That is, it reports the individual patient's PRS results but does not contextualize those results amidst any other clinical risk factors (e.g., smoking status for cardiovascular disease) or protective factors (e.g., recent negative colonoscopy for colorectal cancer) the patient might have, information often unavailable to a clinical laboratory. On the other hand, the overall delivery of the PRS report back to the primary care context was designed to support its use in clinical decision-making, as described below.

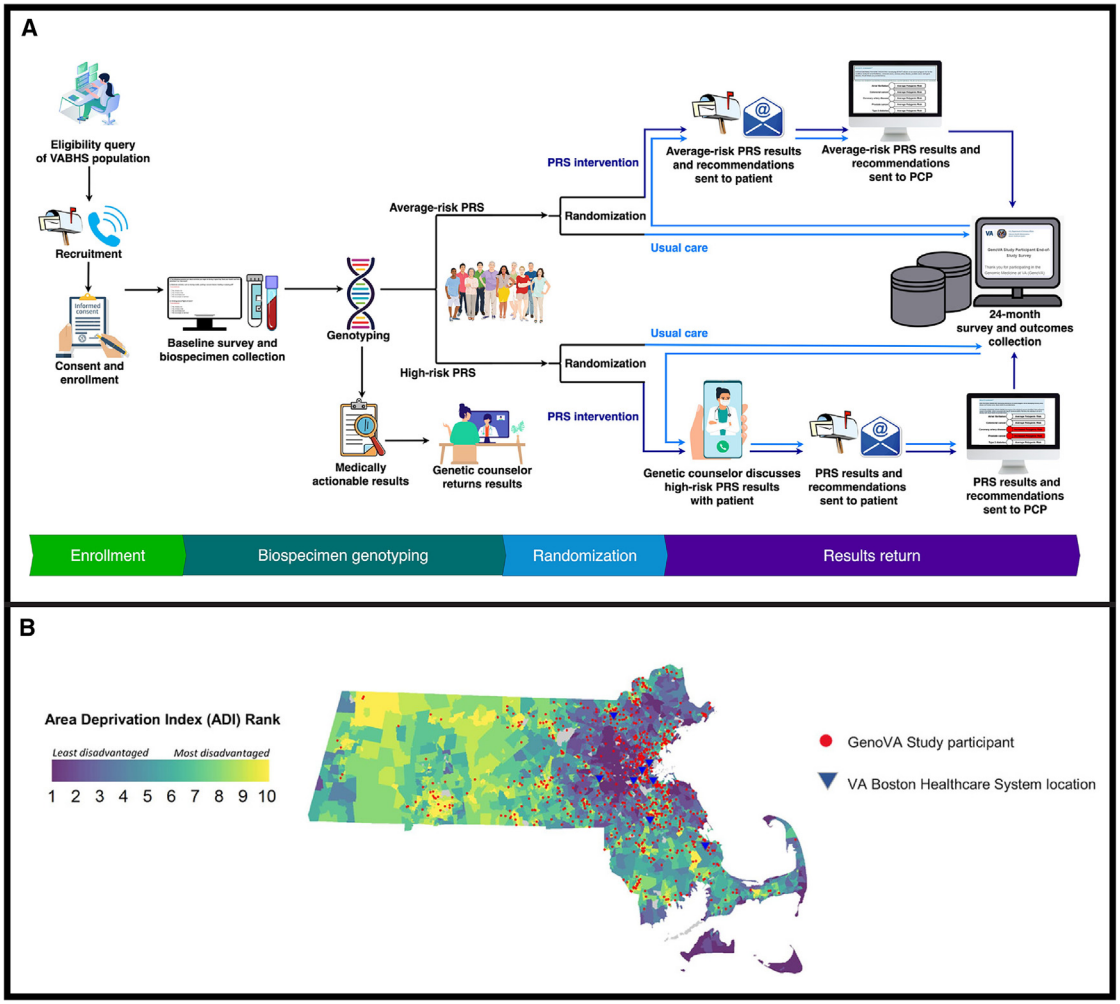


Figure 1. Processes and geographic catchment in the GenoVA Study

A flow diagram of the GenoVA Study process (A) illustrates how eligible participants are identified from an electronic health record (EHR) query and recruited via mailings, emails, and phone calls. Enhanced recruitment efforts target underrepresented gender, race, and ethnicity patient populations. Consent documents are signed remotely prior to baseline survey completion and collection of either blood or saliva. Genotyping identifies the subset of participants with an actionable monogenic disease result, as defined by the American College of Medicine Genetics and Genomics. These participants receive their monogenic and PRS results from a genetic counselor, who refers them to appropriate care. All other participants undergo randomization to the PRS intervention or usual-care arm and are stratified by risk status (at least one high-risk PRS versus only average-risk PRS results). In the intervention arm, any participant with at least one high-risk PRS result receives their results by phone or video from a genetic counselor; participants with only average-risk PRS results receive their results by mail or e-mail. PCPs are notified of results and recommendations by e-mail and through the EHR. End-of-study data collection from the EHR and surveys occurs 24 months after randomization. Participants in the usual-care arm receive their PRS results and recommendations through the same procedures as the intervention arm at the end of study. (B) State-level area deprivation index (ADI) values and relative geolocation of GenoVA Study participants across Massachusetts. ADI is a census block group level neighborhood disadvantage measure composed of 17 factors, including income, education, employment, and housing quality factors derived from American Community Survey data; values range from 1 to 10. Also shown are the eight healthcare facility locations of the VA Boston Health System. Abbreviations: PCP, primary care provider; PRS, polygenic risk score; VABHS, Veterans Affairs Boston Healthcare System.

Modeling the primary preventive-care context

Given that most preventive care is delivered in primary-care settings, we aimed to model the GenoVA Study PRS intervention and processes on how they might plausibly be introduced within this constrained clinical context. The following elements of the GenoVA Study reflect this goal. First, eligible patients are 50–70 years of age, a window during which many common preventive screenings occur in adult primary care.[19–22] Second, eligible patients have no known diagnoses of the target diseases, presenting the opportunity for primary prevention or early detection. The study does not, therefore, model the scenario where patients receive PRS results for diseases they are already known to have, an increasingly likely occurrence if PRSs for multiple diseases are more widely implemented.[28] Third, we chose six target diseases that are seen commonly in adult primary care and for which PCPs have established

guidelines or practice for their prevention, screening, and diagnosis. These choices allow an examination of PRSs as a complementary tool for PCPs' preventive practices, as opposed to as an isolated technology without familiar clinical anchoring. A high-risk PRS result for a disease is presented as an additional risk factor for the PCP to consider. The choice of familiar diseases also lessens the concern that unprepared PCPs will overinterpret PRS results and order unnecessary, costly, or even harmful follow-up tests or procedures; any test a PCP might recommend upon learning of a high-risk PRS result is likely to fall within guideline-recommended care (e.g., hemoglobin A1c testing for diabetes screening or mammography for breast cancer screening) and simply be an addition to currently used risk-stratification tools. Fourth, the GenoVA Study supports PRS test ordering by removing PRS consent and order entry from the PCP's responsibilities but still reports the PRS results back to them, similar to other clinical-decision support or population-management programs healthcare systems often use to promote the systematic use of established preventive-care interventions, such as EHR alerts to prompt cholesterol test ordering and nurse-led lung cancer screening programs to identify and consent eligible primary-care patients for computed tomography.[29]

While the content of GenoVA Study PRS report itself focuses on the technical and laboratory aspects of the results, how the PRS reports and accompanying supportive information are delivered to the patient and PCP supported the interpretation of the individual patient's PRS results, contextualization amidst other clinical factors, and clinical decision-making. The PRS results are sent to both the PCP and the patient, in addition to being entered in the EHR, both as a portable document format (pdf) report and as structured data in the laboratory information-management system (Figure S2).[12] Any patient with at least one high-risk PRS is also contacted by a genetic counselor to discuss the results, potential implications for the patient's health and health care, and recommendations for next steps in talking about the results with their PCPs. A disease-specific layperson information sheet is provided to the patient and outlines potential clinical management options.[12] Similarly, the PCP is sent a provider-oriented disease-specific information sheet with details about the PRS, its limitations, recommendations for contextualizing the results among other patient characteristics and risk factors, and management suggestions, including information about current screening guidelines. In contrast, any patient with no high-risk results simply receives their report and a brief letter stating that none of their PRS results indicated high risk. This delivery was modeled on common primary-care practices in lab-results reporting: patients with abnormal results often receive phone calls or have follow-up visits to discuss the results and next steps in management, whereas patients with normal results often receive their results via brief letters or notifications via patient portal.

Determining clinical utility

There is no clear agreement on how to define the clinical utility of PRSs in the preventive medicine setting, nor how to measure that utility. Proponents of PRSs argue that, because scores in the upper tail of the normal distribution can indicate risk comparable to rare variants associated with monogenic forms of disease, PRSs could similarly influence clinical screening, prevention, and management strategies.[1,2,30] Critics argue that PRSs achieve similar discrimination for disease risk as other risk factors already used in clinical care (e.g., body-mass index, family history, and smoking) or readily available without additional testing (e.g., socioeconomic status).[31,32] Thus, for some diseases, it is not clear whether PRSs improve current clinical standards of care in disease prediction and prevention. As a concept, the clinical utility of genetic testing in general and PRSs specifically have been variably defined on a spectrum from narrow to expansive, depending on the context and purpose of the definition. The clinical utility of PRSs most often refers to their ability to improve patient outcomes, often through the prevention or amelioration of mortality, morbidity, or disability through the adoption of effective interventions based on the test results.[1,3,4,33]

In the GenoVA Study context, we define clinical utility as the ability of a PRS test to result in earlier diagnosis of clinically significant cases of the target diseases; earlier diagnosis and treatment are associated with the improved outcomes of lower morbidity and mortality.[34–37] The trial's primary outcome operationalizes the measurement of that utility: time to diagnosis of both undiagnosed prevalent cases of the six diseases and incident cases during the 24-month observation period after randomization. Despite the apparent paradox that the use of PRSs in preventive medicine might accelerate the diagnosis of disease instead of preventing disease onset, the trial's primary hypothesis is that, among participants with at least one high-risk PRS result, the time to diagnosis will be shorter for participants receiving PRS results than for those receiving usual care. This choice of outcome best fits the population of 50- to 70-year-old adults, given their baseline annual diagnosis rate of 6.2% (Note S1) and likely high proportion of undiagnosed prevalent cases. On the basis of *a priori* assumptions that 33% of participants would have at least one high-risk PRS result and that 12% of high-risk patients in the usual-care arm will receive a new target diagnosis during the 24-month observation period, the GenoVA Study has 80% power at a two-tailed $\alpha = 0.05$ to detect a 2-fold increase in new diagnoses among high-risk patients receiving the PRS intervention.

Participants and their PCPs are not blinded to participant allocation to the intervention or study arms. Although essential for drug trials, blinding is impossible for a trial whose aim is to determine whether participants and their PCPs act on PRS results in a way that improves health outcomes. Indeed, the study hypotheses assume

that this unblinded, differential knowledge about disease risk will lead to increased surveillance and disease diagnosis among high-risk patients. This does not represent a detection bias but, rather, the intention of the PRS intervention. The primary threat to study validity from this lack of blinding, then, would be differential outcomes assessment at the end of the trial. The GenoVA Study minimizes the risk of detection bias by including in the primary outcome only strictly defined clinically significant cases (e.g., only prostate cancer cases classified as intermediate risk or higher by National Comprehensive Cancer Network guidelines[38] and only cases of atrial fibrillation meeting guidelines for clinical management[39]), as adjudicated by expert reviewers blinded to randomization status or PRS results. This measure minimizes the overdetection of clinically insignificant disease, the diagnosis of which is less certain to be associated with improved outcomes.

In choosing new diagnoses of clinically significant disease as its primary outcome, the GenoVA Study has adopted a narrow definition of clinical utility. An even narrower definition might require demonstration of reduced mortality or morbidity, often measured with quality-adjusted life years; rigorous trials powered to detect meaningful differences in these outcomes will require longer follow-up. Although narrow definitions of clinical utility are most often used by evidence-based guidelines and healthcare payers,[40,41] broader definitions of the clinical utility of PRS include their ability to inform clinical decision-making, a type of clinical actionability.[42] Absent evidence for narrower-sense clinical utility, actionability is an admittedly subjective term, despite efforts to generate expert-informed consensus around its definition and quantification.[43,44] In the GenoVA Study, the secondary outcome of diagnostic testing ordered by treating clinicians reflects a measurement of actionability. In its broadest sense, clinical utility can also refer to the ability of a test to improve any outcomes considered important to individuals and families; such outcomes might include psychosocial wellbeing and reproductive decision-making and are often termed personal utility.[1,45] Additional GenoVA Study outcomes including patient activation in their healthcare, self-assessed health status, and quality of life capture this broader scope of the utility patients might derive from PRS testing. Table S2 shows the secondary and exploratory outcomes, including follow-up diagnostic testing, patient medication adherence, healthcare costs, and quality of life, each representing process and implementation outcomes relevant to determining the value and costs of integrating PRS into adult preventive medicine. Note S4 includes the GenoVA Study statistical analysis plan, which provides greater detail about outcomes measurement and statistical approach.

The choice of a pragmatic randomized clinical trial

There is also no agreement on the type of study design and evidence needed to demonstrate the clinical utility of PRSs. Although observational studies provide valuable evidence when RCTs are not feasible, RCTs remain at the top of the evidence hierarchy, given their ability to minimize bias and confounding.[46] However, RCTs have their own limitations, namely that controlled experimental conditions limit the generalizability of the findings to real-world contexts. Whether RCT evidence is needed to demonstrate the clinical utility of PRSs is an unsettled question. On one hand, RCTs are generally accepted as the gold standard for determining the effectiveness of interventions. On the other hand, most laboratory tests used routinely in clinical medicine, such as kidney function testing and complete blood counts, are not supported by RCT evidence.[47,48] The question of appropriate study design, then, might hinge on whether a PRS is considered a laboratory test or as one component of a preventive genomics intervention. The GenoVA Study models PRS testing as the latter, the first step of an intervention that also includes interpreted PRS reporting, targeted genetic counseling for high-risk individuals, communication with each participant's PCP, and patient- and provider-oriented materials to support decision-making around PRS results. An RCT design is thus appropriate for the GenoVA Study in this context. It is worth noting that the ultimate outcomes of the GenoVA Study RCT will need to be interpreted in the setting of this overall intervention, not in terms of a PRS in isolation.

Moreover, a pragmatic design is appropriate, given the study aim to determine the clinical effectiveness of PRS testing in a real-world primary-care context and the multifaceted, preventive nature of the intervention.[49] This contrasts with a treatment trial warranting a more explanatory trial design to demonstrate biological impact.[49] The pragmatic design also affords the opportunity to collect implementation outcomes relevant to stakeholders interested in the adoption of PRSs in clinical care; such outcomes include healthcare costs and participant- and provider-reported outcomes.[49,50] Pragmatic design elements of the GenoVA Study include embedding into existing clinical workflows the PRS test ordering, the send-out to a reference laboratory, and results reporting. Another pragmatic element is the collection of trial-outcome EHR data, supplemented with end-of-study survey data. Figure S3 displays a *pr*agmatic *e*xplanatory *c*ontinuum *i*ndicator *s*ummary 2 (PRECIS-2) wheel illustrating the degree to which the GenoVA Study design is considered pragmatic versus explanatory.[50] Table S3 further elucidates each of the trial's design features and pragmatic elements. Our intention is that the pragmatic design will increase the likelihood that the RCT results are relevant to how PRSs might reasonably be implemented into routine primary care.

Promoting health equity in clinical PRS implementation

The GenoVA Study affords the opportunity to address a pressing ethical challenge to the clinical implementation

of PRS: the risk of exacerbating health disparities among populations already at higher risk of poor health outcomes. The associations between PRSs and disease risk are most robustly validated for populations descended from European continental ancestry groups.[11,51] Despite advances in dataset diversity, statistical methods, and trans-ancestry PRS development and validation, this disparity in PRS performance is reduced, but not eliminated.[7–10]

Challenges addressed

We have taken several specific actions to leverage the GenoVA Study as an opportunity to promote health equity in the clinical implementation of PRSs. First, recognizing that most data from genome-wide association studies are derived from European-ancestry populations, we paid significant attention to the handling of genetic ancestry in constructing the PRSs and in validating our proposed PRS in the multiracial Mass General Brigham Biobank.[12] As described previously, instead of developing multiple population-specific PRSs (e.g., by continental genetic ancestry group or self-reported racial or ethnic group), we chose to validate a single, genetic principal-components-adjusted PRS for each disease for application across populations. At the same time, we transparently include in our PRS laboratory report a description of the populations in which the PRSs were developed and validated, highlighting the limited population diversity for some PRSs. In these efforts, we are intentional in how we use population descriptors such as racial categories and genetic ancestry groups so as not to conflate biological and social constructs or suggest that racial categories have biological meaning.[52]

Second, we developed recruitment strategies to address underrepresentation in biomedical research. Even if PRSs of equal accuracy across all populations are developed, existing healthcare inequities, including disparate access to care and legacies of untrustworthy healthcare systems, are still likely to impede equitable implementation. The learning healthcare system of the Veterans Health Administration offers a unique setting in which to address these challenges. Although racial and ethnic disparities in healthcare and health outcomes persist, VA as an "equal-access" healthcare system outperforms other systems on several disparity measures.[53,54] Even within this setting, the GenoVA Study implemented enhanced recruitment measures to increase representation of patient populations less likely to participate in biomedical research. That is, we preferentially directed recruitment efforts (e.g., mailings and phone calls) to VABHS patients identified as non-white, Hispanic or Latino, or female, all minority populations at VABHS.

Third, we promote gender identity equity in allowing participants to describe their sex assigned at birth and gender identity and use inclusive, specific language to describe which participants may receive PRS results for prostate cancer risk (i.e., those born with a prostate) and breast cancer risk (i.e., those born of natal female sex, given that the validation of breast cancer PRSs has been limited to this population).

Equitable implementation outcomes

Analysis of recruitment and enrollment data to date illustrates these efforts. From June 17, 2020 to May 10, 2023, a total of 10,036 patients across VABHS were deemed eligible for study participation by a computable eligibility classifier described previously (Table 1).[55] Among this eligible population, VABHS administrative data categorize 15.5% as having a race other than white, 3.4% as having Hispanic or Latino ethnicity, and 12.2% as female. Figure S4 shows the GenoVA Study recruitment and enrollment efforts as of May 10, 2023. Among the 2,083 participants who actively declined participation, 1,165 (56%) individuals offered a reason. Common reasons included time constraints (n = 407, 35%); ethical, legal, and social considerations (n = 173, 15%); lack of interest (n = 162, 14%); health reasons (n = 114, 9.8%); dislike of research (n = 60, 5.2%); and VA- or government-related reasons (n = 38, 3.2%).

We observed that our enhanced recruitment efforts were necessary—as we did observe that participants of non-white race or Hispanic ethnicity were less likely to enroll than expected—and successful, in that they ultimately yielded proportional or higher representation of women and non-white or Hispanic participants from the VABHS population overall. Table S4 shows overall observed and expected rates of study acceptance (defined as agreement to receive a consent packet among eligible phone call respondents, n = 3,855) and enrollment (defined as return of consent documents among those who agreed to receive a consent packet, n = 2,107). Overall, women accepted study participation ($p < 0.001$, Cramer's V 0.072) and enrolled ($p < 0.046$, Cramer's V 0.039) in slightly greater proportions than would be expected if acceptance and enrollment were proportional across demographic categories.[56] Non-white or Hispanic participants accepted study participation ($p < 0.001$, Cramer's V 0.087) in slightly greater proportions than expected but were less likely to enroll ($p < 0.001$, Cramer's V 0.104) by returning consent documents than expected. Table 1 shows baseline demographic and clinical characteristics of the 966 participants who have enrolled and provided a viable biospecimen, in comparison to the overall eligible VABHS patient population who did not enroll. The study achieved overrepresentation of women and Hispanic/Latino enrollees. Among enrollees, 182 (19%) identify as women, including three transgender women, compared to 12.8% of the overall eligible VABHS population ($p < 0.001$, Cramer's V 0.058), and 53 (5.5%) report Hispanic/Latino ethnicity ($p < 0.001$, Cramer's V 0.042), in comparison to 3.6% of the overall eligible VABHS population. Racial representation was achieved: 144 (14.9%) of enrollees report one or more racial identities other than white, in comparison to 15.4% of the overall eligible population ($p = 0.149$, Cramer's V 0.019).

Table 1. Baseline demographic and clinical characteristics of first 966 GenoVA Study participants with collected biospecimens and the remaining 9,070 eligible members of the VABHS patient population

	GenoVA patients with collected biospecimens (n = 966)		GenoVA eligible patients (n = 9,070)	
	Mean	**SD**	**Mean**	**SD**
Age (50–70 years)	60.47	5.65	60.68	5.83
Body mass index (BMI)	30.37	6.30	29.65	5.95
Systolic blood pressure (SBP)	130.1	14.45	131.03	16.22
Diastolic blood pressure (DBP)	79.74	8.73	80.00	9.49
Low-density lipoprotein cholesterol (LDL-C)	110.65	33.09	112.69	34.67
	n	**%**	**N**	**%**
Gender				
Female[a]	182	18.5	1,106	12.2
Male	784	81.5	7,964	87.8
Race				
American Indian/Alaskan Native	3	0.3	35	0.4
Asian	9	0.9	46	0.5
Black/African American	112	11.6	1,120	12.4
Native Hawaiian/Pacific Islander	5	0.5	34	0.4
White	766	79.3	6,965	76.8
Multiracial	15	1.6	162	1.8
Unknown or declined	56	5.8	708	7.8
Ethnicity				
Hispanic of Latino	53	5.5	306	3.4
Not Hispanic or Latino	824	85.3	7,616	84.0
Unknown or declined	89	9.2	1,148	12.7
Rurality[b]				
Rural	48	5.0	796	8.8
Urban	917	95.0	8,269	91.2
Area deprivation index[c] category (ADI; state rank)				
Least disadvantaged (ADI 1–3)	228	23.6	2,109	23.3
Moderately disadvantaged (ADI 4–6)	359	37.2	3,363	37.1
Most disadvantaged (ADI 7–10)	364	37.7	3,466	38.2
Unknown or suppressed	15	1.5	132	1.5

[a]Including three transgender women with male biological sex.
[b]Rural or urban designation is attributed to geocoded patient location data as validated by the VA Geospatial Service Support Center. Rural status includes "rural" and "highly rural" designations from VA data. Six individuals included in the total cohort (n = 10,036) have undesignated rurality status.
[c]Area deprivation index (ADI) derived via 2020 FIPS-level ADI, 2020 US Census Block Group shapefile boundaries, and VA Boston Healthcare System (station 523) geocoded patient location data as validated by the VA Geospatial Service Support Center. State rank is based on validated geolocated state of residence. 84.16% of the total cohort (n = 8,446) and 95.24% of participants with collected biospecimens (n = 920) were designated as Massachusetts residents. 147 individuals in the total cohort have missing geolocation data or suppressed ADI designations of high group quarters, low population or housing, both high group quarters and low population or housing, or questionable data integrity.

We also examined differences in expected and observed recruitment and enrollment outcomes by neighborhood disadvantage, as measured with the state-level area deprivation index (ADI; 1 = least deprived to 10 = most deprived). ADI is a census block group-level neighborhood disadvantage measure composed of 17 factors, including income, education, employment, and housing quality factors derived from American Community Survey (ACS) data.[57] Similar proportions of participants within each ADI category were as likely to accept study participation ($p = 0.241$, Cramer's V 0.015) and enroll ($p = 0.982$, Cramer's V < 0.001) as expected (Table S4). Among

enrollees with available data (n = 951), large proportions reside in highly disadvantaged (ADI 7–10, 38.7%) and moderately disadvantaged (ADI 4–6, 37.6%) areas. Figure 1 maps the ADI and relative geolocation of GenoVA Study enrollees across Massachusetts. In contrast, fewer rural participants have enrolled, in comparison to the remainder of the eligible patient cohort (5.0% versus 8.8%, respectively); rural status was defined by the VA Geospatial Service Support Center.[58] We did not implement enhanced efforts to recruit participants from rural areas and did in fact observe lower representation of rural enrollees. Although rural participants were as likely to accept study participation as expected ($p = 0.243$, Cramer's V 0.012), they were ultimately slightly less likely to enroll than expected ($p = 0.031$, Cramer's V 0.044). Recruitment and retention of socioeconomically deprived and rural populations is a well-known challenge in clinical research, including among Veterans,[59,60] but our success in targeted recruitment from other underrepresented populations gives hope that similar efforts would help reach these groups.

As of May 10, 2023, study staff have received interpreted PRS reports for 840 participants. Thirteen (2%) of these had at least one positive result from the current ACMG actionable secondary findings gene list (Table S5).[14] Table S6 shows the distribution of all 840 PRS results received as of May 10, 2023; 307 (37.1%) participants have at least one high-risk PRS result, and 54 (6.5%) of these have two or more high-risk PRS results, consistent with expected results.[12] Of these 307 high-risk participants with available demographic data, 238 are white (77.5%) and 60 are non-white or Hispanic (19.5%); this racial and ethnic demographic composition is consistent with overall recruitment percentages and is not significantly different from participants with average-risk results ($p = 0.483$, Cramer's $V < 0.001$).

Enrolled participants are now being followed for 24 months for the study outcomes, including disease diagnoses (primary outcome) and diagnostic testing (secondary outcome). We will report these outcomes in aggregate and stratified by gender; race and ethnicity; neighborhood deprivation; and rural status.[61] Planned analyses of primary and secondary outcomes will include participant sex as a covariate because of the study's sex-stratified randomization (Note S4). We will also perform exploratory analyses to investigate whether heterogeneous effects of the PRS intervention exist among different demographic groups. For example, we will include participant race, ethnicity, ADI, and rurality separately and in combination as factor variables in our statistical models for the pre-specified outcomes to identify between-group differences and generate hypotheses for how the introduction of PRSs in primary care might differentially impact certain groups. These analyses of GenoVA Study process and outcome measures will facilitate the identification of points in the PRS clinical-implementation pathway where disparities might exist and should be addressed.

Conclusions

The clinical implementation of PRSs is moving forward through clinical programs, research projects, and commercial laboratory and direct-to-consumer offerings,[28,62–64] and a limited number of important RCTs have or will inform the clinical utility of PRS in single-disease settings.[65–70] As a pragmatic RCT implementing a multi-disease PRS intervention, the GenoVA Study makes a unique contribution to informing the equitable implementation of PRSs for preventive medicine in the time-constrained primary-care context. Its design as a pragmatic trial enhances the generalizability of its ultimate findings, and its RCT design adds rigor to hypothesis-testing about the impact of PRS testing on preventive-medicine processes and outcomes. The VA is the largest healthcare system in the United States. Although this setting limits the generalizability of the GenoVA Study's findings to other settings in some respects, lessons learned from the study still offer potential solutions for assessing the clinical utility of implementing PRSs into adult primary care while attending to the potential of that implementation to hinder or promote health equity.

Supplemental information

Supplemental information can be found online at https://doi.org/10.1016/j.ajhg.2023.10.001.

Acknowledgments

This work was supported by NIH (National Institutes of Health) National Human Genome Research Institute (NHGRI) grant R35HG010706. P.N. is supported by NIH NHGRI grant U01HG011719. S.A.L. previously received support from NIH grants R01HL139731 and R01HL157635, and American Heart Association 18SFRN34250007 during this project. K.D.C. was previously supported by a research grant from Sanford Health. R.C.G.'s research is supported by grant funding from the National Institutes of Health, the Department of Defense, and the Franca Sozzani Fund for Preventive Genomics.

Author contributions

J.L.V., C.A.B., M.S.L., N.V.J.A., M.G., R.C.G., P.K., S.A.L., and P.N. conceived the study and contributed to its design. M.S.L. analyzed the data from the MGBB and GenoVA samples. A.A.A., C.A.B., M.D., N.E.J., P.K., E.H., K.M., M.P.C., and J.L.V. collected GenoVA data. J.L.V., C.A.B., K.M., N.E.J., and A.A.A. drafted the manuscript, and all authors reviewed the scientific content of the manuscript prior to submission.

Declaration of interests

P.N. reports research grants from Allelica, Apple, Amgen, Boston Scientific, Genentech/Roche, and Novartis; and personal fees from Allelica, Apple, AstraZeneca, Blackstone Life Sciences, Eli Lilly & Co, Foresite Labs, Genentech/Roche, GV, HeartFlow, Magnet Biomedicine, and Novartis; scientific advisory board

membership of Esperion Therapeutics, Preciseli, and TenSixteen Bio; status as a scientific co-founder of TenSixteen Bio; equity in MyOme, Preciseli, and TenSixteen Bio; and spousal employment at Vertex Pharmaceuticals, all unrelated to the present work. S.A.L. is an employee of Novartis as of July 18, 2022, has previously received sponsored research support from Bristol Myers Squibb, Pfizer, Boehringer Ingelheim, Fitbit, Medtronic, Premier, and IBM, and has consulted for Bristol Myers Squibb, Pfizer, Blackstone Life Sciences, and Invitae. A.C.F.L. reports owning stock in Fabric Genomics. R.C.G. has received compensation for advising Allelica, Atria, Fabric, Genome Web, Genomic Life, and Juniper Genomics and is co-founder of Genome Medical and Nurture Genomics. The other authors declare no competing interests.

Web resources

2020 12-Digit FIPS code block group area deprivation index (ADI) data, https://www.neighborhoodatlas.medicine.wisc.edu/

2020 United States Census Bureau cartographic boundary files, https://www.census.gov/geographies/mapping-files/time-series/geo/cartographic-boundary.2020.html#list-tab-1883739534

References

1. Torkamani, A., Wineinger, N.E., and Topol, E.J. (2018). The personal and clinical utility of polygenic risk scores. Nat. Rev. Genet. *19*, 581–590. https://doi.org/10.1038/s41576-018-0018-x.
2. Lambert, S.A., Abraham, G., and Inouye, M. (2019). Towards clinical utility of polygenic risk scores. Hum. Mol. Genet. *28*, R133–R142. https://doi.org/10.1093/hmg/ddz187.
3. Lewis, C.M., and Vassos, E. (2020). Polygenic risk scores: from research tools to clinical instruments. Genome Med. *12*, 44. https://doi.org/10.1186/s13073-020-00742-5.
4. Abu-El-Haija, A., Reddi, H.V., Wand, H., Rose, N.C., Mori, M., Qian, E., Murray, M.F.; and ACMG Professional Practice and Guidelines Committee. Electronic address: documents@acmg.net (2023). The clinical application of polygenic risk scores: A points to consider statement of the American College of Medical Genetics and Genomics (ACMG). Genet. Med. *25*, 100803. https://doi.org/10.1016/j.gim.2023.100803.
5. Huntley, C., Torr, B., Sud, A., Rowlands, C.F., Way, R., Snape, K., Hanson, H., Swanton, C., Broggio, J., Lucassen, A., et al. (2023). Utility of polygenic risk scores in UK cancer screening: a modelling analysis. Lancet Oncol. *24*, 658–668. https://doi.org/10.1016/S1470-2045(23)00156-0.
6. Grosse, S.D., and Khoury, M.J. (2006). What is the clinical utility of genetic testing? Genet. Med. *8*, 448–450. https://doi.org/10.1097/01.gim.0000227935.26763.c6.
7. Wang, Y., Tsuo, K., Kanai, M., Neale, B.M., and Martin, A.R. (2022). Challenges and Opportunities for Developing More Generalizable Polygenic Risk Scores. Annu. Rev. Biomed. Data Sci. *5*, 293–320. https://doi.org/10.1146/annurev-biodatasci-111721-074830.
8. Ge, T., Irvin, M.R., Patki, A., Srinivasasainagendra, V., Lin, Y.-F., Tiwari, H.K., Armstrong, N.D., Benoit, B., Chen, C.-Y., Choi, K.W., et al. (2022). Development and validation of a trans-ancestry polygenic risk score for type 2 diabetes in diverse populations. Genome Med. *14*, 70. https://doi.org/10.1186/s13073-022-01074-2.
9. Weissbrod, O., Kanai, M., Shi, H., Gazal, S., Peyrot, W.J., Khera, A.V., Okada, Y., Biobank Japan Project, Martin, A.R., Finucane, H.K., and Price, A.L. (2022). Leveraging fine-mapping and multipopulation training data to improve cross-population polygenic risk scores. Nat. Genet. *54*, 450–458. https://doi.org/10.1038/s41588-022-01036-9.
10. Ding, Y., Hou, K., Xu, Z., Pimplaskar, A., Petter, E., Boulier, K., Privé, F., Vilhjálmsson, B.J., Olde Loohuis, L.M., and Pasaniuc, B. (2023). Polygenic scoring accuracy varies across the genetic ancestry continuum. Nature *618*, 774–781. https://doi.org/10.1038/s41586-023-06079-4.
11. Martin, A.R., Kanai, M., Kamatani, Y., Okada, Y., Neale, B.M., and Daly, M.J. (2019). Clinical use of current polygenic risk scores may exacerbate health disparities. Nat. Genet. *51*, 584–591. https://doi.org/10.1038/s41588-019-0379-x.
12. Hao, L., Kraft, P., Berriz, G.F., Hynes, E.D., Koch, C., Korategere V Kumar, P., Parpattedar, S.S., Steeves, M., Yu, W., Antwi, A.A., et al. (2022). Development of a clinical polygenic risk score assay and reporting workflow. Nat. Med. *28*, 1006–1013. https://doi.org/10.1038/s41591-022-01767-6.
13. Vassy, J.L., Chun, S., Advani, S., Ludin, S.A., Smith, J.G., and Alligood, E.C. (2019). Impact of SLCO1B1 Pharmacogenetic Testing on Patient and Healthcare Outcomes: A Systematic Review. Clin. Pharmacol. Ther. *106*, 360–373. https://doi.org/10.1002/cpt.1223.
14. Miller, D.T., Lee, K., Abul-Husn, N.S., Amendola, L.M., Brothers, K., Chung, W.K., Gollob, M.H., Gordon, A.S., Harrison, S.M., Hershberger, R.E., et al. (2022). ACMG SF v3.1 list for reporting of secondary findings in clinical exome and genome sequencing: A policy statement of the American College of Medical Genetics and Genomics (ACMG). Genet. Med. *24*, 1407–1414. https://doi.org/10.1016/j.gim.2022.04.006.
15. Lewis, A.C.F., Green, R.C., and Vassy, J.L. (2021). Polygenic risk scores in the clinic: Translating risk into action. HGG Adv. *2*, 100047. https://doi.org/10.1016/j.xhgg.2021.100047.
16. Lee, A., Mavaddat, N., Wilcox, A.N., Cunningham, A.P., Carver, T., Hartley, S., Babb de Villiers, C., Izquierdo, A., Simard, J., Schmidt, M.K., et al. (2019). BOADICEA: a comprehensive breast cancer risk prediction model incorporating genetic and nongenetic risk factors. Genet. Med. *21*, 1708–1718. https://doi.org/10.1038/s41436-018-0406-9.
17. Stone, N.J., Robinson, J.G., Lichtenstein, A.H., Bairey Merz, C.N., Blum, C.B., Eckel, R.H., Goldberg, A.C., Gordon, D., Levy, D., Lloyd-Jones, D.M., et al. (2014). 2013 ACC/AHA Guideline on the Treatment of Blood Cholesterol to Reduce Atherosclerotic Cardiovascular Risk in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. Circulation *129*, S1–S45. https://doi.org/10.1161/01.cir.0000437738.63853.7a.
18. Senol-Cosar, O., Schmidt, R.J., Qian, E., Hoskinson, D., Mason-Suares, H., Funke, B., and Lebo, M.S. (2019). Considerations for clinical curation, classification, and reporting of low-penetrance and low effect size variants associated with disease risk. Genet. Med. *21*, 2765–2773. https://doi.org/10.1038/s41436-019-0560-8.
19. US Preventive Services Task Force, Grossman, D.C., Curry, S.J., Owens, D.K., Bibbins-Domingo, K., Caughey, A.B., Davidson, K.W., Doubeni, C.A., Ebell, M., Epling, J.W., Jr., et al. (2018). Screening for Prostate Cancer: US Preventive Services Task Force Recommendation Statement. JAMA *319*, 1901–1913. https://doi.org/10.1001/jama.2018.3710.

20. US Preventive Services Task Force, Davidson, K.W., Barry, M.J., Mangione, C.M., Cabana, M., Caughey, A.B., Davis, E.M., Donahue, K.E., Doubeni, C.A., Krist, A.H., et al. (2021). Screening for Colorectal Cancer: US Preventive Services Task Force Recommendation Statement. JAMA *325*, 1965–1977. https://doi.org/10.1001/jama.2021.6238.
21. Siu, A.L.; and U.S. Preventive Services Task Force (2016). Screening for Breast Cancer: U.S. Preventive Services Task Force Recommendation Statement. Ann. Intern. Med. *164*, 279–296. https://doi.org/10.7326/M15-2886.
22. ElSayed, N.A., Aleppo, G., Aroda, V.R., Bannuru, R.R., Brown, F.M., Bruemmer, D., Collins, B.S., Hilliard, M.E., Isaacs, D., Johnson, E.L., et al. (2023). 2. Classification and Diagnosis of Diabetes: Standards of Care in Diabetes-2023. Diabetes Care *46*, S19–S40. https://doi.org/10.2337/dc23-S002.
23. Arnett, D.K., Blumenthal, R.S., Albert, M.A., Buroker, A.B., Goldberger, Z.D., Hahn, E.J., Himmelfarb, C.D., Khera, A., Lloyd-Jones, D., McEvoy, J.W., et al. (2019). 2019 ACC/AHA Guideline on the Primary Prevention of Cardiovascular Disease: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Circulation *140*, e596–e646. https://doi.org/10.1161/CIR.0000000000000678.
24. Fahed, A.C., Wang, M., Homburger, J.R., Patel, A.P., Bick, A.G., Neben, C.L., Lai, C., Brockman, D., Philippakis, A., Ellinor, P.T., et al. (2020). Polygenic background modifies penetrance of monogenic variants for tier 1 genomic conditions. Nat. Commun. *11*, 3635. https://doi.org/10.1038/s41467-020-17374-3.
25. Bevers, T.B., Niell, B.L., Baker, J.L., Bennett, D.L., Bonaccio, E., Camp, M.S., Chikarmane, S., Conant, E.F., Eghtedari, M., Flanagan, M.R., et al. (2023). NCCN Guidelines® Insights: Breast Cancer Screening and Diagnosis, Version 1.2023: Featured Updates to the NCCN Guidelines. J. Natl. Compr. Canc. Netw. *21*, 900–909. https://doi.org/10.6004/jnccn.2023.0046.
26. Weiss, J.M., Gupta, S., Burke, C.A., Axell, L., Chen, L.-M., Chung, D.C., Clayback, K.M., Dallas, S., Felder, S., Gbolahan, O., et al. (2021). NCCN Guidelines® Insights: Genetic/Familial High-Risk Assessment: Colorectal, Version 1.2021. J. Natl. Compr. Canc. Netw. *19*, 1122–1132. https://doi.org/10.1164/jnccn.2021.0048.
27. Watts, G.F., Gidding, S.S., Hegele, R.A., Raal, F.J., Sturm, A.C., Jones, L.K., Sarkies, M.N., Al-Rasadi, K., Blom, D.J., Daccord, M., et al. (2023). International Atherosclerosis Society guidance for implementing best practice in the care of familial hypercholesterolaemia. Nat. Rev. Cardiol., 1–25. https://doi.org/10.1038/s41569-023-00892-0.
28. Linder, J.E., Allworth, A., Bland, H.T., Caraballo, P.J., Chisholm, R.L., Clayton, E.W., Crosslin, D.R., Dikilitas, O., DiVietro, A., Esplin, E.D., et al. (2023). Returning integrated genomic risk and clinical recommendations: The eMERGE study. Genet. Med. *25*, 100006. https://doi.org/10.1016/j.gim.2023.100006.
29. Saab, M.M., O'Driscoll, M., Sahm, L.J., Leahy-Warren, P., Noonan, B., FitzGerald, S., Kilty, C., O'Malley, M., Lyons, N., and Hegarty, J. (2022). Referring high-risk individuals for lung cancer screening: A systematic review of interventions with healthcare professionals. Eur. J. Cancer Prev. *31*, 540–550. https://doi.org/10.1097/CEJ.0000000000000755.
30. Khera, A.V., Chaffin, M., Aragam, K.G., Haas, M.E., Roselli, C., Choi, S.H., Natarajan, P., Lander, E.S., Lubitz, S.A., Ellinor, P.T., and Kathiresan, S. (2018). Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nat. Genet. *50*, 1219–1224. https://doi.org/10.1038/s41588-018-0183-z.
31. Khoury, M.J., Janssens, A.C.J.W., and Ransohoff, D.F. (2013). How can polygenic inheritance be used in population screening for common diseases? Genet. Med. *15*, 437–443. https://doi.org/10.1038/gim.2012.182.
32. Inouye, M., Abraham, G., Nelson, C.P., Wood, A.M., Sweeting, M.J., Dudbridge, F., Lai, F.Y., Kaptoge, S., Brozynska, M., Wang, T., et al. (2018). Genomic Risk Prediction of Coronary Artery Disease in 480,000 Adults: Implications for Primary Prevention. J. Am. Coll. Cardiol. *72*, 1883–1893. https://doi.org/10.1016/j.jacc.2018.07.079.
33. Reddi, H.V., Wand, H., Funke, B., Zimmermann, M.T., Lebo, M.S., Qian, E., Shirts, B.H., Zou, Y.S., Zhang, B.M., Rose, N.C., et al. (2023). Laboratory perspectives in the development of polygenic risk scores for disease: A points to consider statement of the American College of Medical Genetics and Genomics (ACMG). Genet. Med. *25*, 100804. https://doi.org/10.1016/j.gim.2023.100804.
34. Crosby, D., Bhatia, S., Brindle, K.M., Coussens, L.M., Dive, C., Emberton, M., Esener, S., Fitzgerald, R.C., Gambhir, S.S., Kuhn, P., et al. (2022). Early detection of cancer. Science *375*, eaay9040. https://doi.org/10.1126/science.aay9040.
35. Chatterjee, S., Khunti, K., and Davies, M.J. (2017). Type 2 diabetes. Lancet *389*, 2239–2251. https://doi.org/10.1016/S0140-6736(17)30058-2.
36. Kirchhof, P., Camm, A.J., Goette, A., Brandes, A., Eckardt, L., Elvan, A., Fetsch, T., van Gelder, I.C., Haase, D., Haegeli, L.M., et al. (2020). Early Rhythm-Control Therapy in Patients with Atrial Fibrillation. N. Engl. J. Med. *383*, 1305–1316. https://doi.org/10.1056/NEJMoa2019422.
37. Fihn, S.D., Gardin, J.M., Abrams, J., Berra, K., Blankenship, J.C., Dallas, A.P., Douglas, P.S., Foody, J.M., Gerber, T.C., Hinderliter, A.L., et al. (2012). 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS Guideline for the Diagnosis and Management of Patients With Stable Ischemic Heart Disease: Executive Summary: A Report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines, and the American College of Physicians, American Association for Thoracic Surgery, Preventive Cardiovascular Nurses Association, Society for Cardiovascular Angiography and Interventions, and Society of Thoracic Surgeons. J. Am. Coll. Cardiol. *60*, 2564–2603. https://doi.org/10.1016/j.jacc.2012.07.012.
38. Moses, K.A., Sprenkle, P.C., Bahler, C., Box, G., Carlsson, S.V., Catalona, W.J., Dahl, D.M., Dall'Era, M., Davis, J.W., Drake, B.F., et al. (2023). NCCN Guidelines® Insights: Prostate Cancer Early Detection, Version 1.2023: Featured Updates to the NCCN Guidelines. J. Natl. Compr. Canc. Netw. *21*, 236–246. https://doi.org/10.6004/jnccn.2023.0014.
39. Noseworthy, P.A., Kaufman, E.S., Chen, L.Y., Chung, M.K., Elkind, M.S.V., Joglar, J.A., Leal, M.A., McCabe, P.J., Pokorney, S.D., Yao, X.; and American Heart Association Council on Clinical Cardiology Electrocardiography and Arrhythmias Committee; Council on Arteriosclerosis, Thrombosis and Vascular Biology; Council on Cardiovascular and Stroke Nursing; and Stroke Council (2019). Subclinical and Device-Detected Atrial Fibrillation: Pondering the Knowledge Gap: A Scientific Statement From the American Heart Association. Circulation *140*, e944–e963. https://doi.org/10.1161/CIR.0000000000000740.

40. Pitini, E., De Vito, C., Marzuillo, C., D'Andrea, E., Rosso, A., Federici, A., Di Maria, E., and Villari, P. (2018). How is genetic testing evaluated? A systematic review of the literature. Eur. J. Hum. Genet. *26*, 605–615. https://doi.org/10.1038/s41431-018-0095-5.
41. Walcott, S.E., Miller, F.A., Dunsmore, K., Lazor, T., Feldman, B.M., and Hayeems, R.Z. (2021). Measuring clinical utility in the context of genetic testing: a scoping review. Eur. J. Hum. Genet. *29*, 378–386. https://doi.org/10.1038/s41431-020-00744-2.
42. Goddard, K.A.B., Lee, K., Buchanan, A.H., Powell, B.C., and Hunter, J.E. (2022). Establishing the Medical Actionability of Genomic Variants. Annu. Rev. Genomics Hum. Genet. *23*, 173–192. https://doi.org/10.1146/annurev-genom-111021-032401.
43. Berg, J.S., Foreman, A.K.M., O'Daniel, J.M., Booker, J.K., Boshe, L., Carey, T., Crooks, K.R., Jensen, B.C., Juengst, E.T., Lee, K., et al. (2016). A semiquantitative metric for evaluating clinical actionability of incidental or secondary findings from genome-scale sequencing. Genet. Med. *18*, 467–475. https://doi.org/10.1038/gim.2015.104.
44. Hunter, J.E., Irving, S.A., Biesecker, L.G., Buchanan, A., Jensen, B., Lee, K., Martin, C.L., Milko, L., Muessig, K., Niehaus, A.D., et al. (2016). A standardized, evidence-based protocol to assess clinical actionability of genetic disorders associated with genomic variation. Genet. Med. *18*, 1258–1268. https://doi.org/10.1038/gim.2016.40.
45. ACMG Board of Directors (2015). Clinical utility of genetic and genomic services: a position statement of the American College of Medical Genetics and Genomics. Genet. Med. *17*, 505–507. https://doi.org/10.1038/gim.2015.41.
46. Burke, W. (2014). Genetic tests: clinical validity and clinical utility. Curr. Protoc. Hum. Genet. *81*, 9.15.1–9.15.8. https://doi.org/10.1002/0471142905.hg0915s81.
47. Sackett, D.L., Rosenberg, W.M., Gray, J.A., Haynes, R.B., and Richardson, W.S. (1996). Evidence based medicine: what it is and what it isn't. BMJ *312*, 71–72.
48. Elwenspoek, M.M.C., Scott, L.J., Alsop, K., Patel, R., Watson, J.C., Mann, E., and Whiting, P. (2020). What methods are being used to create an evidence base on the use of laboratory tests to monitor long-term conditions in primary care? A scoping review. Fam. Pract. *37*, 845–853. https://doi.org/10.1093/fampra/cmaa074.
49. Ford, I., and Norrie, J. (2016). Pragmatic Trials. N. Engl. J. Med. *375*, 454–463. https://doi.org/10.1056/NEJMra1510059.
50. Loudon, K., Treweek, S., Sullivan, F., Donnan, P., Thorpe, K.E., and Zwarenstein, M. (2015). The PRECIS-2 tool: designing trials that are fit for purpose. BMJ *350*, h2147. https://doi.org/10.1136/bmj.h2147.
51. Duncan, L., Shen, H., Gelaye, B., Meijsen, J., Ressler, K., Feldman, M., Peterson, R., and Domingue, B. (2019). Analysis of polygenic risk score usage and performance in diverse human populations. Nat. Commun. *10*, 3328. https://doi.org/10.1038/s41467-019-11112-0.
52. Using Population Descriptors in Genetics and Genomics Research: A New Framework for an Evolving Field (2023). (National Academies Press) https://doi.org/10.17226/26902.
53. Peterson, K., Anderson, J., Boundy, E., Ferguson, L., McCleery, E., and Waldrip, K. (2018). Mortality Disparities in Racial/Ethnic Minority Groups in the Veterans Health Administration: An Evidence Review and Map. Am. J. Public Health *108*, e1–e11. https://doi.org/10.2105/AJPH.2017.304246.
54. Wong, M.S., Hoggatt, K.J., Steers, W.N., Frayne, S.M., Huynh, A.K., Yano, E.M., Saechao, F.S., Ziaeian, B., and Washington, D.L. (2019). Racial/Ethnic Disparities in Mortality Across the Veterans Health Administration. Health Equity *3*, 99–108. https://doi.org/10.1089/heq.2018.0086.
55. Alexander, N.V.J., Brunette, C.A., Guardino, E.T., Yi, T., Kerman, B.J., MacIsaac, K., Harris, E.J., Antwi, A.A., and Vassy, J.L. (2022). Performance of EHR classifiers for patient eligibility in a clinical trial of precision screening. Contemp. Clin. Trials *121*, 106926. https://doi.org/10.1016/j.cct.2022.106926.
56. Kamiran, F., and Calders, T. (2012). Data preprocessing techniques for classification without discrimination. Knowl. Inf. Syst. *33*, 1–33. https://doi.org/10.1007/s10115-011-0463-8.
57. Kind, A.J.H., and Buckingham, W.R. (2018). Making Neighborhood-Disadvantage Metrics Accessible - The Neighborhood Atlas. N. Engl. J. Med. *378*, 2456–2458. https://doi.org/10.1056/NEJMp1802313.
58. Gonsoulin, M. (2018). VIReC Factbook: Corporate Data Warehouse (CDW) Patient 3.1 Domain (US Department of Veteran Affairs).
59. Brockman, T.A., Shaw, O., Wiepert, L., Nguyen, Q.A., Kelpin, S.S., West, I., Albertie, M., Williams, S., Abbenyi, A., Stephenson, N., et al. (2023). Community engagement strategies to promote recruitment and participation in clinical research among rural communities: A narrative review. J. Clin. Transl. Sci. *7*, e84. https://doi.org/10.1017/cts.2023.16.
60. Weeks, W.B., Wallace, A.E., West, A.N., Heady, H.R., and Hawthorne, K. (2008). Research on Rural Veterans: An Analysis of the Literature. J. Rural Health *24*, 337–344. https://doi.org/10.1111/j.1748-0361.2008.00179.x.
61. Fain, K.M., Nelson, J.T., Tse, T., and Williams, R.J. (2021). Race and ethnicity reporting for clinical trials in ClinicalTrials.gov and publications. Contemp. Clin. Trials *101*, 106237. https://doi.org/10.1016/j.cct.2020.106237.
62. Homburger, J.R., Neben, C.L., Mishne, G., Zhou, A.Y., Kathiresan, S., and Khera, A.V. (2019). Low coverage whole genome sequencing enables accurate assessment of common variants and calculation of genome-wide polygenic scores. Genome Med. *11*, 74. https://doi.org/10.1186/s13073-019-0682-2.
63. Chen, S.-F., Dias, R., Evans, D., Salfati, E.L., Liu, S., Wineinger, N.E., and Torkamani, A. (2020). Genotype imputation and variability in polygenic risk score estimation. Genome Med. *12*, 100. https://doi.org/10.1186/s13073-020-00801-x.
64. Hughes, E., Tshiaba, P., Gallagher, S., Wagner, S., Judkins, T., Roa, B., Rosenthal, E., Domchek, S., Garber, J., Lancaster, J., et al. (2020). Development and Validation of a Clinical Polygenic Risk Score to Predict Breast Cancer Risk. JCO Precis. Oncol. *4*, 585–592. https://doi.org/10.1200/PO.19.00360.
65. Grant, R.W., O'Brien, K.E., Waxler, J.L., Vassy, J.L., Delahanty, L.M., Bissett, L.G., Green, R.C., Stember, K.G., Guiducci, C., Park, E.R., et al. (2013). Personalized genetic risk counseling to motivate diabetes prevention: a randomized trial. Diabetes Care *36*, 13–19. https://doi.org/10.2337/dc12-0884.
66. Kullo, I.J., Jouni, H., Austin, E.E., Brown, S.A., Kruisselbrink, T.M., Isseh, I.N., Haddad, R.A., Marroush, T.S., Shameer, K., Olson, J.E., et al. (2016). Incorporating a Genetic Risk Score into Coronary Heart Disease Risk Estimates: Effect on LDL Cholesterol Levels (the MIGENES Clinical Trial). Circulation *133*, 1181–1188. https://doi.org/10.1161/CIRCULATIONAHA.115.020109.
67. Turner, A.R., Lane, B.R., Rogers, D., Lipkus, I., Weaver, K., Danhauer, S.C., Zhang, Z., Hsu, F.C., Noyes, S.L., Adams, T., et al. (2016). Randomized trial finds that prostate cancer genetic risk score feedback targets prostate-specific antigen screening

among at-risk men. Cancer *122*, 3564–3575. https://doi.org/10.1002/cncr.30162.

68. Smit, A.K., Allen, M., Beswick, B., Butow, P., Dawkins, H., Dobbinson, S.J., Dunlop, K.L., Espinoza, D., Fenton, G., Kanetsky, P.A., et al. (2021). Impact of personal genomic risk information on melanoma prevention behaviors and psychological outcomes: a randomized controlled trial. Genet. Med. *23*, 2394–2403. https://doi.org/10.1038/s41436-021-01292-w.
69. Gray, M.P., Berman, Y., Bottà, G., Grieve, S.M., Ho, A., Hu, J., Hyun, K., Ingles, J., Jennings, G., Kilov, G., et al. (2023). Incorporating a polygenic risk score-triaged coronary calcium score into cardiovascular disease examinations to identify subclinical coronary artery disease (ESCALATE): Protocol for a prospective, nonrandomized implementation trial. Am. Heart J. *264*, 163–173. https://doi.org/10.1016/j.ahj.2023.06.009.
70. Esserman, L.J.; and WISDOM Study and Athena Investigators (2017). The WISDOM Study: breaking the deadlock in the breast cancer screening debate. NPJ Breast Cancer *3*, 34. https://doi.org/10.1038/s41523-017-0035-5.

The American Journal of Human Genetics, Volume *110*

Supplemental information

The GenoVA study: Equitable implementation of a pragmatic randomized trial of polygenic-risk scoring in primary care

Jason L. Vassy, Charles A. Brunette, Matthew S. Lebo, Katharine MacIsaac, Thomas Yi, Morgan E. Danowski, Nicholas V.J. Alexander, Mark P. Cardellino, Kurt D. Christensen, Manish Gala, Robert C. Green, Elizabeth Harris, Natalie E. Jones, Benjamin J. Kerman, Peter Kraft, Preetika Kulkarni, Anna C.F. Lewis, Steven A. Lubitz, Pradeep Natarajan, and Ashley A. Antwi

PARTNERS HEALTHCARE | PERSONALIZED MEDICINE
MASSACHUSETTS GENERAL HOSPITAL | BRIGHAM AND WOMEN'S HOSPITAL

LABORATORY FOR MOLECULAR MEDICINE
65 Landsdowne St, Cambridge, MA 02139
Phone: (617) 768-8500 / Fax: (617) 768-8513
http://www.partners.org/personalizedmedicine/lmm

Name: LAST, FIRST
DOB: MM/DD/YYYY
Sex: Female/Male
Family #: F000000

MRN: XXXXX
Referring Facility: XXXXX
Referring Physician: XXXXX

LMM Accession ID: PM-23-X00000
Specimen: Blood, Peripheral
Received: 09/07/2022

Test Performed: GenoVA Polygenic Risk Assessment
Test Codes: lmGenoVA-b_L

RESULTS SUMMARY*

HIGH POLYGENIC DISEASE RISK: Genotyping indicated an increased polygenic risk for developing coronary artery disease. Result details are provided below.

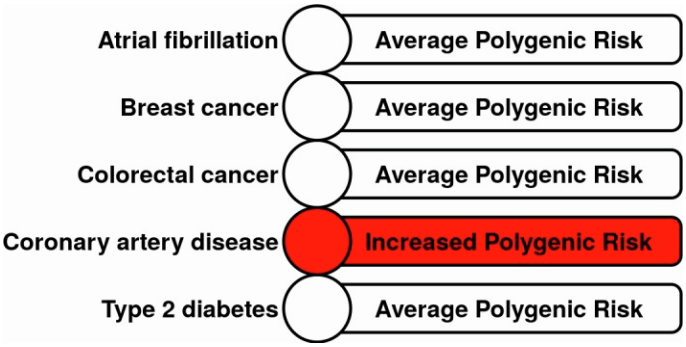
***Polygenic risk calculated using data from predominantly European ancestry individuals. Results are known to be less accurate for individuals of non-European ancestry. See details below.**

DETAILED GENOMIC RESULTS

A. POLYGENIC DISEASE RISK

Polygenic risk describes the chance of developing certain health conditions based on a large number of genetic variants across the genome. This test assessed the risk for developing the following conditions: atrial fibrillation, colorectal cancer, coronary artery disease, prostate cancer and type II diabetes.

This test identified an increased polygenic risk for coronary artery disease (see methodology for complete description of the analysis). It did NOT indicate increased polygenic risk for the remaining conditions.



Diseases WITH an increased polygenic risk

Disease	This patient's result	Lifetime Risk
Coronary artery disease	Increased polygenic risk	1 in 5 men aged 60-79 1 in 8 women aged 60-79
Risk Interpretation: The patient's calculated polygenic risk score, derived from 6630016 loci, has been associated with an INCREASED risk for coronary artery disease, defined here as greater than 2-fold risk. Individuals with similar polygenic risk scores have been shown to have an increased risk for coronary artery disease above baseline. Values of this polygenic risk score that fall among the top 10% were associated with a greater than 2-fold greater risk of developing coronary artery disease among >400,000 British volunteers of European ancestry when compared to the average individual (Khera 2018 PMID: 30104762). Having an ancestry-adjusted score in the top 5th percentile has been associated with an odds ratio of early myocardial infarction (before age 55) of 5.09, 2.02, 3.38, and 3.33 in people of white, black, Hispanic, and Asian ancestry, respectively (Khera 2019 PMID: 30586733).		
Disease Information: Coronary artery disease (CAD) is the most common type of heart disease in the United States, caused by plaque buildup in the walls of the coronary arteries, which supply blood to the heart. Risk of developing CAD increases with age. Symptoms of CAD include chest pain (angina), weakness, light-headedness, nausea, pain or discomfort in the arms or shoulder, shortness of breath, and heart attack (adapted from Centers for Disease Control and Prevention https://www.cdc.gov/heartdisease/coronary_ad.htm).		

Figure S1. Example polygenic risk score report from the GenoVA Study. All GenoVA Study participants receive a report indicating on the first page any actionable monogenic disease risk variants identified from their genotype data, if applicable, and the PRS risk interpretation for each of 5 diseases. High-risk PRS is defined as one associated with a published odds ratio >2 for the target disease, compared to the median PRS value. Technical details about the genotyping assay and analysis; more detailed description about the risk model, analysis, and interpretations; references to relevant publications; and limitations are included in subsequent pages of the report.

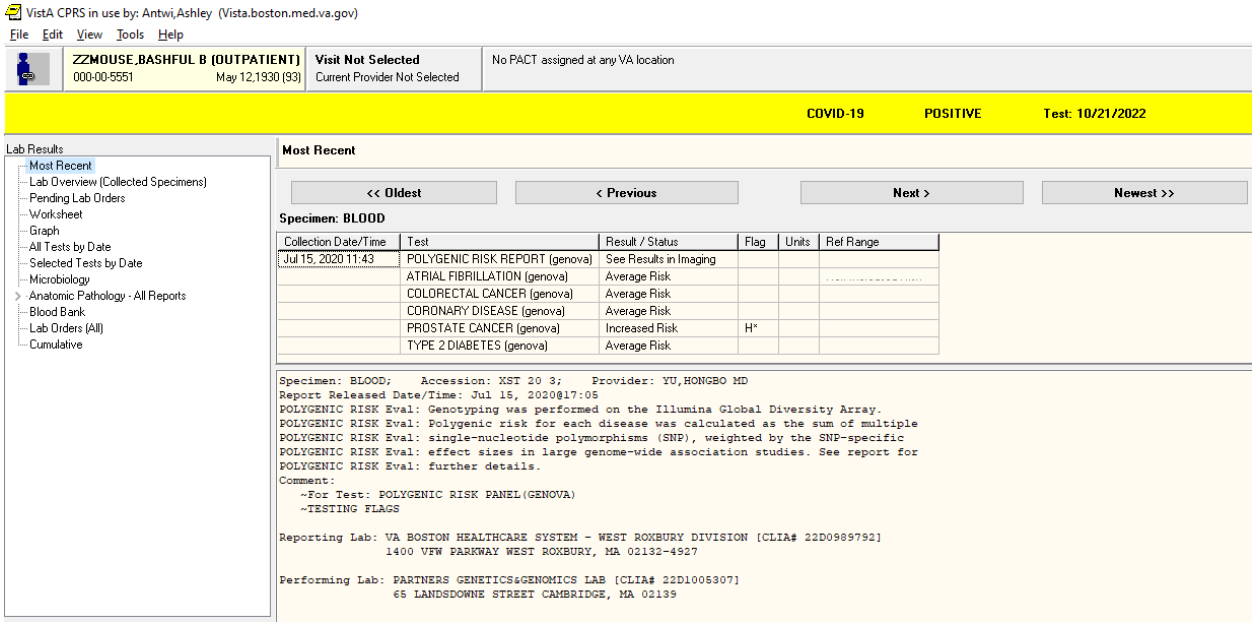


Figure S2. Polygenic risk score results in electronic health record. The figure shows the representation of PRS categories (increased risk vs. average risk) as structured data in the VA Boston Healthcare System electronic health record (EHR). The full PRS report is additionally uploaded as a .PDF report in the EHR.

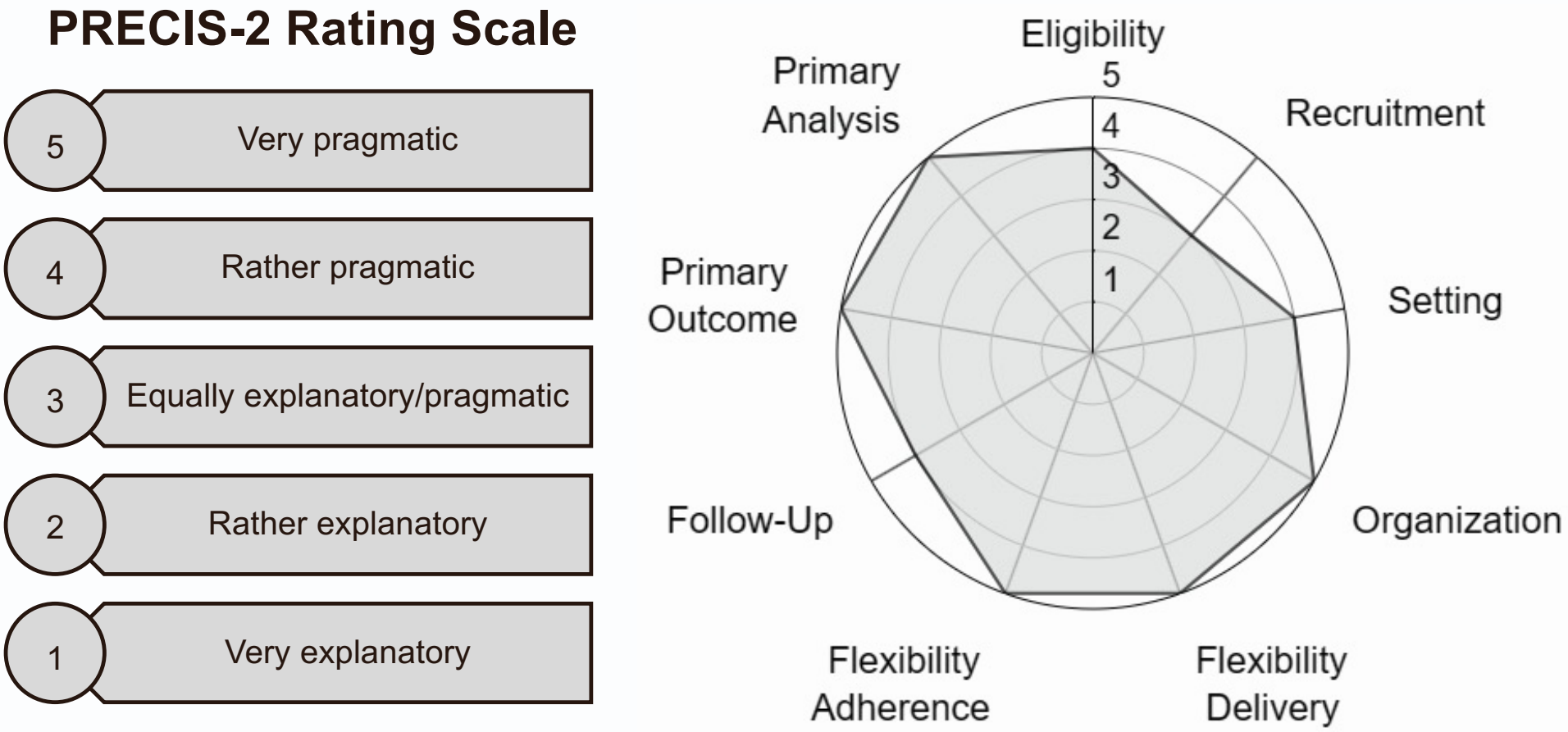


Figure S3. PRagmatic-Explanatory Continuum Indicator Summary 2 (PRECIS-2) rating scale (left) and mapping of the Genomic Medicine at VA (GenoVA) Study design to the PRECIS-2 wheel (right). Clinical trial elements (see Table S3) are rated on a scale from 1 (very explanatory) to 5 (very pragmatic).

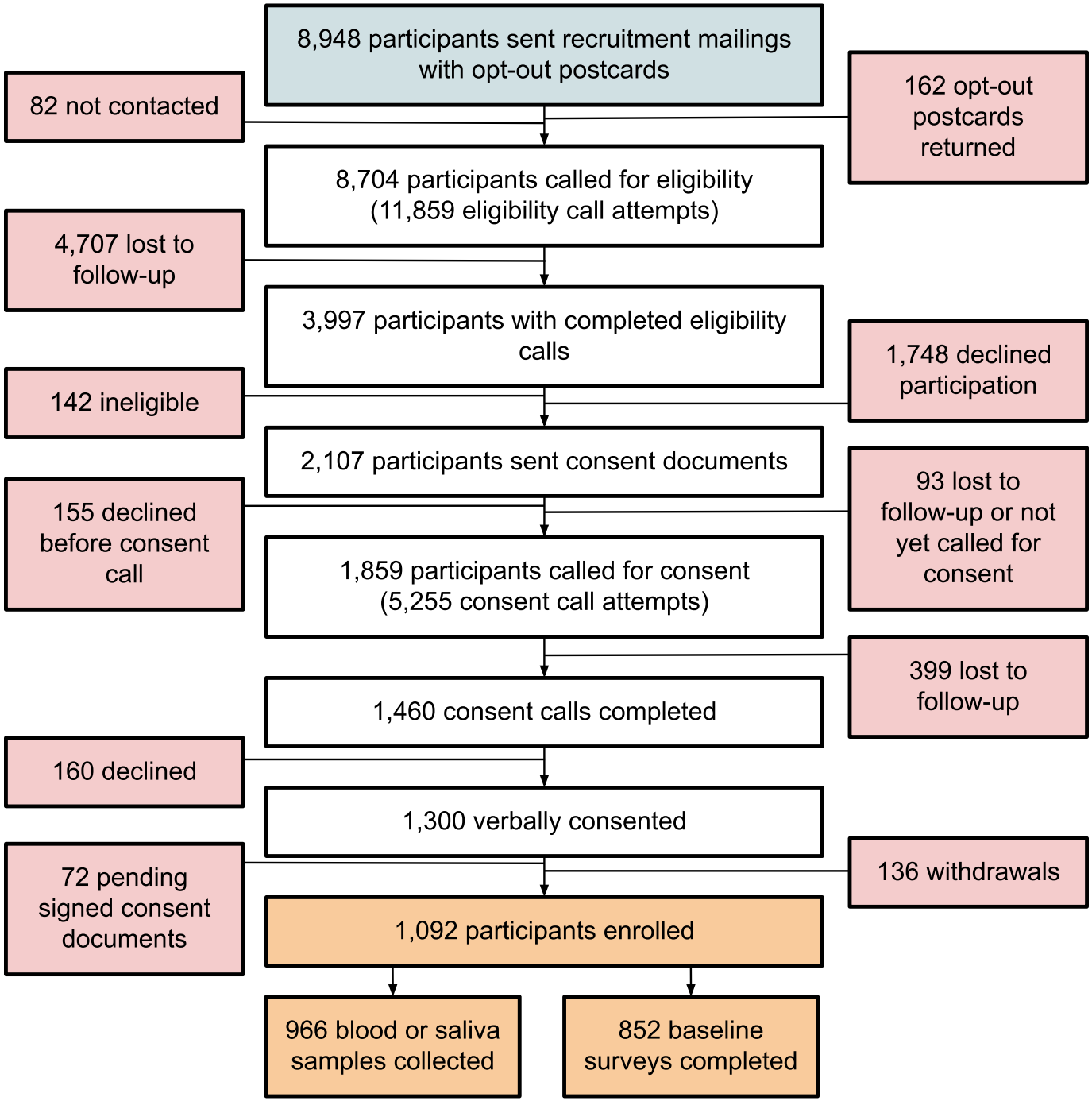


Figure S4. GenoVA Enrollment Diagram. Shown are the outcomes from 8,948 apparently eligible VABHS patients recruited for study participation. Recruitment enrollment, and biospecimen collection activities are ongoing until the prespecified sample size of 1,076 enrollees with biospecimens is reached

Eligibility criteria	Ascertainment methods
1.) Age: 50-70 years old	CDW: Age between 50-69 at time of recruitment
2.) Actively receiving primary care at any VABHS location	CDW: Determined through database of practicing VABHS clinicians, defined as having a VABHS primary care provider and had at least one clinical care visit or admission at a VABHS facility within the prior 12 months
3.) No known diagnosis of the following six conditions:	
1.) Coronary artery disease	CDW: ICD-9, ICD-10 codes, ICD-9 procedure codes, or CPT codes
2.) Colorectal cancer	CDW: ICD-9, ICD-10 codes, or ICD-9 procedure codes
3.) Atrial fibrillation	CDW: ICD-9, ICD-10 codes, or ICD-9 procedure codes
4.) Type 2 diabetes	CDW: ICD-9, ICD-10 codes, or use of medication to treat diabetes mellitus
5.) Breast cancer	CDW: ICD-9, ICD-10 codes, ICD-9 procedure codes, or CPT codes
6.) Prostate cancer	CDW: ICD-9, ICD-10 codes, ICD-9 procedure codes, or CPT codes

Table S1. GenoVA Study eligibility criteria and ascertainment methods. Abbreviations: CDW, corporate data warehouse; CPT, Current Procedural Terminology®; ICD, International Classification of Diseases; VABHS, VA Boston Healthcare System

Measure	Data source(s)	Description
Primary outcome		
Time-to-new diagnosis of common complex disease	CDW, expert chart review of VA and external records	A composite outcome of only clinically significant diagnoses of the 6 target conditions, as adjudicated by expert clinical chart review using gold-standard diagnostic criteria
Secondary outcomes		
Diagnostic testing	CDW, expert chart review of VA and external records	Any evidence that the participant underwent diagnostic testing for the 6 target diseases after enrollment: stress testing, cardiac CT for coronary artery calcium, coronary angiography, ECG, heart rhythm monitoring, hemoglobin A1c, blood glucose, colonoscopy, sigmoidoscopy, fecal blood testing, CT colonography, mammography, breast MRI, breast ultrasound, breast biopsy, PSA testing, prostate biopsy
Patient activation	Baseline and end-of-study surveys	Patient Activation Measure (13-item short form): Self-reported understanding, competence, and willingness to participate in health care decisions and processes
Healthcare costs	Billing and administrative data; empiric estimates of infrastructure and personnel costs	Combination of administrative data, microcosting approaches and empiric estimates to estimate costs of intervention and subsequent healthcare costs during 24 months after enrollment
Medication adherence	Baseline and end-of-study surveys	Voils Medication Adherence Survey: Self-report of taking medications as prescribed assessed on the baseline and end-of-study surveys
Exploratory outcomes		
Blood pressure	CDW	Most recent systolic and diastolic blood pressure values prior to or on date of enrollment and prior to or on the date 24 months after enrollment
Body-mass index	CDW	Most recent BMI values prior to or on the date of enrollment and prior to or on the date 24 months after enrollment
Aspirin use	Baseline and end-of-study surveys	Self-reported use of prescription or over-the-counter aspirin
Physical activity	Baseline and end-of-study surveys	Self-reported physical activity on single item with ordinal Likert scale
Alcohol intake	Baseline and end-of-study surveys	Self-reported alcohol intake on single item with ordinal Likert scale
Processed meat consumption	Baseline and end-of-study surveys	Self-reported processed meat consumption on single item with ordinal Likert scale
Low-density lipoprotein cholesterol (LDL-C)	CDW	Most recent LDL-C values recorded in the medical record prior to or on the date of enrollment and prior to or on the date 24 months after enrollment
Smoking status	Baseline and end-of-study surveys	Self-reported smoking status on the 5-item “Tobacco Use” instrument from the Behavioral Risk Factor Surveillance System (BRFSS,[1] Core Section 9)
Risk-reducing medication prescriptions	CDW, medical record review, baseline and end-of-study surveys	Relevant prescription medication changes, including antihypertensives, cholesterol-lowering medications, anticoagulants, antiplatelet medications, 5-alpha reductase inhibitors, selective estrogen receptor modulators, aromatase inhibitors
Health status and quality of life	Baseline and end-of-study surveys	Veterans RAND 12-item Health Survey (VR-12)[2]

Table S2. Primary, secondary, and exploratory outcomes and measures in the GenoVA Study. Abbreviations: CDW, corporate data warehouse; CT, computed tomography; ECG, electrocardiogram; PRS, polygenic risk score; PSA, prostate-specific antigen; VA, U.S. Department of Veterans Affairs; VR-12, Veterans RAND 12-item Health Survey.

Domain	Domain description	Assessment of pragmatism	Rationale for PRECIS-2 scoring of GenoVA Study	Score
Eligibility	Specifies inclusion and exclusion criteria for the trial and frames the target population(s) for which its results are intended to apply.	Are participants in the trial similar to those who would receive the intervention if it were available in usual care?	All patients between 50 and 70 years of age at enrollment with a VABHS primary care provider that have had at least one outpatient care visit or hospital admission at a VABHS facility within the past 12 months, and do not have a known diagnosis of any of the six target diseases (CAD, T2D, AFib, CRC, BrCA, PrCA).	4
Recruitment	Outlines the steps for the identification, consent, and enrollment of participants into the trial.	How much extra effort is made to recruit participants into the trial above what would occur in usual care?	The trial leverages a routine query of the VA CDW to identify potentially eligible patients for study participation, after which study staff mail a recruitment letter and opt-out postcard. Following a 10 day wait period, eligibility is confirmed via telephone call and patients deemed eligible are sent informed consent documents to review, complete, and return via e-mail or postal mail. Additionally, participants must provide a saliva or blood specimen to fully engage in the trial.	3
Setting	Context under which the trial is carried out, including factors such as geographic location and clinical infrastructure of the study site(s).	How different is the setting of the trial and the usual care setting?	The context under which the trial intervention is administered is similar to the context under which routine care is delivered across the VABHS healthcare system, with only slightly greater engagement by the research team for enrollment, specimen collection, and follow-up than what occurs routinely. Moreover, the resources, clinical infrastructure, and reach of primary care services at VABHS are comparable to those found in other similarly sized healthcare systems.	4
Organization	Structure and delivery of the intervention, including the clinical resources required to provide the intervention.	How different are the resources, provider expertise, and organization of care delivery in the intervention arm of the trial and usual care?	The use of an external reference laboratory, return of an interpreted laboratory report to patients and their PCPs either through standard mail or encrypted email, provision of support and reference materials, and inclusion of the result in the EHR are typical for any specialized laboratory testing at VABHS. No specialized training is provided to PCPs beyond the provision of reference materials.	5
Flexibility in delivery	How the trial intervention is delivered to study participants.	How different is the flexibility in how the intervention is delivered and the flexibility likely in usual care?	Reports and supporting materials are provided to patient-participants and PCPs similar to any specialized laboratory result, but their use within the routine medical care of patient-participants is not protocolized.	5
Flexibility in adherence	How closely study participants are monitored for compliance to the trial intervention and the measures used to maintain or improve adherence.	How different is the flexibility in how participants must adhere to the intervention and the flexibility likely in usual care?	Participant adherence to the study intervention is not monitored or required. Patient-participants and their PCPs are provided an intervention package containing an interpreted PRS report and supporting materials, but are not obligated to adhere to recommendations.	5
Follow-up	The rigor of measurement and amount of contact between the study staff and trial participants for the purposes of event tracking and data collection.	How different is the intensity of follow-up of participants in the trial and the likely follow-up in usual care?	The intensity of participant follow-up is minimally greater than what might occur in usual care. Most outcomes will be assessed through CDW and chart review. Study staff contact participants at enrollment and after 24 months for baseline and end-of-study surveys, respectively, for collection of patient-reported outcomes.	4
Primary outcome	The main variable to be measured for use in assessing the effect of the study intervention.	To what extent is the trial's primary outcome relevant to participants?	The primary outcome is time-to-diagnosis of at least one of 6 common complex diseases, and is assessed 24 months post-randomization for high-risk participants. Disease diagnosis is highly relevant to patient-participants and the future course of their medical care.	5
Primary analysis	The approach used for the analysis of final results.	To what extent are all data included in the analysis of the primary outcome?	The primary outcome will be analyzed using an intention-to-treat approach. The high risk PRS and usual care arms are compared with respect to time-to-diagnosis of at least one of the six common complex diseases.	5

Table S3. PRagmatic-Explanatory Continuum Indicator Summary 2 (PRECIS-2) domains and pragmatism assessment for the design of the GenoVA Study.

Domain scores range from 1 (very explanatory) to 5 (very pragmatic).[3,4] Abbreviations: AFib, atrial fibrillation; BrCA, breast cancer; CAD, coronary artery disease; CDW, Corporate Data Warehouse; CRC, colorectal cancer; PRS, polygenic risk score; PCP, primary care provider; PrCA, prostate cancer, T2D, type 2 diabetes; UC, usual care; VABHS, Veterans Affairs Boston Healthcare System.

	n (%)	Expected acceptance	Observed acceptance	Expected decline	Observed decline	p[a], effect size[b]
Completed eligibility call and accepted or declined receipt of consent documents[c]	3855 (100%)		2107 (54.7%)		1748 (45.3%)	
Gender						
Female	509 (13.2%)	7.2%	326 (8.5%)	6.0%	183 (4.7%)	<0.001, 0.072
Male	3346 (86.8%)	47.4%	1781 (46.2%)	39.4%	1565 (40.6%)	
Race/ethnicity						
Non-white or Hispanic	752 (19.5%)	10.7%	478 (12.4%)	8.8%	274 (7.1%)	<0.001, 0.087
White and Non-Hispanic	3098 (80.5%)	44.0%	1629 (42.3%)	36.5%	1474 (38.2%)	
Rurality[d]						
Rural	218 (5.7%)	3.1%	128 (3.3%)	2.6%	90 (2.3)	0.243, 0.012
Urban	3635 (94.3%)	51.5%	1978 (51.3%)	42.8%	1657 (43.0%)	
State-level ADI rank (1-10)[e]						
Least deprived (1-3)	834 (21.6%)	11.8%	475 (12.3%)	9.8%	359 (9.3%)	0.241, 0.015
Moderately deprived (4-6)	1489 (38.6%)	21.1%	794 (20.6%)	17.5%	695 (18.0%)	
Most deprived (7-10)	1477 (38.3%)	21.0%	809 (21.0%)	17.4%	668 (17.3%)	
	n (%)	**Expected enrolled**	**Observed enrolled**	**Expected not enrolled**	**Observed not enrolled**	
Received consent materials by mail or email and returned or did not return signed documents	2107 (100%)		1092 (51.8%)		1015 (48.2%)	
Gender						
Female	326 (15.5%)	8.0%	186 (8.8%)	7.5%	140 (6.6%)	0.046, 0.039
Male	1781 (84.5%)	43.8%	906 (43.0%)	40.7%	875 (41.5%)	
Race/ethnicity						
Non-white or Hispanic	478 (22.7%)	11.8%	201 (9.5%)	10.9%	277 (13.1%)	<0.001, 0.104
White and Non-Hispanic	1629 (77.3%)	40.1%	891 (42.3%)	37.2%	738 (35.0%)	
Rurality[f]						
Rural	128 (6.1%)	3.1%	54 (2.6%)	2.9%	74 (3.5%)	0.031, 0.044
Urban	1978 (93.9%)	48.6%	1037 (49.2%)	45.2%	941 (44.7%)	
State-level ADI rank (1-10)[g]						
Least deprived (1-3)	475 (22.5%)	11.7%	259 (12.3%)	10.9%	216 (10.3%)	0.217, 0.023
Moderately deprived (4-6)	794 (37.7%)	19.5%	414 (19.6%)	18.2%	380 (18.0%)	
Most deprived (7-10)	809 (38.4%)	19.8%	401 (19.0%)	18.6%	408 (19.4%)	

Table S4. Expected and observed recruitment and enrollment outcomes

Expected rates for each recruitment outcome computed using underlying demographic group memberships [*e.g.*, total proportion of sex (male, female) multiplied by total proportions of agreement to receive a consent packet (accept, decline)].

[a]*p*-values derived from between group comparisons using Pearson's Chi-square test for categorical data.
[b]Effect size estimates between groups presented as Cramer's V (equivalent to the *phi* coefficient for 2x2 contingency tables; ≤ 0.2 indicates a relatively weak association) for categorical data.
[c]Does not include individuals determined ineligible after completion of eligibility screen (n=142).
[d]Two individuals have undesignated rural status.
[e]55 participants with suppression due to low PH, high GQ, or both or QDI or missing ADI rank.
[f]One individual has undesignated rural status.
[g]29 participants with suppression due to low PH, high GQ, or both or QDI or missing ADI rank.
Abbreviations: ADI, Area Deprivation Index; GQ, group quarters; PH, population and/or housing; QDI, questionable data integrity.

Gene transcript	Variant(s)	Classification	Disease	High-risk PRS results
BRCA1 NM_007294.3	c.2748delT (p.Asn916Lysfs*84)	Likely pathogenic	Hereditary breast and ovarian cancer	None
BRCA2 NM_000059.3	c.3545_3546delTT (p.Phe1182*)	Pathogenic	Hereditary breast and ovarian cancer	None
BTD NM_000060.2	c.1330G>C (p.Asp444His)	Pathogenic	Biotinidase deficiency	None
CACNA1S NM_000069.2	c.3256C>A (p.Arg1086Ser)	Likely pathogenic	Malignant hyperthermia	None
HFE[a] NM_000410.3	c.845G>A (p.Cys282Tyr)	Established risk allele	Hereditary hemochromatosis	CAD
HFE[a] NM_000410.3	c.845G>A (p.Cys282Tyr)	Established risk allele	Hereditary hemochromatosis	None
HFE[a] NM_000410.3	c.845G>A (p.Cys282Tyr)	Established risk allele	Hereditary hemochromatosis	None
KCNH2 NM_000238.3	c.1744C>T (p.Arg582Cys)	Likely pathogenic	Long QT syndrome	None
KCNQ1 NM_000218.2	c.1085A>G (p.Lys362Arg)	Likely pathogenic	Long QT syndrome	CAD, PrCa
LDLR NM_000527.4	c.1898G>A (p.Arg633His)	Likely pathogenic	Familial hypercholesterolemia	CAD
MSH6 NM_000179.2	c.845_846insT (p.Asp284Glyfs*2)	Pathogenic	Lynch syndrome	CRC
MSH6 NM_000179.2	c.3768T>G (p.Tyr1256*)	Pathogenic	Lynch syndrome	None
RYR1 NM_000540.2	c.7300G>A (p.Gly2434Arg)	Pathogenic	Malignant hyperthermia	None

Table S5. American College of Medical Genetics and Genomics actionable gene list for 13 GenoVA participants

[a]Homozygous for *HFE* c.845G>A. Abbreviations: CAD, coronary artery disease; CRC, colorectal cancer; PrCa, prostate cancer

	Breast cancer	Colorectal cancer	Prostate cancer	Atrial fibrillation	Coronary artery disease	Type 2 diabetes
Total *n*	165	840	675	840	840	840
High-risk PRS, *n* (%)	20 (12.1%)	50 (6.0%)	77 (11.4%)	70 (8.3%)	78 (9.3%)	75 (8.9%)

Table S6. Distribution of polygenic risk score results of the first 840 GenoVA Study enrollees with completed genetic analyses. Percentages indicate proportions of participants with a high-risk polygenic risk score (PRS) result for the given disease out of 840 participants with completed genetic analyses as of May 10, 2023 Only biologically female participants (*n*=165) were eligible to receive a breast cancer PRS, and only biologically male participants (*n*=675) were eligible to receive a prostate cancer PRS.

SUPPLEMENTAL REFERENCES

1. Behavioral Risk Factor Surveillance System (2009). Centers for Disease Control and Prevention. http://www.cdc.gov/brfss/annual_data/pdf-ques/2009brfss.pdf. Accessed September 23, 2023.
2. Selim, A.J., Rogers, W., Fleishman, J.A., Qian, S.X., Fincke, B.G., Rothendler, J.A., and Kazis, L.E. (2009). Updated U.S. population standard for the Veterans RAND 12-item Health Survey (VR-12). Qual Life Res *18*, 43–52. 10.1007/s11136-008-9418-2.
3. Ford, I., and Norrie, J. (2016). Pragmatic Trials. N. Engl. J. Med. *375*, 454–463. 10.1056/NEJMra1510059.
4. Loudon, K., Treweek, S., Sullivan, F., Donnan, P., Thorpe, K.E., and Zwarenstein, M. (2015). The PRECIS-2 tool: designing trials that are fit for purpose. BMJ *350*, h2147. 10.1136/bmj.h2147.

The GenoVA Study: Pragmatic randomized trial of polygenic risk scoring for common disease in primary care

Principal Investigator: Jason L Vassy, MD, MPH, SM
Phone: 857-364-2561
E-mail: Jason.vassy@va.gov

Protocol Version Number: 19
Protocol Version Date: February 17, 2023

Funding Mechanism: NIH

ClinicalTrials.gov number: NCT04331535

TABLE OF CONTENTS

1. Protocol Summary/Abstract

Objectives:	To determine the clinical effectiveness of polygenic risk score testing among patients at high genetic risk for at least one disease, measured by time-to-diagnosis of prevalent or incident disease over 24 months.
Research Design:	Pragmatic randomized controlled trial
Methodology	In this project, primary care patients without a known diagnosis of any one of 6 target diseases (coronary artery disease, atrial fibrillation, type 2 diabetes mellitus, breast cancer, colorectal cancer, and prostate cancer) will undergo clinical genotype testing to calculate their polygenic risk scores (PRS) for these 6 diseases. Patients will undergo stratified randomization to PRS reporting to their primary care provider (PCP) at baseline vs. usual care (UC). The stratum of patients with at least one PRS indicating high risk (odds ratio, OR_{PRS}, >2.0) will be randomized to have them and their PCP receive their high-PRS results at baseline (PRS-high arm) or after a 24-month observation period (UC-high arm), reported along with evidence-based screening and management recommendations. Similarly, the stratum of patients with no PRS indicating high genetic risk will be randomly allocated to having them and their PCPs receive their results at baseline (PRS-average arm) or after 24 months (UC-average arm). All patients will be observed for 24 months for the primary clinical outcome of time-to-diagnosis of any one of the 6 target diseases.
Clinical Implications:	The outcomes of this trial will inform whether and how polygenic risk scores should be incorporated into the routine practice of medicine.

2. Aims/Objectives

The objective of this study is to determine the clinical effectiveness of polygenic risk score testing among patients at high genetic risk for at least one disease, measured by time-to-diagnosis of prevalent or incident disease over 24 months (primary outcome). Secondary clinical outcomes will include changes in clinical management and patient health behaviors.

3. Background Information

One of the most pressing questions in genomics today is the clinical utility of polygenic risk scores (PRS). Broadening the scope of genomic risk testing beyond monogenic diseases, PRS combine information from hundreds or even millions of genetic loci (SNPs), each with a very small effect size on the risk of common complex disease. The result is a continuous quantitative risk factor for susceptibility to conditions such as coronary artery disease (CAD) and type 2 diabetes (T2D). Compared to rarer monogenic disease variants, PRS have greater transformative potential for public health and healthcare in their ability to identify much larger proportions of the population at significantly elevated risk for disease, potentially facilitating evidence-based prevention and management. Moreover, their prediction

ability has vastly improved over the last 5 years compared to earlier PRS that included only a limited number of genetic variants.

However, while the associations between PRS and a wide range of common diseases are well established (clinical validity), the potential impact of this information on patient health outcomes (clinical utility) remains contested and understudied. Proponents argue that, because PRS in the upper tails of the normal distribution confer an equivalent risk to rare variants associated with monogenic forms of disease, they should similarly impact clinical screening and prevention strategies. Opponents argue that PRS achieve similar discrimination for disease risk as other risk factors already used in clinical care (*e.g.* body-mass index and smoking) or readily available without additional testing (*e.g.* socioeconomic status). Nonetheless, the invariability of PRS over the entire life course and the possibility of deriving PRS for multiple conditions from a single, relatively inexpensive test make them attractive candidates for novel risk factors in an era of increasing access to genotyping.

Despite disagreement about the readiness of PRS for clinical use, there is more agreement that patient outcomes data are needed to demonstrate their clinical utility, ideally prospectively collected from real-world medical practice. It is also recognized that PRS alone will be insufficient to achieve improvements in patient health, if they lack actionability to facilitate their use. In this project, Dr. Vassy (PI) will extend his point-of-care pragmatic trial methodology to examine the clinical effectiveness of the use of PRS for 6 common diseases that are screened for by PCPs and have established prevention strategies: CAD, atrial fibrillation (AFib), T2D, colorectal cancer, prostate cancer, and breast cancer.

4. Rationale and Purpose

Laboratories and healthcare systems in the US and abroad are racing to bring PRS to patient care. The results of this trial, whether positive or null, provide critical outcomes data to inform whether and how PRS should be used in clinical medicine.

5. Relevance to Veterans Health

The conditions studied in this project are highly prevalent in the Veteran patient population, and interventions that would facilitate their prevention and early diagnosis and treatment could reduce their morbidity and associated costs.

6. Study Design

6.1 Brief Summary

This study is a point-of-care pragmatic randomized controlled trial of polygenic risk scores (PRS) and linked clinical recommendations compared to usual care (UC). As shown in Figure 1, enrolled patients without a known diagnosis of CAD, T2D, Afib, or colorectal, breast, and prostate cancer undergo genotyping for PRS for each of these conditions. Patients undergo stratified randomization to the PRS or UC arms. Patients with at least one PRS indicating high risk (odds ratio, OR_{PRS}, >2.0) are randomized to have them and their primary care providers (PCPs) receive their high-PRS results report at baseline (PRS-high arm) or after a 24-month observation period (UC-high arm). Providers will also receive evidence-based management recommendations. Similarly, the stratum of patients with no PRS indicating high genetic risk are randomly allocated to having them and their PCPs receive their results at baseline (PRS-

average arm) or after 24 months (UC-average arm). All patients are observed for 24 months for the primary clinical outcome of time-to-diagnosis of any one of the 6 target diseases. Any participant with an actionable genetic variant will be withdrawn from the RCT portion of the study and will instead be followed along with the concurrent control group, although their data will not be analyzed with the concurrent control group.

6.2 Study Sites

This study is conducted at all VABHS locations.

6.3 Study Design

This study is a pragmatic randomized controlled trial.

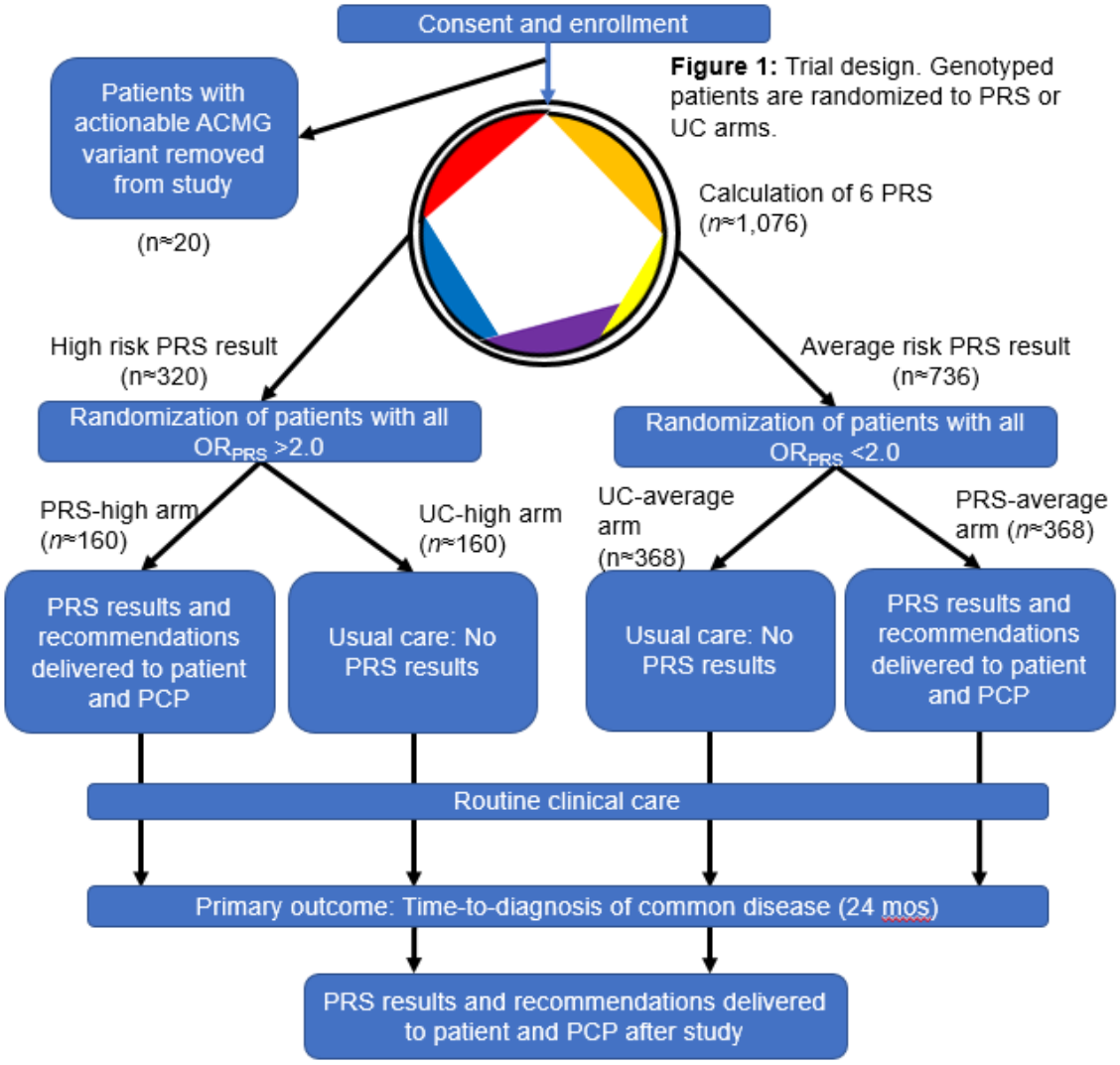


Figure 1: Trial design. Genotyped patients are randomized to PRS or UC arms.

6.4 Interventions

6.4.1. Usual care

In this study, usual care (UC) is defined as the routine medical care that patient-participants are already receiving from the primary care and other providers. This may include regular screenings and health maintenance activities such as recommended cancer screening, physical examinations, and counseling about health behaviors such as diet and exercise. These clinical activities are not protocolized by this study and are not a part of the research activities.

6.4.2 Intervention

The intervention in this study is the delivery of genetic susceptibility results to patient-participants and their primary care providers. Consented patient-participants will undergo SNP array genotyping, performed on either a saliva sample or a blood sample. An external CLIA-certified laboratory will perform the genotyping and calculate a PRS for 5 diseases for each patient-participant: CAD, Afib, T2D, colorectal cancer, and breast cancer (for women) or prostate cancer (for men). Participants undergo stratified randomization. Participants with a PRS odds ratio >2.0 for at least one of the target diseases are randomized to receive the study intervention at baseline (PRS-high arm) or after 24 months (UC-high arm). In the intervention, the participant receives a clinical report with their high-risk PRS result(s) via patient portal and/or letter, along with educational information about the disease(s) for which they have high genetic risk. Their primary care provider also receives the genetic report via the electronic health record and encrypted email, in addition to evidence-based recommendations for disease risk assessment, screening, and diagnosis. Similarly, the stratum of patients with no PRS indicating high genetic risk and their PCPs will receive their genetic reports at baseline (PRS-average arm) or after 24 months (UC-average arm). The genetic risk results report is also stored in the participant's medical record.

6.5 Study Procedures

6.5.1 Identification of eligible patients

Study staff perform a regular query of the VA Corporate Data Warehouse (CDW) to generate patient eligibility tables. This query assesses potential eligibility: age 50-70 years and absence of the ICD codes and other data indicating an established diagnosis of one of the 6 target diseases. Patient eligibility is further confirmed by review of the patient's electronic medical records (Computerized Patient Record System, CPRS).

6.5.2. Patient recruitment

Study staff mail recruitment materials to potentially eligible patients. This mailing includes a recruitment letter (see "GenoVA Recruitment Letter") and a pre-stamped, self-addressed postcard with a coded study participant ID giving the patients the ability to opt-out (see "GenoVA Opt-Out Postcard"). After at least ten days, study staff follow each letter with a telephone call to the potential participant to confirm interest in participating in the study and screen for eligibility (see "GenoVA Telephone Scripts"). If the patient confirms interest and is found to be eligible to participate in the study, study staff then sends a second mailing. This mailing includes a cover letter (see "GenoVA Cover Letter"), an Informed Consent form (see "GenoVA Informed Consent Form"), a HIPAA Authorization form (see "GenoVA HIPAA Authorization Form"), an Authorization to Release Health Information form (VA Form 10-5345) if participant opts to have results information sent to non-VA provider, and a pre-stamped, self-addressed envelope the patient will use to return the signed Informed Consent and HIPAA Authorization forms.

After at least 10 days, study staff will call the patient to review the Informed Consent and HIPAA authorization forms, answer any questions, obtain consent to enroll in the study, and ask the participant to sign and date the Informed Consent and HIPAA Authorization forms (and, if applicable, VA Form 10-5345) and return by mail in an included stamped envelope (see "GenoVA Telephone Scripts").

After the initial contact by standard mail and the opportunity to opt out of contact, patient-participants have the option to request that study staff send recruitment and other study materials (i.e. cover letter, Informed Consent form, HIPAA Authorization form, and VA Form 10-5345) via Azure RMS encrypted e-mail, DocuSign envelope (Azure RMS encrypted e-mail containing links to study documents for patient-participants to review and sign after study staff have obtained consent), or, alternatively, may schedule a remote visit via phone or video call with study staff. VA-approved technologies such as Webex, VA Video Connect (VVC), and Doximity will be used to conduct video calls with prospective participants. Before the remote visit, a copy of the Informed Consent and HIPAA Authorization forms, and, if applicable, VA Form 10-5345 will be sent to the prospective participant for their review. To help ensure that the forms are signed properly, signature lines will be flagged with an "X" and/or highlighted. After the consent information is reviewed with the prospective participant and study staff have addressed any questions or concerns, study staff will confirm the participant's willingness to participate and ask the participant to sign and date the consent documents. The participant may scan or take a photo of the signed signature page(s) and e-mail them back to the study staff using Azure RMS encrypted e-mail. Alternatively, during a videoconference call study staff can ask the participant to hold up the document so they can take a screen capture of the signature page using a VA-approved web-camera. Patient-participants may also opt to return their signed Informed Consent and HIPAA Authorization forms and VA Form 10-5345, if applicable, via Azure RMS encrypted e-mail or DocuSign.

Once study staff receive a participant's completed, signed and dated Informed Consent and HIPAA Authorization forms, they may call the participant and conduct the baseline telephone survey and schedule the biospecimen collection. Study staff may not access the participant's medical records, send genetic results to their provider, or collect any protected health information (PHI) or baseline data prior to receiving properly completed, signed and dated Informed Consent and HIPAA Authorization forms from participants. Once study staff receive participants' Informed Consent and HIPAA Authorization forms, they may scan or download the Informed Consent and HIPAA Authorization forms and store the scanned or downloaded pdfs in a secured folder behind the VA firewall. Hard copy versions of the Informed Consent and HIPAA Authorization forms will be shredded after they are scanned and stored to the secured folder. All participants may opt to have their results sent to a non-VA healthcare provider. Participants who opt to have their results sent to a non-VA provider will be sent a Request for and Authorization to Release Health Information form, which will be included in the second recruitment mailing, alongside the cover letter, and Informed Consent and HIPAA Authorization forms.

There are two additional methods for participant recruitment that do not use the opt-out postcard method (first mailing), as alternatives to the process described above:

1. Study staff utilize VABHS social media to recruit study participants (see "GenoVA Recruitment Flyer" and "GenoVA Social Media Outreach"). In response, VABHS patients may contact study staff in two ways:
    a. Interested Veterans may contact study staff directly via telephone call or e-mail after seeing the social media posts to inquire about the study. In this case, study staff will perform an eligibility screen with the patient and then send them the Informed Consent

and HIPAA Authorization forms, to be reviewed and signed with the study staff as described above.

b. Interested Veterans may click on a link provided in the social media post to access an eligibility screen and enrollment portal, hosted by Ipsos which has a VA authority to operate (ATO) to perform the following functions. An individual will access an eligibility screening survey asking them to confirm their Veteran status, age, their association with VA Boston, and the absence of the six diseases of interest. If survey responses indicate eligibility, the Veteran will be given the option to provide their name, phone number, and/or e-mail address for study staff to contact them about next steps.

2.

a. A VABHS provider may also refer a patient directly to study staff for recruitment, if he/she thinks the patient would be eligible and interested. To do so, the provider will obtain and document permission to send the patient's information to GenoVA study staff, and then send the patient's information and documentation of permission to share this information to the study staff by Azure RMS encrypted e-mail, or Microsoft Teams, or by adding study staff as an additional signer to a clinical note in the electronic health record. Study staff will screen the patient for eligibility via chart review prior to contacting the patient to perform an eligibility screen via telephone call following the "GenoVA Telephone Scripts." Then, if the patient is interested and requests more information, study staff will send them the Informed Consent and HIPAA Authorization forms, to be reviewed and signed with the study staff via telephone or approved teleconference method following a 10-day waiting period as described above.

b. If a provider refers a patient to study staff by sending patient information to study staff without documenting that they received permission from their patient to share this information, study staff will send the patient a recruitment letter and opt-out postcard and wait a period of 10 days before contacting the patient to gauge if they are eligible and interested in study participation before sending them the study's informed consent information.

6.5.3. Baseline telephone survey

After receiving the signed Informed Consent and HIPAA Authorization forms, study staff call the participant (see "GenoVA Telephone Scripts") to administer the baseline patient survey (see "GenoVA Baseline Survey") and schedule biospecimen collection.

6.5.4. Genotyping

Consented patient-participants undergo genome-wide genotyping, performed on either a mailed saliva sample or a blood sample obtained by phlebotomists at their local VABHS facility. The genotyping array includes millions of genetic markers, including those used to calculate the polygenic risk scores for the 6 target diseases and a small number of markers associated with medically actionable findings for 59 conditions (described below in 6.5.5.) Patient-participants may choose either method for specimen collection (saliva or blood) depending on personal preference and/or convenience. Samples are coded using unique study IDs and do not include patient identifiers before shipment to an external VA-approved Clinical Laboratory Improvement Amendments (CLIA)-certified clinical laboratory, which performs the genotyping and calculates the polygenic risk scores. This laboratory generates a clinical polygenic risk report for each participant and sends it to the study staff.

6.5.5. Incidental actionable findings

For patients undergoing clinical genome sequencing, the American College of Medical Genetics and Genomics (ACMG) currently recommends identifying and reporting incidental genetic findings for 59 conditions deemed medically actionable, primarily associated with cardiovascular disease and hereditary cancer syndromes (Kalia, *Genetics in Medicine* 2017). This number might increase over the course of the study. Although the GenoVA Study is not using genome sequencing, certain pathogenic and likely pathogenic variants in actionable ACMG genes are present on the genotype array the study is using. It is estimated that about 1-2% of individuals carry one of these variants. If the clinical laboratory identifies an actionable ACMG variant in a GenoVA Study participant's specimen, it confirms the finding with Sanger sequencing and report any confirmed result back to the study staff. Adapting processes for return of incidental results developed for the Mass General Brigham Biobank, a GenoVA Study genetic counselor (GC) calls the participant to notify them that a genetic result that may be important to their health has been identified. The participant may choose to receive the result or decline (see "GenoVA Return of Incidental Results Phone Scripts"). If the participant declines to receive the result, the GC collects the reason for declining. The GC will review this information with the PI, who will use clinical judgment and consultation with the IRB to decide about any further action. If the participant agrees to receive the result, the GC conducts a standard genetic counseling session including: collection of family history, description of the disease associated with the actionable finding, discussion of potential implications for family members, facilitation of appropriate clinical follow-up (see "GenoVA Return of Incidental Results Phone Scripts"), and a letter relaying what was discussed during the genetic counseling session (see "GenoVA Patient Incidental Result Letter"). The participant also has the option of having the GC share the result with the participant's primary care provider (see "GenoVA PCP Incidental Result Letter") and family members (see "GenoVA Family Incidental Result Letter"). Any participant with an actionable genetic variant will be withdrawn from the RCT portion of the study and will instead be followed along with the concurrent control group, although their data will not be analyzed together with the control group.

6.5.6. Randomization

Any participant with a confirmed actionable ACMG variant (estimated 1-2% of participants) is ineligible for subsequent stratified randomization and will instead be followed along with the concurrent control group. Among the remaining participants, those with at least one polygenic risk score (PRS) indicating and odds ratio (OR) >2.0 for any of the 6 target diseases are randomized to the PRS-high or usual care (UC)-high arm. Similarly, the stratum of patients with no PRS indicating high genetic risk will be randomly allocated to having them and their PCP receive their results at baseline (PRS-average arm) or after 24 months (UC-average arm). Study staff use pre-generated randomization tables to assign participants to a study arm. Participants with an actionable ACMG variant are observed as concurrent controls. Study staff use pre-generated randomization tables to assign participants to a study arm.

6.5.7. Delivery of intervention

Patients assigned to the PRS-high and PRS-average arms receive a copy of their PRS report via the online patient portal and by letter (see "GenoVA Patient Average-Risk Results Letter (Immediate Results)" and "GenoVA Patient High-Risk Results Letter (Immediate Results)"). Their VABHS primary care providers also receive the report via CPRS and encrypted email (see "GenoVA PCP Average-Risk Results Letter" and "GenoVA PCP High-Risk Results Letter"). For participants with high genetic results (PRS-high arm), a clinician member of the study team (MD or genetic counselor) will contact them by phone (see "GenoVA High-Risk Results Phone Scripts) prior to sending the patient and his/her provider(s) the patient and provider reports, accompanied by evidence-based recommendations for screening, prevention, and

diagnosis of the target conditions (see “Atrial Fibrillation Patient Information Sheet,” “Breast Cancer Patient Information Sheet,” “Colorectal Cancer Patient Information Sheet,” “Coronary Artery Disease Patient Information Sheet,” “Prostate Cancer Patient Information Sheet,” “Type 2 Diabetes Patient Information Sheet,” “Atrial Fibrillation Provider Information Sheet,” “Breast Cancer Provider Information Sheet,” “Colorectal Cancer Provider Information Sheet,” “Coronary Artery Disease Provider Information Sheet,” “Prostate Cancer Provider Information Sheet,” and “Type 2 Diabetes Provider Information Sheet”). Patient-participants have the option to request that study staff send their PRS results and accompanying disease information sheet(s), and if applicable any non-VA providers patient-participants have authorized to release their information to, via Azure RMS encrypted e-mail. If a clinician member of the study team is unable to contact a patient with a high-risk PRS result after three attempts, they will leave a voicemail message informing the patient that their GenoVA study results will be mailed to them requiring a signature of receipt of the study intervention package. A brief letter including contact information and describing why the patient is receiving this letter and their study intervention package will also be included in this mailing (see “GenoVA High-Risk PRS Unable to Contact Letter”). Genetic results reports are included in CPRS as a laboratory order that refers providers to a scanned pdf report in the patients’ medical record (VistA). Patients assigned to the UC-high and UC-average arms receive the same intervention at the end of the study (after 24 months), after completion of the end-of-study survey. All participants with an actionable genetic variant will receive a copy of their PRS report prior to withdrawal from the RCT. All participants may opt to have their results sent to a non-VA healthcare provider.

6.5.8. End-of-study survey

24 months after enrollment, study staff will e-mail a link to the end-of-study survey that patient-participants will use to access and complete the end-of-study survey electronically via Qualtrics or Ipsos, or, alternatively, study staff administer the end-of-study survey via telephone call to patient-participants who do not have an active e-mail address or otherwise prefer to complete the survey with study staff on the phone (see “GenoVA End-of-Study Survey_Online” and “GenoVA End-of-Study Survey_Telephone”). End-of-study survey data collected via Qualtrics or Ipsos will be stored in a VA Box account (also approved by VA OIT) for the purpose of serving as a centralized location in which study staff can clean, organize, extract, and analyze end-of-study survey data. The end-of-study survey also asks patient-participants whether they had a new diagnosis of any of the six diseases during the study period. For any affirmative response, if applicable, the study staff may ask participants to request that their relevant medical records from outside healthcare providers be sent to study staff for review and study staff will send a copy of VHA-FL-10-212 to fill and return to study staff by postal mail or e-mail to retrieve relevant medical records from outside healthcare providers.

6.5.9. End-of-study chart review

Clinical chart reviewers blinded to patient-participant randomization status independently review each patient's medical record for the 24 months after enrollment for any evidence that one of the target diseases has been diagnosed since enrollment. VA databases including the Corporate Data Warehouse (CDW) and HERC Managerial Cost Accounting (MCA) are also accessed for other clinical and economic study outcomes. Centers for Medicaid and Medicare (CMS) data may also be requested, to identify study outcomes occurring outside of VA.

6.5.10. End-of-study results reporting

Participants randomized to the UC-high and UC-average arms and their providers receive the study intervention after completion of the end-of-study survey (see “GenoVA Patient Average-Risk Results

Letter (Delayed Results),""GenoVA Patient High-Risk Results Letter (Delayed Results)," "GenoVA PCP Average-Risk Results Letter," and "GenoVA PCP High-Risk Results Letter").

7. Study Subject Selection

7.1 Sample Description

This study recruits patients actively receiving primary care at any VA Boston location.

7.2 Subject Inclusion Criteria

- Age 50-70 years at enrollment
- No known diagnosis of the following conditions, initially screened by the International Classification of Disease (ICD) codes and the EHR data and then confirmed with potential patient-participants during recruitment:
  - Coronary artery disease: ICD-9 Codes 410-414 or ICD-10 Codes I20-I25 or ICD-9 Procedure Codes 36, 00.66 or CPT Codes 33510-33536, 9292x, 9293x, 9294x, 92973, 92974, 92975
  - Atrial fibrillation: ICD-9 Codes 427.3 or ICD-10 Codes I48 or ICD-9 Procedure Codes 37.33, 37.34
  - Type 2 diabetes: ICD-9 Codes 250 or ICD-10 Codes E10-E11, E13 or use of medication to treat diabetes mellitus
  - Colorectal cancer: ICD-9 Codes 153, 154.0, 154.1, 159.0, 230.3, 230.4, V10.05, V10.06 or ICD-10 Codes C18, C19, C20, C26, D01.0, D01.1, D01.2, Z85.038, Z85.048 or ICD-9 Procedure Codes 17.31-17.36, 45.71-45.76, 45.81-45.83 or CPT Codes 44140-44160, 44204-44212
  - Breast cancer: ICD-9 Codes 174, 175, 233.0, V10.3 and ICD-10 Codes C50 - C50.9, D05, Z853 or ICD-9 Procedure Codes 85.20, 85.21, 40.22, 40.23, 85.22, 85.23, 85.33-85.36, 85.41-85.48 or CPT Codes 19120, 19125, 19126, 19160, 19162, 19180, 19182, 19200, 19220, 19240, 19300-19307
  - Prostate cancer: ICD-9 Codes 185, 233.4, V10.46 and ICD-10 Codes C61, D07.5, Z85.46 or ICD-9 Procedures Codes 60.21, 60.29, 60.3, 60.4, 60.5, 60.62, 60.69 or CPT codes 55801, 55810, 55812, 55815, 55821, 55831, 55840, 55842, 55845

7.3 Subject Exclusion Criteria

There are no other exclusion criteria for this study. The age range of 50-70 years was chosen after review of the current incidence rates of the six target diseases at VABHS. Patients younger than 50 have a low overall rate of diagnosis of at least one of the 6 diseases; per VABHS estimates between 2014 and 2017, patients younger than 50 had a 2% aggregate rate of diagnosis for these conditions. This low prevalence rate limits the adequate assessment of the hypotheses and interventions under study. The exclusion of patients over the age of 70 is due to the potentially decreased relevance of polygenic risk scoring (PRS) in this age range, as many evidence-based recommendations for disease screening and prevention are only limited to individuals younger than 70. If the objective of this study is to examine the impact of PRS testing on existing preventive care, it is important to target the relevant age range for this usual care. Thus, the age range between 50 and 70 balances both the prevalence of these conditions to detect a significant intervention effect and the clinical significance of PRS results in the determination of patient risk and the potentially positive impact on patient outcomes through the mechanisms of earlier disease detection and intervention.

7.4 Recruitment

Study staff perform a regular query of the CDW and CPRS to generate patient eligibility tables. Study staff then mail a recruitment letter to potentially eligible patients (see "GenoVA Recruitment Letter"). This letter introduces the study and gives the potential to opt out of further contact about the study through a coded postcard. After 10 days, if the potential participant has not returned an opt-out postcard, the staff call to assess interest and eligibility and then mail a second mailing with a cover letter (see "GenoVA Cover Letter, Informed Consent and HIPAA Authorization forms, and a pre-stamped self-addressed return envelope. After at least 10 days, study staff follow the second mailing with a telephone call to the potential participant. During this call, staff review the informed consent information in detail, answer any questions, and obtain informed consent to participate. Study staff will ask the participant to sign and date the Informed Consent and HIPAA Authorization forms and return by mail in an included stamped envelope. No protected health information (PHI) or baseline surveys will be administered to patients prior to study staff receiving a properly completed, signed, and dated copy of participants' Informed Consent and HIPAA Authorization forms. After receiving properly completed, signed and dated Informed Consent and HIPAA Authorization forms from participants, study staff will then call participants and administer the baseline survey. Social media are used in recruitment. This study does not recruit participants with impaired decision-making capacity.

7.5 Participant incentives

Participants receive cash or a gift card for $30 after completion of the baseline survey and biospecimen collection and cash or a gift card for $30 after completion of the end-of-study survey.

8. Data Collection / Study Measures

8.1. Baseline telephone survey

The baseline survey (see "GenoVA Baseline Survey") is administered by study staff over the phone and takes approximately 15 minutes. It collects the following data:

- Family health history
- Smoking status
- Alcohol consumption
- Physical activity
- Processed meat consumption
- Aspirin use
- Medication adherence
- Patient activation

8.2 Specimen collection

By phone, study staff work with each participant to arrange for DNA specimen collection either by blood draw or by saliva sample. Participants may present to a VA Boston laboratory and undergo a blood draw of one EDTA tube (5 mL). If a participant already has an available EDTA blood sample in the laboratory (typically stored for 7 days after phlebotomy), that extant sample may be used instead. The VA Boston laboratory ships blood samples to an external VA-approved CLIA-certified laboratory for genotyping. Alternatively, the participant may choose to receive a coded saliva collection kit by mail, which he/she

can ship to the laboratory using a pre-paid shipping package. Participants who opt to receive a coded saliva collection kit by mail will receive a package including the saliva collection kit, a user guide (see "Oragene Saliva Kit User Guide"), instructions for sending their saliva specimen to the laboratory using a pre-paid shipping package (see "GenoVA Saliva Kit Packaging Instructions"), and a pre-paid shipping label.

8.3 Genotyping and reporting

The external laboratory performs genotyping on the DNA sample using a single-nucleotide polymorphism (SNP) array. The laboratory uses these resulting genotype data to calculate a polygenic risk score for 5 diseases for each participant, as described above, using the methods described by Khera, *Nat Comm.* 2018. The laboratory returns a clinical report with these scores for each participant to the study staff. The laboratory also returns the full uninterpreted genotype array data to the study staff, although these "raw" data are not returned to participants or providers or entered into the medical record.

8.4 End-of-study survey

The end-of-study survey is administered electronically via Qualtrics or Ipsos, or by study staff via telephone for patient-participants who do not have an active e-mail address, 24 months after enrollment and takes approximately 20 minutes. It collects the following data:

- Smoking status
- Alcohol consumption
- Physical activity
- Processed meat consumption
- Aspirin use
- Medication adherence
- Patient activation

The survey also asks patient-participants whether they had a new diagnosis of any of the six diseases during the study period; if needed, study staff may request that participants obtain and submit additional outside medical records for any affirmative response for review by study staff.

8.5 End-of-study chart review

Clinical chart reviewers blinded to patient-participant randomization status independently review each patient's medical record for the 24 months after enrollment for any evidence that one of the target diseases has been diagnosed since enrollment. Reviewers also abstract data on the medical care received during the study period, including diagnostic testing relevant to the target conditions.

8.6 Database query

Databases including the Corporate Data Warehouse (CDW), Centers for Medicaid and Medicare (CMS), and HERC Managerial Cost Accounting (MCA) databases are also accessed for the following data:

- Demographics
- Pharmacy data: Prescriptions and fill history
- ICD and CPT codes

- Vital signs: Blood pressure and body-mass index measurements
- Laboratory test results: LDL cholesterol, blood chemistries including glucose, hemoglobin A1c values, prostate-specific antigen (PSA) values, carcinoembryonic antigen (CEA) values
- Healthcare costs

8.7 Data repository

Study data are stored in two data repositories, as described below. Both data repositories comply with VHA Handbook 1200.12 and local VABHS IRB SOP.

8.7.1 NCBI Database of Genotypes and Phenotypes (dbGaP)

Deidentified study data from the patient-enrollees in this study are submitted to the NCBI dbGaP. This includes the SNP array data and the presence/absence of diagnosis of the 6 target diseases in the study. External researchers may request access to these data through a Data Use Certification (DUC) Agreement on the dbGaP website.

8.7.2 GenoVA Data Repository

A separate data repository is stored behind the VA firewall that includes de-identified individual-level trial data, including SNP array data, demographics, diagnoses, and survey data. Researchers outside VHA with an IRB-approved protocol may request access to these data.

9. Statistical Analysis Plan

9.1. Primary outcome: Time-to-new diagnosis of common complex disease

9.1.1 Statistical analysis plan

Intention-to-treat analyses compare the PRS-high and UC-high arms. The primary endpoint for efficacy is the time to a new diagnosis of one of 6 common complex diseases among patients with at least one PRS indicating high genetic risk. The analysis is based on the rate of new diagnosis at month 24 after randomization for the PRS-high and UC-high arms. We use the Cox model to analyze the data with the time to diagnosis. Further analyses examine time-to-new diagnosis for specific diseases separately. Moreover, analyses where summary PRS scores are included in statistical models, as well as demographic factors such as age, gender, socioeconomic status, and baseline health as covariates, are conducted. Additional analyses make outcomes comparisons between the randomized PRS-average and UC-average arms. Differences between the UC-high group and the UC-average group quantify the disease risk elevation among patients with $OR_{PRS}>2.0$ compared to those with all $OR_{PRS}\leq 2.0$. Differences between the UC-high and PRS-high arms will quantify the impact of telling high-risk patients and their PCPs about their high risk. Data from the 1-2% of participants excluded from the RCT because of a medically actionable finding will be analyzed separately in exploratory analyses.

	2014	2015	2016	2017	Average
Total eligible patients	6,145	6,083	6,016	5,770	6,003
Total patients diagnosed	372 (6.1%)	340 (5.6%)	388 (6.4%)	400 (6.9%)	375 (6.2%)
By disease					
CAD	113 (1.8%)	106 (1.7%)	126 (2.1%)	130 (2.3%)	119 (2.0%)
Atrial fibrillation	51 (0.8%)	60 (1.0%)	71 (1.2%)	62 (1.1%)	61 (1.0%)
T2D	156 (2.5%)	113 (1.9%)	127 (2.1%)	156 (2.7%)	138 (2.3%)
Colorectal cancer	8 (0.1%)	10 (0.2%)	13 (0.2%)	11 (0.2%)	11 (0.2%)
Breast cancer	1 (0.0%)	1 (0.0%)	3 (0.0%)	5 (0.0%)	2 (0.0%)
Prostate cancer	43 (0.7%)	50 (0.8%)	48 (0.8%)	36 (0.6%)	44 (0.7%)

Table 2 Annual rates of new diagnoses for six target diseases in patients between the ages of 50-70 years old across VA Boston Healthcare System, 2014-2017, by disease

Diagnosis estimates based on age of first diagnosis of at least one of the six target diseases. Annual estimates based on primary care relationship over two-year period with no new diagnosis of a target disease during first year.

9.1.2 Power and sample size calculation

A total of 1,076 patient-participants are genotyped. Based on published estimates of the prevalence of high-risk PRS values, we make the conservative assumption that 33% of genotyped patient-participants have at least one OR_{PRS}>2.0 across all diseases.

In review of data VABHS from 2014-2017 (Table 2), based on published algorithms using ICD and CPT codes, among patients 50-70 years old, an average 6.2% had a new target diagnosis per year. If we assume that 12% of patients in the control arm will have a new target diagnosis over the 2-year study period and that delivery of PRS results and recommendations will increase this to 24% of patients in the intervention arm during the same period (RR=2), a total sample size of 320 patient-participants must be included in the RCT to detect this difference at a two-tailed α=0.05 and β=0.20 (power of 80%). It is likely that patients with higher genetic risk (OR_{PRS}>2.0) have a higher disease incidence than that observed in the general VABHS population. If the incidence in the OR_{PRS}>2.0 group is 15%, then 320 patient-participants give us the same power to detect an increase to 28% in the PRS arm (RR 1.87). If 33% of genotyped patient-participants have at least one OR_{PRS}>2.0 and are enrolled in the RCT, a total sample size of 960 genotyped patient-participants is needed. To account for potential clustering effect among patients receiving care from the same providers, we include a design effect of 1.10, based on an estimate of 7 enrollees per PCP and an intraclass correlation coefficient of 0.02 (Glynn, Medical Care, 2007). As a result, a total of 960 x 1.1 = 1,056 participants would be needed. We will enroll a total of 1,076 to account for the 1-2% of participants who will be excluded from randomization due to an actionable ACMG variant.

9.2 Secondary and other outcomes

Other intention-to-treat analyses compare the randomized PRS-high and UC-high randomization arms to examine secondary and other outcomes. Similar analyses compare the randomized PRS-average and UC-average arms. The Wilcoxon Rank-Sum test is used to compare ordinal measures (e.g. processed meat consumption) between groups. For continuous outcomes (e.g. change in systolic blood pressure and healthcare costs), we use linear regression to compare the randomization arms. Logistic regression is used for dichotomous outcomes (e.g. occurrence of diagnostic testing). Regression models include a term for PRS versus UC randomization status. Covariates are included if they improve model precision. Missing data are imputed using fully conditional specification. Additional analyses may include examination of associated research questions of interest to the research team.

10. Ethical Issues

10.1 Potential Risks

Patient-participants are subject to the following risks:

- The patient-participants' providers may order unnecessary screening tests in response to PRS results or the medically actionable findings identified in 1-2% of participants. However, in routine clinical care, there is already much variation in provider behavior around disease screening for the 6 target diseases, and any screening test ordered in response to high-risk PRS results likely falls into the range of what would be considered reasonable medical management with a favorable benefit/risk ratio (e.g. hemoglobin A1c testing or colonoscopy). It is unlikely that providers initiate new medication therapy in response to PRS without first confirming a new disease diagnosis. For the medically actionable findings, the study genetic counselor will provide information and consultation to the patient-participant and his/her healthcare providers for recommended clinical management. Thus, this study poses risks not dissimilar to those of current standard of care for the screening and management of these diseases.
- Patient-participants who chose to submit a blood sample for genotyping may experience bruising, lightheadedness, or infection from phlebotomy.
- They may experience psychological distress upon learning they have a high genetic susceptibility for a certain disease, including an unanticipated medically actionable finding.
- If a medically actionable finding is identified in a participant (estimated 1-2% of participants), then his/her first-degree family members each have a 50% chance of also carrying the finding. There is the risk that these family members will not want to learn this information or will experience distress upon learning it. The genetic counselor will discuss these risks with the participant before he/she consents to learn about the medically actionable result.
- Although federal law prevents health insurance companies from discriminating against patients on the basis of genetic information, some insurance companies may deny life, disability, and long-term care coverage on the basis of genetic information, such as unanticipated medically actionable findings or the PRS used in this study.
- There is the risk of breach of data privacy.
- For active-duty military participants, study-related information that is included in the VA medical record is subject to fewer protections, including access by DOD personnel.

10.2 Protections Against Risks

The risks to participating patients are minimal and not dissimilar from routine clinical care, where there is already much variation in provider behavior around screening and risk management for these 6 common diseases and variation in patient behavior around adherence to management recommendations. The misinterpretation and misuse of PRS results is minimized through the reporting of clear, concise test interpretations, coupled to evidence-based screening and risk management recommendations consistent with accepted medical practice. Risk of mismanagement is further minimized because patients are in the care of their primary care providers, using clinical judgment for patient management.

Risk of breach of confidentiality is minimized through the appropriate management and security of clinical data per VABHS and HIPAA protocols for use of research data. Data are securely transmitted using VA approved methods, including FIPS 140-2 validated encryption. This includes transmission of PHI and other patient-participant data, including PRS results, between VABHS and the external clinical laboratory, where clinical genotyping and interpretation are performed. Patient data files (source and analytic) are stored behind the VA firewall, on a drive created specifically to house the data for this research project.

A copy of patient mailing data only will be downloaded outside of the drive specifically created to house data files in a VA secured, study-specific SharePoint site, and behind the VA firewall where strict permissions will be set to limit viewing to IRB-approved study personnel. This will be done to allow for the use of the Microsoft mail merge software so patient letters and address labels can be created and printed in batches, increasing patient enrollment numbers to meet the study's grant time table. Patient mailing data will be in the form of CSV files and may include identifying variables for both patients and providers. Variables for patients/providers may include: ID, full name, title, institution code/ID, gender, mailing address, and any associated flags (i.e. temporary address), patient-provider relationship information, or other similar variables that are required to be able to send mail or that are named in the IRB-approved patient letter template. The use of the mail merge system can be completed within the secure SharePoint environment.

Patient protected health information (PHI) are delinked from the final analytic dataset. All data are retained within the VA except in 2 instances. First, coded biospecimens with DNA are sent to an external VA-approved laboratory for genotyping. Although these specimens will have DNA, they will not be labeled with any patient identifiers. Biospecimens will be shipped by commercial shippers using chain of custody, minimizing the risk of data breach. Second, deidentified data (including genetic risk scores but not the full genetic array data) will be submitted to the dbGaP data repository, per NIH regulations.

Only study personnel credentialed and approved by the IRB have access to study data stored in either physical or electronic environments. Once study team members are no longer a part of the research team, their access to data and research materials is terminated. We do not allow any unauthorized access to our servers or our datasets. No PHI is released to the public, nor is it published in any medical journal. Suspected information security and privacy incidents are reported within one hour to the Information Security and Privacy Officers. Data are kept indefinitely or until the law allows their destruction in accordance with the VA Record Control Schedule. Electronic records are destroyed, when allowed, in a manner in which they cannot be retrieved. Mobile devices will not be used in this study and thus will not contain the only copy of research information.

Active-duty military participants will be made aware of the potential for access of study-related data entered into the VA medical record by additional parties (DOD personnel) through the informed consent process. Prior to study participation, these participants will provide written consent to take part in study activities.

10.3 Potential Benefits

The benefits to patients participating in this study include the potential for them to engage with their healthcare providers in therapeutic conversations about the risks and benefits of screening for and reducing the risk of 6 common diseases. Receipt of high-risk PRS or incidental actionable results might

prompt providers to order appropriate screening tests they might have otherwise overlooked. Patients might be more adherent to recommended screening and risk-reducing behaviors if they feel the recommendations are personalized to them. Society also benefits from the knowledge to be learned about the impact of introducing PRS testing into clinical care. These potential benefits outweigh the minimal potential risks to providers and patient.

10.4 Stopping Rules

A participant may always withdraw their participation at any time. The study has no stopping rules.

11. Safety Monitoring Plan

Overall, the risks to participating patients are considered minimal and are not dissimilar from the risks inherent to routine clinical care. Moreover, the study occurs within a healthcare system and thus concomitantly includes the oversight of providers in the clinical management of patients enrolled in the study. Study data are collected observationally through the patient electronic health record and through participant surveys and interviews. As such, the principal investigator (PI), Dr. Jason Vassy, a physician by training, and the research staff is responsible for the day-to-day monitoring of patient safety and data protection throughout the conduct of this proposed research. Monitoring of patient safety and data protection occurs in conjunction with the regular operations and conduct of the study and is commensurate with the relative risks associated with the proposed research.

If a participant experiences adverse effects or expresses emotional distress related to study participation, including the receipt of high-risk results/score (PRS) or actionable finding, and requires medical attention based on the judgement of the PI (a physician) and/or the patient's participating provider, they are referred for clinical assessment as appropriate. All serious such cases, including those requiring a referral to a mental health professional or other therapeutic intervention, are reported to the IRB as required. Additionally, though not provided as part of this study, participants' primary care providers may choose to refer their patients to genetic counseling services.

Any concerns regarding the ethical conduct of the study, the safety of participants, or a breach in the protection of study data made by provider or patient participants, the study staff, or others is promptly reported to the study PI and escalated accordingly to the IRB and other relevant research oversight committees.

12. Adverse Event/Unanticipated Problems Reporting Plans

We anticipate very few, if any, adverse effects (AE) during the course of the study, but we nonetheless have a process in place to identify and address AE if they occur. An AE is defined as any unanticipated or unintended medical occurrence, which does not necessarily have a causal relationship with the study condition, procedure(s) or study agent(s), that occurs after the informed consent is obtained. Pre-existing conditions or illnesses which are expected to exacerbate or worsen are not considered adverse events and are accounted for in the subject's medical history. A serious adverse event (SAE) is defined as an AE resulting in one of the following outcomes: death during the 24 months after enrollment, life threatening event (defined as an event that places a participant at immediate risk of death), inpatient hospitalization, and any other condition which, in the judgment of the investigator, represents a significant hazard, such as an important medical event that does not result in one of the above

outcomes. An event may be considered an SAE when it jeopardizes the participant or requires medical or surgical intervention to prevent one of the outcomes listed above. AEs may be observed by the study staff or volunteered by VABHS providers and patients. All AEs or SAEs are assessed for relationship to the study research procedures, to determine whether study participation was likely to have caused the AE/SAE. AEs related to study participation that are reported to research personnel are recorded on an AE form in an electronic database.

The Principal Investigator at VABHS reports unanticipated problems, deaths, study-related AEs, and safety monitors' reports to the IRB in accordance with VHA Handbook 1058.01 and VABHS IRB SOP. These events will also be reported to the Harvard Medical School IRB in accordance with federal and local policies, without using participant identifiers.

Note S2. Genomic Medicine at VA (GenoVA) Study Baseline Survey

You are currently working on the record of GenoVA ID [XXXXX]

The following questions are intended to collect information about you and your health care. You may choose to skip any question that you do not wish to answer.

A. SELF-RELATED HEALTH AND QUALITY OF LIFE

This information will keep track of how you feel and how well you are able to do your usual activities. If you are unsure how to answer a question, please give the best answer you can.

1. In general, how would you describe your health?

* must provide value

- ○ Excellent
- ○ Very Good
- ○ Good
- ○ Fair
- ○ Poor
- ○ Did not answer or declined

2. The following questions are about activities you might do during a typical day. Does your health now limit you in these activities? if so, how much?

a. Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf?

* must provide value

- ○ Yes, limited a lot
- ○ Yes, limited a little
- ○ No, not limited at all
- ○ Did not answer or declined

b. Climbing several flights of stairs?

* must provide value

- ○ Yes, limited a lot
- ○ Yes, limited a little
- ○ No, not limited at all
- ○ Did not answer or declined

3. During the past 4 weeks, have you experienced any of the following problems with your work or other regular daily activities as a result of your physical health?

a. Accomplished less than you would like?

* must provide value

- ○ No, none of the time
- ○ Yes, a little of the time
- ○ Yes, some of the time
- ○ Yes, most of the time
- ○ Yes, all of the time
- ○ Did not answer or declined

b. Were limited in the kind of work or other activities?

* must provide value

- ○ No, none of the time
- ○ Yes, a little of the time
- ○ Yes, some of the time
- ○ Yes, most of the time
- ○ Yes, all of the time
- ○ Did not answer or declined

4. During the past 4 weeks, have you experienced any of the following problems with your work or other regular daily activities as a result of any emotional problems (such as feeling depressed or anxious)?

a. Accomplished less than you would like?

* must provide value

- ○ No, none of the time
- ○ Yes, a little of the time
- ○ Yes, some of the time
- ○ Yes, most of the time
- ○ Yes, all of the time
- ○ Did not answer or declined

b. Didn't do work or other activities as carefully as usual?

* must provide value

- ○ No, none of the time
- ○ Yes, a little of the time
- ○ Yes, some of the time
- ○ Yes, most of the time
- ○ Yes, all of the time
- ○ Did not answer or declined

5. During the past 4 weeks, how much did pain interfere with your normal work (including both work outside the home and house work)?

* must provide value

- ○ Not at all
- ○ A little bit
- ○ Moderately
- ○ Quite a bit
- ○ Extremely
- ○ Did not answer or declined

The next questions are about how you feel and how things have been with you during the past 4 weeks. For each question, please give the one answer that comes closest to the way you have been feeling.

6. How much of the time during the past 4 weeks:

a. Have you felt calm and peaceful?

* must provide value

- ○ All of the time
- ○ Most of the time
- ○ A good bit of the time
- ○ Some of the time
- ○ A little of the time
- ○ None of the time
- ○ Did not answer or declined

b. Did you have a lot of energy?

* must provide value

- ○ All of the time
- ○ Most of the time
- ○ A good bit of the time
- ○ Some of the time
- ○ A little of the time
- ○ None of the time
- ○ Did not answer or declined

c. Have you felt downhearted and blue?

* must provide value

- ○ All of the time
- ○ Most of the time
- ○ A good bit of the time
- ○ Some of the time
- ○ A little of the time
- ○ None of the time
- ○ Did not answer or declined

7. During the past 4 weeks, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting with friends, relatives, etc.)?

* must provide value

- ○ All of the time
- ○ Most of the time
- ○ Some of the time
- ○ A little of the time
- ○ None of the time
- ○ Did not answer or declined

Now, we'd like to ask you some questions about how your health may have changed.

8. Compared to one year ago, how would you rate your physical health in general now?

* must provide value

- ○ Much better
- ○ Slightly better
- ○ About the same
- ○ Slightly worse
- ○ Much worse
- ○ Did not answer or declined

9. Compared to one year ago, how would you rate your emotional problems (such as feeling anxious, depressed, or irritable) now?

* must provide value

- ○ Much better
- ○ Slightly better
- ○ About the same
- ○ Slightly worse
- ○ Much worse
- ○ Did not answer or declined

B. PATIENT ACTIVATION

The following are statements that people sometimes make when they talk about their health. Please indicate how much you disagree or agree with each statement as it applies to you personally. Your answers should be what is true for you and not just what you think others expect of you. Your choices are Strongly Disagree, Disagree, Agree, or Strongly Agree. If the statement does not apply to you, please say Does Not Apply.

1. When all is said and done, I am the person who is responsible for managing my health.

* must provide value

- ○ Strongly Disagree
- ○ Disagree
- ○ Agree
- ○ Strongly Agree
- ○ Does Not Apply
- ○ Did not answer or declined

2. Taking an active role in my own health care is the most important factor in determining my health and ability to function.

* must provide value

- ○ Strongly Disagree
- ○ Disagree
- ○ Agree
- ○ Strongly Agree
- ○ Does Not Apply
- ○ Did not answer or declined

3. I am confident that I can take actions that will help prevent or minimize some symptoms and problems associated with my health.

* must provide value

- ○ Strongly Disagree
- ○ Disagree
- ○ Agree
- ○ Strongly Agree
- ○ Does Not Apply
- ○ Did not answer or declined

4. I know what each of my prescribed medications does.

* must provide value

- ○ Strongly Disagree
- ○ Disagree
- ○ Agree
- ○ Strongly Agree
- ○ Does Not Apply
- ○ Did not answer or declined

5. I am confident that I can tell when I need to go get medical care and when I can handle a health problem myself.

* must provide value

- ○ Strongly Disagree
- ○ Disagree
- ○ Agree
- ○ Strongly Agree
- ○ Does Not Apply
- ○ Did not answer or declined

6. I am confident I can tell a doctor concerns I have even when he or she does not ask.

* must provide value

- ○ Strongly Disagree
- ○ Disagree
- ○ Agree
- ○ Strongly Agree
- ○ Does Not Apply
- ○ Did not answer or declined

7. I am confident that I can follow through on medical treatments I need to do at home.

* must provide value

- ○ Strongly Disagree
- ○ Disagree
- ○ Agree
- ○ Strongly Agree
- ○ Does Not Apply
- ○ Did not answer or declined

8. I understand the nature and causes of my health problems.

* must provide value

- ○ Strongly Disagree
- ○ Disagree
- ○ Agree
- ○ Strongly Agree
- ○ Does Not Apply
- ○ Did not answer or declined

9. I know the different medical treatment options available for my health condition.

* must provide value

- ○ Strongly Disagree
- ○ Disagree
- ○ Agree
- ○ Strongly Agree
- ○ Does Not Apply
- ○ Did not answer or declined

10. I have been able to maintain the lifestyle changes for my health that I have made.

* must provide value

- ○ Strongly Disagree
- ○ Disagree
- ○ Agree
- ○ Strongly Agree
- ○ Does Not Apply
- ○ Did not answer or declined

11. I know how to prevent further problems with my health,

* must provide value

- ○ Strongly Disagree
- ○ Disagree
- ○ Agree
- ○ Strongly Agree
- ○ Does Not Apply
- ○ Did not answer or declined

12. I am confident I can figure out solutions when new situations or problems arise with my health.

* must provide value

- ○ Strongly Disagree
- ○ Disagree
- ○ Agree
- ○ Strongly Agree
- ○ Does Not Apply
- ○ Did not answer or declined

13. I am confident I can maintain lifestyle changes, like diet and exercise, even during times of stress.

* must provide value

- ○ Strongly Disagree
- ○ Disagree
- ○ Agree
- ○ Strongly Agree
- ○ Does Not Apply
- ○ Did not answer or declined

C. MEDICATIONS

These next questions are about any medications you might take.

1. Over the past 7 days, how often did you take your prescribed medications as prescribed?

* must provide value

- ○ Very often
- ○ Somewhat often
- ○ Sometimes
- ○ Rarely
- ○ Never
- ○ Did not answer or declined

2. Over the past 7 days, how often did you skip a dose of prescribed medication?

* must provide value

- ○ Very often
- ○ Somewhat often
- ○ Sometimes
- ○ Rarely
- ○ Never
- ○ Did not answer or declined

3. Over the past 7 days, how often were you unable to take your prescribed medications at all?

* must provide value

- ○ Very often
- ○ Somewhat often
- ○ Sometimes
- ○ Rarely
- ○ Never
- ○ Did not answer or declined

4. Do you regularly take a daily aspirin, either by prescription or over-the-counter?

* must provide value

- ○ Yes
- ○ No
- ○ Unsure/Don't Know
- ○ Did not answer or declined

D. HEALTH BEHAVIORS

These next questions are about your health habits.

1. How often would you say that you consume alcohol?

* must provide value

- ○ Very often
- ○ Somewhat often
- ○ Sometimes
- ○ Rarely
- ○ Never
- ○ Did not answer or declined

2. How often would you say that you exercise?

* must provide value

- ○ Very often
- ○ Somewhat often
- ○ Sometimes
- ○ Rarely
- ○ Never
- ○ Did not answer or declined

3. How often would you say that you consume processed meat?

* must provide value

- ○ Very often
- ○ Somewhat often
- ○ Sometimes
- ○ Rarely
- ○ Never
- ○ Did not answer or declined

4. Have you smoked at least 100 cigarettes in your entire lifetime?

* must provide value

- ○ Yes
- ○ No
- ○ Don't know/Not sure
- ○ No Answer
- ○ Did not answer or declined

5. Do you now smoke cigarettes every day, some days, or not at all?

* must provide value

- ○ Every day
- ○ Some days
- ○ Not at all
- ○ Don't know/Not sure
- ○ No answer

6. During the past 12 months, have you stopped smoking for one day or longer because you were trying to quit smoking?

* must provide value

- ○ Yes
- ○ No
- ○ Don't know/Not sure
- ○ No Answer

7. How long has it been since you last smoked a cigarette, even one or two puffs?

* must provide value

- ○ Within the past month (less than 1 month ago)
- ○ Within the past 3 months (more than 1 month ago but less than 3 months ago)
- ○ Within the past 6 months (more than 3 months ago but less than 6 months ago)
- ○ Within the past year (more than 6 months but less than 1 year ago)
- ○ Within the past 5 years (more than 1 year ago but less than 5 years ago)
- ○ Within the past 10 years (more than 5 years but less than 10 years ago)
- ○ 10 years or more ago
- ○ Never smoked regularly
- ○ Don't know/Not sure
- ○ Did not answer or declined

8. Do you currently use chewing tobacco, snuff, or snus every day, some days, or not at all?

* must provide value

- ○ Every day
- ○ Some days
- ○ Not at all
- ○ Don't know/Not sure
- ○ No Answer
- ○ Did not answer or declined

E. DEMOGRAPHIC INFORMATION

The final questions ask you about your race and ethnicity.

1. What is your race? (Please choose all that apply)

- ☐ White
- ☐ Black / African-American
- ☐ American Indian / Alaska Native
- ☐ Chinese
- ☐ Japanese
- ☐ Asian Indian
- ☐ Other Asian
- ☐ Filipino
- ☐ Pacific Islander
- ☐ Other

2. Are you Spanish, Hispanic, or Latino?

- ☐ No, not Spanish, Hispanic, or Latino
- ☐ Yes, Mexican, Mexican American, Chicano
- ☐ Yes, Puerto Rican
- ☐ Yes, Cuban
- ☐ Yes, other Spanish, Hispanic, Latino

That completes the survey. Thank you for your participation.

Enter the date when the participant answered the survey. If multiple encounters were needed, enter the day of the most recent one.

M-D-Y

* must provide value

Comments on data collection process:

U.S. Department of Veterans Affairs
Veterans Health Administration
Boston Healthcare System

Introduction

GenoVA Study Participant End-of-Study Survey

Thank you for participating in the Genomic Medicine at VA (GenoVA) Study. As a reminder, you enrolled in this study approximately 2 years ago. This study is looking at whether learning about high genetic risk for a disease might help patients and their healthcare providers prevent or detect these diseases even earlier than they might otherwise.

As a study participant, you may have received genetic results approximately 2 years ago related to your risk of 5 common diseases: coronary artery disease, atrial fibrillation, type 2 diabetes, colorectal cancer, prostate cancer (for participants with a prostate), and breast cancer (for participants born with female sex). You may have been told you have high genetic risk of one or more of these diseases, or just average genetic risk for all of them.

We are now inviting you to take the end-of-study survey. It should take approximately 15 - 20 minutes to complete. The following questions are intended to collect information about you and your health care. You may choose to skip any question you do not wish to answer.

Recall

Section A: RECALL

1. First, do you recall whether you received genetic risk results, either normal (average risk) or abnormal (high risk), as a part of this study?

	Yes, I received results	No, I did not receive results	Unsure
Please Select:	☐	☐	☐

Recall 1

2. Did your genetic risk results indicate that you had high genetic risk for any of the following diseases (check all that apply)?

	Coronary artery disease	Atrial fibrillation	Type 2 diabetes	Colorectal cancer	Prostate cancer	Breast cancer	Unsure
Check all that apply	☐	☐	☐	☐	☐	☐	☐

Health Conditions

Section B: HEALTH CONDITIONS

The next questions are about any new health conditions you've been **diagnosed** with in the past 2 years, since enrolling in this study.

Have you been told by a healthcare provider that you have any of the following conditions:

1. Coronary artery disease, such as a heart attack, coronary bypass surgery, or stents in the blood vessels in your heart?

	Yes	No	Unsure
Please Select:	☐	☐	☐

Health Conditions: Coronary Artery Disease

A. Approximately when were you diagnosed with coronary artery disease?

	Month	Year	Unsure
Please Select:	▾	▾	☐

B. Was this diagnosed at a VA facility or at an outside facility?

- ☐ VA facility
- ☐ Outside (non-VA) facility
- ☐ Unsure

Please type the name of the medical facility where you were diagnosed with coronary artery disease.

Health Conditions: Diabetes

2. Have you been told by a healthcare provider that you have diabetes?

	Yes	No	Unsure
Please Select:	☐	☐	☐

Health Conditions: Diabetes

A. Approximately when were you diagnosed with diabetes?

	Month	Year	Unsure
Please Select:			☐

B. Was this diagnosed at a VA facility or at an outside facility?

- ☐ VA facility
- ☐ Outside (non-VA) facility
- ☐ Unsure

Please type the name of the medical facility where you were diagnosed with diabetes.

Health Conditions: Atrial Fibrillation

3. Have you been told by a healthcare provider that you have atrial fibrillation?

	Yes	No	Unsure
Please Select:	☐	☐	☐

Health Conditions: Atrial Fibrillation

A. Approximately when were you diagnosed with atrial fibrillation?

	Month	Year	Unsure -
Please Select:			☐

B. Was this diagnosed at a VA facility or at an outside facility?

- ☐ VA facility
- ☐ Outside (non-VA) facility
- ☐ Unsure

Please type the name of the medical facility where you were diagnosed with atrial fibrillation.

Health Conditions: Colon or Rectal Cancer

4. Have you been told by a healthcare provider that you have colon or rectal cancer?

	Yes	No	Unsure
Please Select:	☐	☐	☐

Health Conditions: Colon or Rectal Cancer

A. Approximately when were you diagnosed with colon or rectal cancer?

	Month	Year	Unsure -
Please Select:			☐

B. Was this diagnosed at a VA facility or at an outside facility?

- ☐ VA facility
- ☐ Outside (non-VA) facility
- ☐ Unsure

Please type the name of the medical facility where you were diagnosed with colon cancer or rectal cancer.

Health Conditions: Prostate Cancer

5. Have you been told by a healthcare provider that you have prostate cancer (if male)?

	Yes	No	Unsure
Please Select:	☐	☐	☐

Health Conditions: Prostate Cancer

A. Approximately when were you diagnosed with prostate cancer?

	Month	Year	Unsure -
Please Select:			☐

B. Was this diagnosed at a VA facility or at an outside facility?

- ☐ VA facility
- ☐ Outside (non-VA) facility
- ☐ Unsure

Please type the name of the medical facility where you were diagnosed with prostate cancer.

Health Conditions: Breast Cancer

6. Have you been told by a healthcare provider that you have breast cancer (if female)?

	Yes	No	Unsure
Please Select:	☐	☐	☐

Health Conditions: Breast Cancer

A. Approximately when were you diagnosed with breast cancer?

	Month	Year	Unsure
Please Select:			☐

B. Was this diagnosed at a VA facility or at an outside facility?

- ☐ VA facility
- ☐ Outside (non-VA) facility
- ☐ Unsure

Please type the name of the medical facility where you were diagnosed with breast cancer.

Screening and Diagnostic Tests and Procedures: Coronary Artery Disease

Section C: SCREENING AND DIAGNOSTIC TESTS AND PROCEDURES

These next questions are about your **screening** and **diagnostic** testing history.

1. Have you had a test for coronary artery disease in the past 2 years, such as a stress test, cardiac CT for coronary artery calcium (CAC), or a coronary angiography?

- ☐ Yes
- ☐ No
- ☐ Don't know/unsure

Screening and Diagnostic Tests and Procedures: Coronary Artery Disease

A. What test(s) have you had for *coronary artery disease* in the past 2 years? Please select all that apply.

- [ ] Stress test [moderate physical exercise using a treadmill or stationary bike where a healthcare provider monitors heart rhythm, blood pressure, and breathing]
- [ ] Cardiac CT for coronary artery calcium (CAC) [computerized tomography detects calcium deposits that can decrease blood flow in the heart's arteries]
- [ ] Coronary angiography [x-ray visible dye is injected into your blood vessels to discern restriction in blood flow]
- [ ] Other

B. Approximately when did this test occur? (If you've had more than one of these tests, approximately when did the first one occur?)

	Month	Year	Unsure
Please Select:			-

C. Did this test take place at a VA facility or at an outside (non-VA) facility?

- [ ] VA facility
- [ ] Outside (non-VA) facility
- [ ] Unsure

D. Please type the name of the medical facility where the appointment was.

Screening and Diagnostic Tests and Procedures: Diabetes

2. Have you had a screening test for diabetes in the past 2 years, such as a hemoglobin A1c test or a fasting glucose test?

- [ ] Yes
- [ ] No
- [ ] Don't know/unsure

Screening and Diagnostic Tests and Procedures: Diabetes

A. What type(s) of screening test have you had for *diabetes* in the past 2 years? Please select all that apply.

- [ ] Hemoglobin A1c [blood test to calculate average blood sugar levels within the past few months]
- [ ] Fasting glucose test [requires an individual to abstain from eating or drinking anything eight hours prior to a blood draw]
- [ ] Other

B. Approximately when did this test occur? (If you've had more than one of these tests, approximately when did the first one occur?)

	Month	Year	Unsure
Please Select:			-

C. Did this test take place at a VA facility or at an outside (non-VA) facility?

- [ ] VA facility
- [ ] Outside (non-VA) facility
- [ ] Unsure

D. Please type the name of the medical facility where the appointment was.

Screening and Diagnostic Tests and Procedures: Atrial Fibrillation

3. Have you had a test for atrial fibrillation in the past 2 years, such as an electrocardiogram (ECG or EKG) or heart rhythm monitoring?

- [ ] Yes
- [ ] No
- [ ] Don't know/unsure

Screening and Diagnostic Tests and Procedures: Atrial Fibrillation

A. What type(s) of tests have you undergone for *atrial fibrillation* in the past 2 years? Please select all that apply.

- [ ] Electrocardiogram (ECG or EKG) [performed at a medical provider's office or in the hospital]
- [ ] ECG patch monitor, Holter monitor, or cardiac event monitor [ordered by your health care provider which records your heart rhythm (sometimes for 24 hours – 30 days)]
- [ ] Other

B. Approximately when did this test occur? (If you've had more than one of these tests, approximately when did the first one occur?)

	Month	Year	Unsure
Please Select:			-

C. Did this test take place at a VA facility or at an outside (non-VA) facility?

- [ ] VA facility
- [ ] Outside (non-VA) facility
- [ ] Unsure

D. Please type the name of the medical facility where the appointment was.

Screening and Diagnostic Tests and Procedures: Colon or Rectal Cancer

4. Have you had a test for colon or rectal cancer in the past 2 years, such as a colonoscopy, sigmoidoscopy, fecal blood testing, or CT colonography?

- [ ] Yes
- [ ] No
- [ ] Don't know/unsure

Screening and Diagnostic Tests and Procedures: Colon or Rectal Cancer

A. What type(s) of screening test have you undergone for *colon* or *rectal cancer* in the past 2 years? Please select all that apply.

- [ ] Colonoscopy [a small camera is inserted into the rectum using a flexible tube in order to view the large intestine. To prepare a provider may ask you to abstain from eating the day prior and recommend a laxative]
- [ ] Sigmoidoscopy [similar to a colonoscopy but less invasive in that the provider only views the lower colon]
- [ ] Fecal blood testing [lab test used to check stool samples for hidden blood]
- [ ] CT colonography [minimally invasive CT scan to screen for cancer of the large intestine]
- [ ] Other

B. Approximately when did this test occur? (If you've had more than one of these tests, approximately when did the first one occur?)

	Month	Year	Unsure -
Please Select:			

C. Did this test take place at a VA facility or at an outside (non-VA) facility?

- [ ] VA facility
- [ ] Outside (non-VA) facility
- [ ] Unsure

D. Please type the name of the medical facility where the appointment was.

Screening and Diagnostic Tests and Procedures: Prostate Cancer

5. Have you had a test for prostate cancer in the past 2 years, such as a PSA blood test, a prostate MRI or ultrasound, or a prostate biopsy?

- [ ] Yes
- [ ] No

- [ ] Don’t know/unsure

Screening and Diagnostic Tests and Procedures: Prostate Cancer

A. What type(s) of test have you had for *prostate cancer* in the past 2 years? Please select all that apply.

- [ ] Prostate-specific antigen (PSA) test [a blood test used to check men for prostate cancer]
- [ ] Prostate ultrasound [imaging test with a small probe that uses sound waves to look at your prostate or your rectum]
- [ ] Prostate MRI [uses a strong magnetic field instead of X-rays to provide clear and detailed pictures the prostate gland]
- [ ] Prostate biopsy [procedure to remove samples of suspicious tissue from the prostate]
- [ ] Other

B. Approximately when did this test occur? (If you’ve had more than one of these tests, approximately when did the first one occur?)

	Month	Year	Unsure -
Please Select:			☐

C. Did this test take place at a VA facility or at an outside (non-VA) facility?

- [ ] VA facility
- [ ] Outside (non-VA) facility
- [ ] Unsure

D. Please type the name of the medical facility where the appointment was.

Screening and Diagnostic Tests and Procedures: Breast Cancer

6. Have you had a test for breast cancer in the past 2 years, such as a mammogram, breast MRI or ultrasound, or breast biopsy?

- [ ] Yes
- [ ] No
- [ ] Don't know/unsure

Screening and Diagnostic Tests and Procedures: Breast Cancer

A. What type(s) of test have you had for ***breast cancer*** in the past 2 years? Please select all that apply.

- [ ] Mammography [An X-ray of the breast done to check for breast cancer]
- [ ] Breast magnetic resonance imaging (MRI) [uses a strong magnetic field instead of X-rays to provide clear and detailed pictures to detect breast cancer and other abnormalities in the breast]
- [ ] Breast ultrasound [imaging test that uses a wand like device on the skin to look at the inside of breasts]
- [ ] Breast biopsy [procedure to remove a sample of breast tissue to test for cancerous cells]
- [ ] Other

B. Approximately when did this test occur? (If you've had more than one of these tests, approximately when did the first one occur?)

	Month	Year	Unsure -
Please Select:			

C. Did this test take place at a VA facility or at an outside (non-VA) facility?

- [ ] VA facility
- [ ] Outside (non-VA) facility
- [ ] Unsure

D. Please type the name of the medical facility where the appointment was.

Health Care and Healthcare Utilization

Section D: HEALTH CARE AND HEALTHCARE UTILIZATION

The following statements are about other medical care you've received since enrolling in this study.

Please indicate if you've seen any of the following providers in the last 2 years. You may respond *"Yes," "No,"* or *"I'm not sure/I don't know."*

	If you've seen any of the following providers in the last 2 years			If yes, how many times you see this kind of prov in the last 2 years
	Yes	No	I'm not sure/I don't know	
a. VA primary care provider	○	○	○	⌄

-

	If you've seen any of the following providers in the last 2 years			If yes, how many times you see this kind of prov in the last 2 years
	Yes	No	I'm not sure/I don't know	
b. Outside (non-VA) primary care provider	○	○	○	⌄

-

	If you've seen any of the following providers in the last 2 years			If yes, how many times you see this kind of prov in the last 2 years
	Yes	No	I'm not sure/I don't know	
c. Cardiologist	○	○	○	⌄

-

	If you've seen any of the following providers in the last 2 years			If yes, how many times you see this kind of prov in the last 2 years
	Yes	No	I'm not sure/I don't know	
d. Geneticist or genetic counselor other than the research genetic counselor for this study	○	○	○	

In the last 2 years, have you been admitted to the hospital?

- [ ] No
- [ ] Yes

Health Care and Healthcare Utilization

1. How many times have you been admitted to the hospital in the last 2 years?

	Count
Please Select:	

Health Care and Healthcare Utilization

A. Admission 1

Please provide us with the reason(s) for this hospital admission:

How many days did you spend in the hospital?

Did you spend any time in the intensive care unit (ICU)?

- [ ] No
- [ ] Yes

If yes, how many days?

B. Admission 2

Please provide us with the reason(s) for this hospital admission:

How many days did you spend in the hospital?

Did you spend any time in the intensive care unit (ICU)?

- [ ] No
- [ ] Yes

If yes, how many days?

C. Admission 3

Please provide us with the reason(s) for this hospital admission:

How many days did you spend in the hospital?

Did you spend any time in the intensive care unit (ICU)?

- [ ] No
- [ ] Yes

If yes, how many days?

Health Care and Healthcare Utilization

A. Admission 1

Please provide us with the reason(s) for this hospital admission:

How many days did you spend in the hospital?

Did you spend any time in the intensive care unit (ICU)?

- ○ No
- ○ Yes

If yes, how many days?

B. Admission 2

Please provide us with the reason(s) for this hospital admission:

How many days did you spend in the hospital?

Did you spend any time in the intensive care unit (ICU)?

- ☐ No
- ☐ Yes

If yes, how many days?

Health Care and Healthcare Utilization

A. Admission 1

Please provide us with the reason(s) for this hospital admission:

How many days did you spend in the hospital?

	Day
Please Select:	

Did you spend any time in the intensive care unit (ICU)?

- [ ] No
- [ ] Yes

If yes, how many days?

Free Text

Please provide any additional information you wish about diagnosis, screening tests, or outpatient or hospital visits you have had in the last 2 years.

Self-Related Health and Quality of Life

Section E: SELF-RELATED HEALTH AND QUALITY OF LIFE

This information will help keep track of how you feel and how well you are able to do your usual activities. If you are unsure how to answer a question, please select the best answer from the choices provided.

1. In general, would you say your health is…?

	Excellent	Very Good	Good	Fair	Poor
Please Select:	☐	☐	☐	☐	☐

2. The following questions are about activities you might do during a typical day. Does your health now limit you in these activities? If so, how much?

a. Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf?

	Yes, limited a lot	Yes, limited a little	No, not limited at all
Please Select:	☐	☐	☐

b. Climbing several flights of stairs?

	Yes, limited a lot	Yes, limited a little	No, not limited at all
Please Select:	☐	☐	☐

3. During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of your physical health?

a. Accomplished *less* than you would like?

- ☐ No, none of the time
- ☐ Yes, a little of the time
- ☐ Yes, some of the time
- ☐ Yes, most of the time
- ☐ Yes, all of the time

b. Were limited in the *kind* of work or other activities?

- ☐ No, none of the time
- ☐ Yes, a little of the time
- ☐ Yes, some of the time
- ☐ Yes, most of the time
- ☐ Yes, all of the time

4. During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of any emotional problems (such as feeling depressed or anxious)

? a. Accomplished *less* than you would like?

- [ ] No, none of the time
- [ ] Yes, a little of the time
- [ ] Yes, some of the time
- [ ] Yes, most of the time
- [ ] Yes, all of the time

b. Didn't do work or other activities as carefully as usual?

- [ ] No, none of the time
- [ ] Yes, a little of the time
- [ ] Yes, some of the time
- [ ] Yes, most of the time
- [ ] Yes, all of the time

5. During the past 4 weeks, how much did pain interfere with your normal work (including both work outside the home and house work)?

- [ ] Not at all
- [ ] A little bit
- [ ] Moderately
- [ ] Quite a bit
- [ ] Extremely

These questions are about how you feel and how things have been with you during the past 4 weeks. For each question, please give the one answer that comes closest to the way you have been feeling.

6. How much of the time during the past 4 weeks:

a. Have you felt calm and peaceful?

- [ ] All of the time
- [ ] Most of the time
- [ ] A good bit of the time
- [ ] Some of the time
- [ ] A little of the time
- [ ] None of the time

b. Did you have a lot of energy?

- [ ] All of the time
- [ ] Most of the time
- [ ] A good bit of the time
- [ ] Some of the time
- [ ] A little of the time
- [ ] None of the time

c. Have you felt downhearted and blue?

- [ ] All of the time
- [ ] Most of the time
- [ ] A good bit of the time
- [ ] Some of the time
- [ ] A little of the time
- [ ] None of the time

7. During the past 4 weeks, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting with friends, relatives, etc.)?

- [ ] All of the time
- [ ] Most of the time
- [ ] A good bit of the time
- [ ] Some of the time

- [ ] A little of the time
- [ ] None of the time

Now, we'd like to ask you some questions about how your health may have changed.

8. Compared to one year ago, how would you rate your physical health in general now?

	Much better	Slightly better	About the same	Somewhat worse	Much worse
Please Select:	☐	☐	☐	☐	☐

9. Compared to one year ago, how would you rate your emotional problems (such as feeling anxious, depressed, or irritable) now?

	Much better	Slightly better	About the same	Somewhat worse	Much worse
Please Select:	☐	☐	☐	☐	☐

Patient Activation

Section F: PATIENT ACTIVATION

The following are statements that people sometimes make when they talk about their health. Please indicate how much you disagree or agree with each statement as it applies to you personally.

Your answers should be what is true for you and not just what you think others expect of you. **Your choices are *Strongly Disagree, Disagree, Agree,*** or ***Strongly Agree***. If the statement does not apply to you, please say ***Does Not Apply***.

1. When all is said and done, I am the person who is responsible for managing my health.

	Strongly Disagree	Disagree	Agree	Strongly Agree	Does Not Apply
Please Select:	☐	☐	☐	☐	☐

2. Taking an active role in my own health care is the most important factor in determining my health and ability to function.

	Strongly Disagree	Disagree	Agree	Strongly Agree	Does Not Apply
Please Select:	☐	☐	☐	☐	☐

3. I am confident that I can take actions that will help prevent or minimize some symptoms and problems associated with my health.

	Strongly Disagree	Disagree	Agree	Strongly Agree	Does Not Apply
Please Select:	☐	☐	☐	☐	☐

4. I know what each of my prescribed medications does.

	Strongly Disagree	Disagree	Agree	Strongly Agree	Does Not Apply
Please Select:	☐	☐	☐	☐	☐

5. I am confident that I can tell when I need to go get medical care and when I can handle a health problem myself.

	Strongly Disagree	Disagree	Agree	Strongly Agree	Does Not Apply
Please Select:	☐	☐	☐	☐	☐

6. I am confident I can tell a doctor concerns I have even when he or she does not ask.

	Strongly Disagree	Disagree	Agree	Strongly Agree	Does Not Apply
Please Select:	☐	☐	☐	☐	☐

7. I am confident that I can follow through on medical treatments I need to do at home.

	Strongly Disagree	Disagree	Agree	Strongly Agree	Does Not Apply
Please Select:	☐	☐	☐	☐	☐

8. I understand the nature and causes of my health problems.

	Strongly Disagree	Disagree	Agree	Strongly Agree	Does Not Apply
Please Select:	☐	☐	☐	☐	☐

9. I know the different medical treatment options available for my health condition.

	Strongly Disagree	Disagree	Agree	Strongly Agree	Does Not Apply
Please Select:	☐	☐	☐	☐	☐

10. I have been able to maintain the lifestyle changes for my health that I have made.

	Strongly Disagree	Disagree	Agree	Strongly Agree	Does Not Apply
Please Select:	☐	☐	☐	☐	☐

11. I know how to prevent further problems with my health.

	Strongly Disagree	Disagree	Agree	Strongly Agree	Does Not Apply
Please Select:	☐	☐	☐	☐	☐

12. I am confident I can figure out solutions when new situations or problems arise with my health.

	Strongly Disagree	Disagree	Agree	Strongly Agree	Does Not Apply
Please Select:	☐	☐	☐	☐	☐

13. I am confident I can maintain lifestyle changes, like diet and exercise, even during times of stress.

	Strongly Disagree	Disagree	Agree	Strongly Agree	Does Not Apply
Please Select:	☐	☐	☐	☐	☐

Medications

Section G: MEDICATIONS

These next questions are about any medications you might take.

1. Over the past 7 days, how often did you take your prescribed medications as prescribed?

	Very often	Somewhat often	Sometimes	Rarely	Never	Does not apply
Please Select:	☐	☐	☐	☐	☐	☐

2. Over the past 7 days, how often did you skip a dose of prescribed medication?

	Very often	Somewhat often	Sometimes	Rarely	Never	Does not apply
Please Select:	☐	☐	☐	☐	☐	☐

3. Over the past 7 days, how often were you unable to take your prescribed medications at all?

	Very often	Somewhat often	Sometimes	Rarely	Never	Does not apply
Please Select:	☐	☐	☐	☐	☐	☐

4. In the past 2 years since enrolling this study, have you been prescribed any of the following medications by a healthcare provider?

	No	Yes, but I no longer take it	Yes, and I still take it	Don't Know/Unsure	-	What is/are the
Blood pressure medication(s)	○	○	○	○		

-

	No	Yes, but I no longer take it	Yes, and I still take it	Don't Know/Unsure	-	What is/are the
A cholesterol medication, such as atorvastatin (Lipitor), rosuvastatin (Crestor), simvastatin (Zocor), or ezetimibe (Zetia)?	○	○	○	○		

-

	-				What is/are the
	No	Yes, but I no longer take it	Yes, and I still take it	Don't Know/Unsure	
A blood thinner, such as warfarin (Coumadin), apixaban (Eliquis), or rivaroxaban (Xarelto)?	○	○	○	○	

-

	-				What is/are the
	No	Yes, but I no longer take it	Yes, and I still take it	Don't Know/Unsure	
A 5-alpha reductase inhibitor?: a medication used to treat an enlarged prostate. Examples include finasteride (Proscar, Propecia) and dutasteride (Avodart).	○	○	○	○	

-

	-				What is/are the
	No	Yes, but I no longer take it	Yes, and I still take it	Don't Know/Unsure	
A selective estrogen receptor modulator: a medication used to lower the risk of breast cancer. Examples include tamoxifen (Nolvadex, Soltamox), raloxifene (Evista), toremifene (Fareston).	○	○	○	○	

5. In the past 2 years since enrolling in this study, were you recommended by a healthcare provider to take a daily aspirin, either by prescription or over-the-

counter?

- [ ] No
- [ ] Yes, but I no longer take it
- [ ] Yes, and I still take it
- [ ] Don't know/Unsure

Health Behaviors

Section H: HEALTH BEHAVIORS

These next questions are about your health habits.

1. How often would you say that you consume alcohol?

- [ ] Very often
- [ ] Somewhat often
- [ ] Sometimes
- [ ] Rarely
- [ ] Never

2. How often would you say that you exercise?

- [ ] Very often
- [ ] Somewhat often
- [ ] Sometimes
- [ ] Rarely
- [ ] Never

3. How often would you say that you consume processed meat?

- [ ] Very often
- [ ] Somewhat often
- [ ] Sometimes
- [ ] Rarely
- [ ] Never

4. Have you smoked at least 100 cigarettes in your entire lifetime?

- [ ] Yes
- [ ] No
- [ ] Don't know/unsure

Health Behaviors

A. Do you now smoke cigarettes every day, some days, or not at all?

- [ ] Every day
- [ ] Some days
- [ ] Not at all

B. How many years have you or did you smoke cigarettes?

C. On average, how many cigarettes do you smoke per day now? (There are 20 cigarettes in a pack.)

D. On average, over the entire time that you smoked, how many cigarettes did you smoke each day? (There are 20 cigarettes in a pack.)

E. How long has it been since you last smoked a cigarette, even one or two puffs?

- [ ] Within the past month (less than 1 month ago)
- [ ] Within the past 3 months (more than 1 month ago but less than 3 months ago)
- [ ] Within the past 6 months (more than 3 months ago but less than 6 months ago)
- [ ] Within the past year (more than 6 months ago but less than 1 year ago)
- [ ] Within the past 5 years (more than 1 year ago but less than 5 years ago)
- [ ] Within the past 10 years (more than 5 years ago but less than 10 years ago)
- [ ] 10 years or more ago

- [ ] Never smoked regularly
- [ ] Don't know/Not sure

During the past 12 months, have you stopped smoking for one day or longer because you were trying to quit smoking?

- [ ] Yes
- [ ] No
- [ ] Don't know/unsure

Health Behaviors

5. Do you currently use chewing tobacco, snuff, or snus every day, some days, or not at all?

- [ ] Every day
- [ ] Some days
- [ ] Not at all
- [ ] Don't know/unsure

Thank you for completing the GenoVA Study end-of-study survey.

If you have any questions or concerns regarding the survey or the study in general, please contact Katharine MacIsaac at 617-676-8936 or Katharine.Macisaac@va.gov.

The GenoVA Study: Pragmatic Randomized Trial of Polygenic Risk Scoring for Common Diseases in Primary Care

Statistical Analysis Plan

Funding Agency: National Institutes of Health (NIH) / National Human Genome Research Institute (NHGRI)

Principal Investigator: Jason L. Vassy, MD, MPH, SM

Senior Biostatistician: Lee-Jen Wei, PhD

Version Number 1.1

September 7, 2022

Role	Name	Signature	Date
Principal Investigator	Jason Vassy, MD	JASON VASSY 619380 (affiliate) Digitally signed by JASON VASSY 619380 (affiliate) Date: 2022.09.07 15:49:29 -04'00'	
Senior Biostatistician	Lee-Jen Wei, PhD		09/07/2022
Author	Charles Brunette, PhD	CHARLES A. BRUNETTE 1431588 Digitally signed by CHARLES A. BRUNETTE 1431588 Date: 2022.09.07 11:01:27 -05'00'	09/07/2022

List of Abbreviations

ACMG	American College of Medical Genetics and Genomics
AE	Adverse event
AFib	Atrial fibrillation
AUC	Area under the curve
CAD	Coronary artery disease
CDW	Corporate Data Warehouse
CMS	Centers for Medicaid and Medicare
CPRS	Computerized patient record system
EHR	Electronic health record
GEE	Generalized estimating equations
HR	Hazard ratio
ITT	Intention-to-treat
PI	Principal investigator
PCP	Primary care provider
PRS	Polygenic risk score
RCT	Randomized controlled trial
SAE	Serious adverse event
SAP	Statistical analysis plan
SQL	Structured Query Language
T2D	Type II diabetes
UC	Usual care
VABHS	VA Boston Healthcare System

Table of Contents

1.0 Administrative Information

1.1 Trial title and registration

The GenoVA study: a pragmatic randomized trial of polygenic risk scoring for common diseases in primary care

ClinicalTrials.gov Identifier: NCT04331535

1.2 Revision history

SAP version	Protocol version	Section changed	Description	Date amended
1.0	8.0		Date created	08/06/2020
1.1	18.0	2.1 Background; 3.2 Sample Size; 4.1 Study Inclusion and Exclusion Criteria; 5.2 Interim Analysis; General updates	Remove references to Harvard Medical School IRB; Include number of events needed; update inclusion and exclusion algorithm codes; clarify outcomes and event monitoring; general updates due to staff attrition and study modifications.	09/07/2022

1.3 Key personnel

1.3.1 Principal investigator

The principal investigator (PI) supervises all aspects of the study. The PI takes responsibility for the scientific development and conduct of the study, including meeting study goals and timelines, monitoring participant safety, and oversight of the dissemination of research findings.

1.3.2 Senior biostatistician

The senior biostatistician advises the study team on the appropriate study design and statistical analysis of study outcomes. The senior biostatistician conducts and/or reviews sample size and power calculations, provides supervision of the drafting of the SAP, and provides supervision to the data analyst in performing data collection, data cleaning, and statistical analysis of study data.

1.3.3 Health economist

The health economist advises the study team on the appropriate study design and statistical analysis of economic outcomes associated with the study. The health economist conducts and provides supervision to the data analyst in performing data collection, data cleaning, and analysis of economic-related study data.

1.3.4 Data manager/analyst

The data manager creates and maintains the database housing study data. The data analyst ensures the capture of study data and performs requisite merging and cleaning of study data. The data manager prepares summary data tables for study planning, reporting, monitoring, and dissemination of results. The data manager prepares and maintains participant randomization tables and mechanisms for treatment allocation. The data manager does not engage in the enrollment or allocation of participants to study treatments. Under the direction of the senior biostatistician and economist, the data analyst may perform statistical analysis of study outcomes.

1.3.5 Research project manager

The research manager is responsible for the maintenance and update of the GenoVA SAP document and day-to-day operations of the GenoVA Study.

2.0 Introduction

This document details the proposed data analysis, presentation, and reporting of outcomes associated with the GenoVA Study. The results reported in the primary study manuscript(s) will adhere to the strategy outlined here. All amendments to this plan will be documented and reviewed by the relevant key personnel listed within this document. Any deviations to this plan will be justified and detailed in the final manuscript(s). Further analysis, including subset and exploratory analyses not included here, may occur as needed and will be justified and described if reported. This document follows the published guidelines for the content of statistical analysis plans in clinical trials.[1]

2.1 Background

In brief, polygenic risk scores (PRS) combine information from hundreds to millions of genetic loci to develop a quantitative risk measure for individuals' susceptibility to common complex diseases such as coronary artery disease (CAD) or type 2 diabetes (T2D). While the association between PRS and many common diseases are well established (clinical validity), the potential impact of PRS on patient health outcomes (clinical utility) remains unclear. Despite this uncertainty, there is agreement that patient outcomes data, ideally using a prospective design, are needed to inform the clinical utility of PRS.

See GenoVA Study protocol VA Boston Healthcare System (VABHS) 3241 for additional detail regarding study background and rationale.

2.2 Study objectives

The objective of this study is to determine the clinical effectiveness of reporting PRS results among patients at VABHS for 6 common complex diseases [coronary artery disease (CAD), atrial fibrillation (AFib), type II diabetes (T2D), colorectal cancer, prostate cancer, and breast cancer]. In this project, we will conduct a randomized controlled trial (RCT) to:

1. Determine the clinical effectiveness of reporting PRS results among patients at high genetic risk for at least one of 6 common diseases, measured by time-to-diagnosis of prevalent or incident disease over 24 months (primary outcome).

2. Measure and assess changes in the following high-priority genomic medicine implementation outcomes: change in clinical management and evidence of diagnostic testing related to disease risk, patient activation in healthcare, self-reported medication adherence, and healthcare costs (secondary outcomes).

3. Measure and assess additional clinical and behavioral outcomes, such as medication use, smoking status, and body-mass index (BMI), and primary care provider (PCP) knowledge and beliefs about PRS (exploratory outcomes).

2.3 Primary outcome and research hypothesis

Primary outcome: Time-to-diagnosis of at least one of 6 common complex diseases.

Null hypothesis: Time-to-diagnosis of at least one of 6 common complex diseases does not differ between patients who receive PRS test results compared to patients undergoing usual care (UC) 24 months after randomization.

Alternative hypothesis: Time-to-diagnosis of at least one of 6 common complex diseases will be lower in patients who receive PRS test results compared to patients undergoing UC 24 months after randomization.

3.0 Trial Methods

3.1 Trial design

The GenoVA Study is a point-of-care pragmatic randomized controlled trial. Patients with at least one PRS indicating high risk (OR_{PRS}>2.0) for any of the 6 target diseases, and without a confirmed American College of Medical Genetics and Genomics (ACMG) actionable monogenic variant, are randomized in a 1:1 ratio to either receive their high-PRS results report at baseline (PRS-high arm) or after a 24-month observation period (UC-high arm). Similarly, the stratum of patients with PRS indicating average genetic risk are randomized in a 1:1 ratio to receive their PRS results report at baseline (PRS-average arm) or after 24 months (UC-average arm). Any participant with a confirmed ACMG variant (estimated 1-2% of participants) is ineligible for randomization and will instead be followed for study outcomes, although their study data will not be analyzed with the concurrent control group.

3.2 Sample size

A total of 1,076 patient-participants are genotyped.

Based on published estimates of the prevalence of high-risk PRS values, we make the conservative assumption that 33% of genotyped patient-participants have at least one OR_{PRS}>2.0 across all diseases of interest (Table 1).

In review of VABHS data from 2014-2017 (Table 2), based on published algorithms using ICD and CPT codes, among patients 50-70 years old, an average 6.2% had a new target diagnosis per year. If we assume that 12% of patients in the control arm will have a new target diagnosis over the 2-year study period and that delivery of high-PRS results and recommendations will increase this to 24% of patients in the intervention arm during the same period (RR=2), a total sample size of 320 patient-participants must be included in the RCT to detect this difference at a two-tailed α=0.05 and β=0.20 (power of 80%). Converting this rate difference to a hazard ratio metric via exponential distributions for time-to-event analysis would result in an estimated hazard ratio (HR) of 0.44 (UC-high versus PRS-high). For this target hazard ratio, a minimum of 46 total events would be needed among high-risk participants for a positive trial. This approach tends to be more powerful than using the event rate difference at 24-months. If 33% of

genotyped patient-participants have at least one OR_{PRS}>2.0 and are enrolled in the RCT, a total sample size of 960 genotyped patient-participants is needed.

To account for potential clustering effect among patients receiving care from the same providers, we include a design effect of 1.10, based on an estimate of 7 high-risk enrollees per PCP and an intraclass correlation coefficient of 0.02.[2] As a result, a total of 960 x 1.1 = 1,056 participants would be needed. We will enroll a total of 1,076 to account for the 1-2% of participants who will be excluded from randomization due to an ACMG variant.

Condition	Cases / controls (n) from recent GWAS	Reported PRS Area Under the Curve (AUC)	Reported prevalence of high-risk PRS
CAD	60,801 / 123,504[3]	0.806[4]	8% (OR≥3)[4] 10% (OR≥2.9)[4] 20% (OR≥2.5)[4]
Atrial fibrillation	65,446 / 522,744[5]	0.773[4]	6.1% (OR≥3)[4] 10% (OR≥2.7)[4] 20% (OR≥2.5)[4]
T2D	74,124 / 824,006[6]	0.66[6] 0.725[4]	2.5% (OR≥3.4)[6] 3.5% (OR≥3)[4] 10% (OR≥2.52)[4] 20% (OR≥2.19)[4]
Colorectal cancer	61,985 / 101,330[7]	0.628[8] 0.733[7,9]	1% (RR≥2.9)[8] 4.3% (OR ≥2)[7]
Breast cancer	137,045 / 119,078[10]	0.685[4]	1.5% (OR≥3)[4] 5% (OR≥2.59)[4] 10% (OR≥2.36)[4]
Prostate cancer	79,194 / 61,112[11,12]	0.68[13]	10% (OR≥2.69) [11,12] 2% (HR≥2.9)[14]

Table 1: PRS performance and prevalence from selected recent reports

OR in Khera 2018[4] compare top percentiles shown to the remainder of population. RR in Frampton 2016[8] compared to population median. OR in Schmit 2018[7] and Schumacher 2018[11,12] compared to 25th–75th percentile. HR in Seibert 2018[14] compares time-to-diagnosis of aggressive prostate cancer to 30th-70th percentile.
Abbreviations: ExWAS, exome-wide association study; HR, hazard ratio; OR, odds ratio; RR, relative risk

	2014	2015	2016	2017	Average
Total eligible patients	6,145	6,083	6,016	5,770	6,003
Total patients diagnosed	372(6.1%)	340(5.6%)	388(6.4%)	400(6.9%)	375(6.2%)
By disease					
CAD	113(1.8%)	106(1.7%)	126(2.1%)	130(2.3%)	119(2.0%)
Atrial	51(0.8%)	60(1.0%)	71(1.2%)	62(1.1%)	61(1.0%)
fibrillation T20	156(2.5%)	113(1.9%)	127(2.1%)	156(2.7%)	138(2.3%)
Colorectal cancer	8(0.1%)	10(0.2%)	13(0.2%)	11(0.2%)	11(0.2%)
Breast cancer	1(0.0%)	1(0.0%)	3(0.0%)	5(0.0%)	2(0.0%)
Prostate cancer	43(0.7%)	50(0.8%)	48(0.8%)	36(0.6%)	44(0.7%)

Table 2. Annual rates of new diagnoses for six target diseases in patients between the ages of 50-70 years old across VA Boston Healthcare System, 2014-2017, by disease

Diagnosis estimates based on age or first diagnosis or at least one or the six target diseases. Annual estimates based on primary care relationship over two-year period with no new diagnosis or a target disease during first year.

3.3 Randomization

Study staff use pre-generated randomization tables to allocate participants to a study arm based on their PRS results. Pre-generated randomization tables are created using standard statistical software (e.g. computerized random block and sequence generation) by the GenoVA Study data manager, under the direction of the senior biostatistician, and stored in a secure file share accessible to select study staff. Randomization occurs upon the receipt of each participant's PRS report and is mechanized through a computerized randomization tool. Study staff enrolling and allocating participants to study treatments are blinded to the pre-generated randomization tables.

3.3.1 High genetic risk and average genetic risk groups

Pre-generated randomization tables stratify participants by PRS threshold (high versus average risk) and sex (male versus female) with 1:1 allocation using a permuted block design with a block size of 4.

3.3.2 Participants with a confirmed ACMG variant

Any participant with a confirmed ACMG variant (estimated 1-2% of participants) is ineligible for randomization but is followed for study outcomes for exploratory analyses.

3.4 Data sources, collection, and storage

Study outcomes data will be collected from the VA Corporate Data Warehouse (CDW)[15], a repository of administrative and clinical data from the VA's nationally deployed electronic health record (EHR) system (computerized patient record system, CPRS), clinical chart review of the EHR, participant baseline and follow-up surveys, Centers for Medicaid and Medicare (CMS) data, and trial operations data recorded by the study team. All study data will be stored, cleaned, and analyzed within a secure VA computing environment and will be accessible to authorized study staff only.

3.5 Stopping guidance

This study has no stopping rules. Enrolled participants may withdraw from the study at any time.

3.6 Protocol deviations

Protocol deviations are characterized as circumstances that depart from planned study procedures and anticipated events (e.g. participant withdrawal, loss to follow-up). Protocol deviations may include, but are not limited to, the following:

1. Deviation from inclusion or exclusion criteria (e.g. ineligible patient enrolled and/or randomized)
2. Patient receipt of treatment other than treatment as randomized

The number of ineligible patients randomized, patients receiving a treatment other than as randomized, or other yet to be determined protocol deviations, if any, will be characterized and reported in the final manuscript(s). For the purposes of primary and secondary outcomes, data from patients who experience a protocol deviation will be included in the final datasets. Their outcomes data will be analyzed as part of the treatment group to which they were randomly allocated. The inclusion or exclusion of these patients' data in subsequent secondary or subgroup analyses will be detailed in the final manuscript(s) as needed.

3.7 Adverse events

An Adverse Event (AE) will be defined as any unanticipated or unintended medical occurrence, which does not necessarily have a causal relationship with the study condition, procedure(s), or study agent(s), that occurs after informed consent is obtained. A serious adverse event (SAE) will be defined as an AE resulting in one of the following outcomes:

1. Death during the 24 months after enrollment
2. Life-threatening event (defined as an event that places a participant at immediate risk of death)
3. Inpatient hospitalization
4. Any other condition which, in the judgment of the investigator, represents a significant hazard, such as an important medical event that does not result in one of the above outcomes.

The number of AEs or SAEs (e.g. deaths, hospitalizations), if any, will be characterized and reported in the final manuscript(s). For the purposes of primary and secondary outcomes, data from patients who experience an AE or SAE will be included in the final datasets. Their outcomes data will be analyzed as part of the treatment group to which they were randomly allocated, regardless of the treatment they actually receive. The inclusion or exclusion of these

patients' data in subsequent secondary or subgroup analyses will be detailed in the final manuscript(s) as needed.

4.0 Trial Population

The overall study population includes all Veteran patients from primary care and women's health clinics across the VABHS healthcare system, between the ages of 50-70, and with no known diagnoses of one or more of the 6 target diseases (CAD, AFib, T2D, colorectal cancer, prostate cancer, and breast cancer) at the time of study consent and enrollment.

4.1 Study inclusion and exclusion criteria

Inclusion criteria:

1) Age 50-70 years at enrollment

2) Has had a VABHS admission or visit for primary or specialized care within the previous 24 months

3) No known diagnosis of the following conditions, initially screened by International Classification of Disease (ICD) codes or other EHR data using validated methods and then confirmed with potential patient-participants during recruitment:

a) Coronary artery disease: At least 2 occurrences of any of the following codes, documented on different dates: ICD-9 Codes 410-412, 414 (except 414.1x), 429.7, 996.03, V45.81-V45.82 or ICD-10 Codes I20-I24, 125 (except I25.3 and I25.4), T82.21, T82.855, Z95.1, Z95.5, Z98.61 or ICD-9 Procedure Codes 36.0 - 36.3, 00.66, ICD-10 Procedure Codes 0210-0213, 021K-021L, 0270-0273, 02C0-02C3, 3E07x1x, 3E07xPx, 3E074GC, or CPT Codes 33510-33545, 92973, 92975, 92977, 92980-92984, 92995-92996

b) Atrial fibrillation: At least 2 occurrences of any of the following codes, documented on different dates by providers other than pharmacists and pharmacy technicians: ICD-9 Codes 427.3 or ICD-10 Codes I48 or ICD-9 Procedure Codes 37.33, 37.34, 99.61 or ICD-10 Procedure Codes 0256-0257, 5A2204Z, or CPT codes 33250-33259, 33260-33266, 92960-92961

c) Type 2 diabetes: At least 2 occurrences of any of the following codes, documented on different dates: ICD-9 Codes 250, 357.2, 362.0, 366.41 or ICD-10 Codes E10-E11 or inpatient dose, outpatient VA prescription, or non-VA self-reported prescription of medications ever listed in the VA National formulary under the antidiabetic/hypoglycemic classes HS500, HS501, HS503, and HS509

d) Colorectal cancer: 1 occurrence of any of the following codes documented in the administrative oncology tables or at least 2 occurrences of any of the following codes, documented on different dates: ICD-9 Codes 153.x (except 153.5), 154.0, 154.1, 230.3, 230.4, V10.05, V10.06 or ICD-10 Codes C18.x (except C18.1), C19, C20, C26, D01.0 - D01.4, Z85.038, Z85.048 or ICD-O-3 site coded C18.x (except C18.1), C19.9, C20.9, or

ICD-9 Procedure Codes 17.31-17.36, 45.71-45.76, 45.81-45.83 or ICD-10 Procedure Codes ODT[EFGHKLMNP], ODB[EFGHKLMNP] or CPT Codes 44140-44160, 44204-44212

e) Breast cancer: 1 occurrence of any of the following codes: ICD-9 Codes 174, 175, 233.0, V10.3 and ICD-10 Codes C50 - C50.9, D05, Z853 or ICD-O-3 Oncology tumor locations C50.x or ICD-9 Procedure Codes 85.20, 85.21, 40.22, 40.23, 85.22, 85.23, 85.33-85.36, 85.41-85.48 or ICD-10 Procedure Codes 0HTT*, 0HTU*, 0HTV* or CPT Codes 19120, 19125, 19126, 19160, 19162, 19180, 19182, 19200, 19220, 19240, 19300-19307

f) Prostate cancer: 1 occurrence of any of the following codes: ICD-9 Codes 185, 222.2, 236.5, 233.4, V10.46 and ICD-10 Codes C61, D07.5, D29.1, D40.0, Z85.46 or ICD-9 Procedure Codes 60.2-60.5, 60.62, 60.69 or ICD-10 Procedure Codes 0V[5BT]0 or CPT codes 52500, 52601, 52606, 52612, 52614, 52620, 52630, 52640, 52647-52649, 52650, 52873, 52859, 52860, 52862, 52865, 52866, 55801, 55810, 55812, 55815, 55821, 55831, 55840, 55842, 55845 or ICD-O-3 sites coded C19.x

There are no other exclusion criteria for this study.

4.2 Screening, recruitment, and withdraw

EHR data from the CDW is screened using a Structured Query Language (SQL) algorithm per the study's inclusion and exclusion criteria to identify potentially eligible patients for study recruitment (see GenoVA Study protocol VABHS 3241 for additional detail regarding the study recruitment process). A description of the final screening algorithm, including ICD codes, procedure codes, medications, and other criteria, and its performance will be reported in the final study manuscript(s). Duration of the study recruitment period, the total number of patients screened, the number of screened patients not recruited and reason for non-recruitment, and other screening and recruitment metrics will be collected and reported for the overall study. In addition to protocol deviations and AEs or SAEs, if any, the number, and reasons (if known) for participant withdrawal and/or loss to follow-up prior to the conclusion of the study's period of enrollment, will be reported in the final manuscript(s). Participant flow is detailed via the proposed diagram and is reported using CONSORT guidelines for the reporting of clinical trials[16,17]:

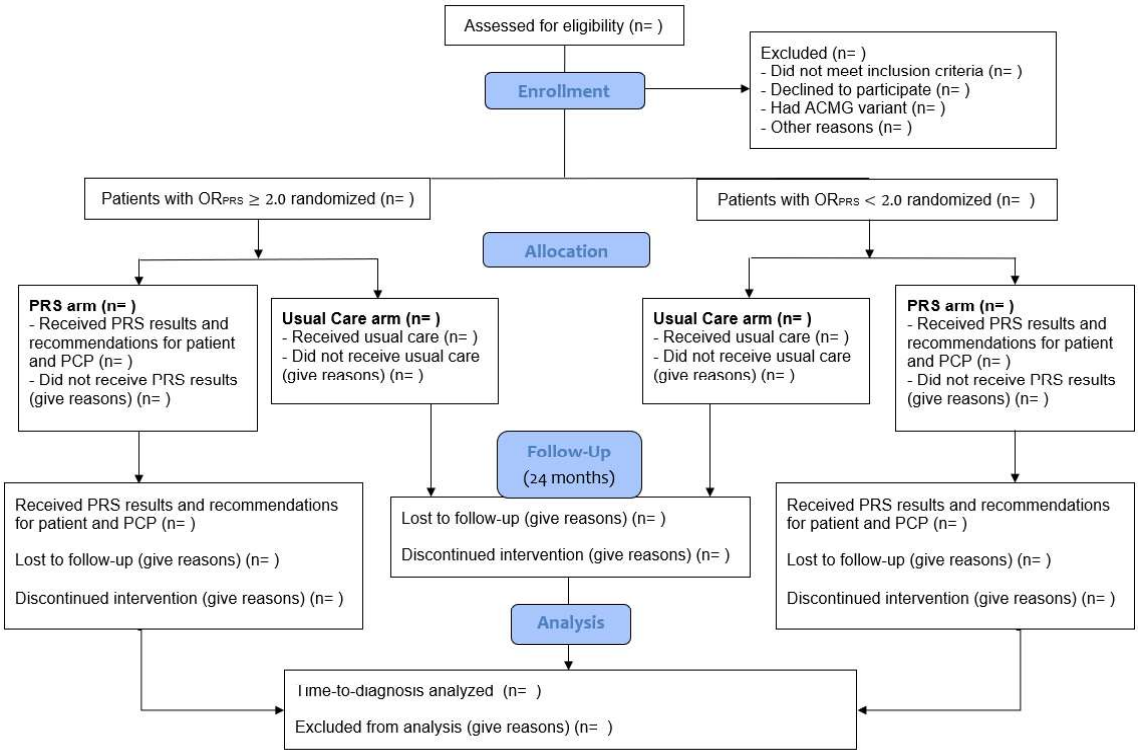


Figure 1. CONSORT Template for the GenoVA Study

4.3 Reference start and end dates

Participant study enrollment occurs on the date of consent. Randomization occurs upon genotype analysis and receipt of the PRS report by study staff, since participants with an ACMG result are excluded from randomization and because randomization is stratified by the presence or absence of at least one high-risk PRS result. Study participation concludes upon completion of the 24-month observation period from a participant's date of randomization and end-of-study survey. End-of-study surveys may be completed up to 3 months after the end of the 24-month observation period. However, any relevant diagnosis or diagnostic procedure reported on the end-of-study survey will only count towards the primary and secondary outcomes, respectively, if clinician review of medical records confirms the outcome to have occurred within the 24-month observation period.

4.3.1 Baseline assessment

For the purposes of primary and secondary outcomes assessment, baseline is defined as the most recent measurement of a study-related outcome (see Section 6.0) on or prior to a participant's date of randomization, unless otherwise specified. In the event a participant is randomized to a treatment arm prior to the measurement of a pre-specified baseline measure (e.g. study measures assessed by survey), the study team will consider the measurement of a study-related outcome within 60 days from the participant's date of randomization as their baseline measurement. The total number of participants with baseline measurements obtained

post-randomization and/or any statistical analysis including a baseline measurement obtained post-randomization will be reported. Baseline measurements not obtained after 60 days from the participant's date of randomization will be considered missing.

4.3.2 End-of-study follow-up and period of observation

For the primary outcome (time-to-diagnosis of a common complex disease), dates of diagnosis will correspond to the date that clinical chart reviewers agree that a clinically significant diagnosis has been made in the medical record. Additional medical records will be requested from the participant, if needed, to substantiate the diagnosis and date. For each diagnosis, time-to-diagnosis will be defined as the difference between the dates of diagnosis and randomization. Diagnosis dates occurring later than 24 months from the randomization date will be considered censored.

For secondary outcomes assessment, end-of-study follow-up or observation is defined as the most recent measurement of a study-related outcome (Section 6.0) on or prior to the date 24 months from a participant's date of randomization, unless otherwise specified. 24-month observation for participants with a confirmed ACMG variant is defined as the date 24 months from the date of the study staff's receipt of the participant's genetic result.

For self-reported outcomes (e.g. study outcomes assessed by survey), the study team will consider the measurement of a study-related outcome prior to or within 90 days from the participant's 24-month date from randomization as their end-of-study follow-up measurement. The total number of participants with follow-up measurements obtained after their 24-month date will be reported. Follow-up measurements not obtained after 90 days from the participant's 24-month date of follow-up will be considered missing.

4.4 Analysis populations

4.4.1 Intention-to-treat populations

Intention-to-treat (ITT) populations include all patients who undergo randomization and are characterized by the treatment they were randomized to receive (PRS-high versus UC-high arms or PRS-average versus UC-average arms).

4.4.2 Complete case populations

The complete case populations consist of patients who undergo randomization (PRS-high versus UC-high arms or PRS-average versus UC-average arms), and complete all study assessments, including both the baseline and end-of-study patient surveys.

4.4.3 Subgroup populations

Additional subgroups of the study population may include the analyses of patients stratified by demographic (e.g. age), randomization (e.g. PRS arm(s), UC arm(s)), and/or outcome characteristics (e.g. high versus average disease risk, diagnosed disease). Each subgroup population will be described in detail as reported in the final study manuscript(s).

4.5 Baseline patient characteristics

Baseline characteristics are summarized and presented for participants in the ITT populations. Standard statistical summaries, depending on data type and distribution, are presented as 1) total numbers of participants with each characteristic (n) and as a proportion (%) of each group stratified by randomization arm or 2) as means and standard deviations (if normally distributed) or medians and interquartile ranges (IQR) (if nonnormally distributed) stratified by randomization arm. No statistical testing will be carried out for participant baseline characteristics or measures between treatment groups.

The below participant baseline characteristics, including the pre-specified baseline measurements of the study outcomes described in Section 6.0, will be derived and reported:

Baseline characteristic	**How derived**	**Presentation**
Age in years	As calculated using EHR administrative data relative to date of consent and enrollment.	mean (SD) / median (IQR)
Gender / Sex	As determined by EHR administrative data.	n (%)
Race	As determined by EHR administrative data and/or data collected from baseline survey.	n (%)
Ethnicity	As determined by EHR administrative data and/or data collected from baseline survey.	n (%)
Socioeconomic status	As determined by EHR administrative data.	n (%)
Self-reported health status and quality of life	As determined by data collected from baseline survey. (VR-12)[18-20]	mean (SD) / median (IQR)
Self-reported patient activation	As determined by data collected from baseline survey. (PAM-13)[21]	mean (SD) / median (IQR) / n (%)
Self-reported medication adherence	As determined by data collected from baseline survey. (Voils 3-item medication adherence survey)[22]	mean (SD) / median (IQR)
Self-reported aspirin use	As determined by data collected from baseline survey.	n (%)
Self-reported physical activity	As determined by data collected from baseline survey. (Single-item Likert response)	mean (SD) / median (IQR) / n (%)

Self-reported alcohol intake	As determined by data collected from baseline survey. (Single-item Likert response)	mean (SD) / median (IQR) / n (%)
Self-reported processed meat consumption	As determined by data collected from baseline survey. (Single-item Likert response)	mean (SD) / median (IQR) / n (%)
Self-reported smoking status	As determined by data collected from baseline survey. (BRFSS Code Section 9, Tobacco Use)[23]	n (%)
Systolic blood pressure	As determined by EHR data.	mean (SD) / median (IQR)
Diastolic blood pressure	As determined by EHR data.	mean (SD) / median (IQR)
Body-mass index	As determined by EHR data.	mean (SD) / median (IQR)
Low-density lipoprotein cholesterol	As determined by EHR data.	mean (SD) / median (IQR)

5.0 Statistical Considerations

5.1 Statistical framework

The principal intention-to-treat (ITT)[24] analysis compares the PRS-high and UC-high arms. Using a Cox model and post-randomization stratification[25,26] by disease categories, analysis of the primary outcome uses a Wald statistic, and two-sided type I error rate of 0.05 to test the null hypothesis of no difference between the high-PRS intervention (PRS-high arm) compared to the control group (UC-high arm) based on total rate of new diagnoses at month 24 after randomization. To quantify the treatment arm difference for the primary endpoint, a hazard ratio will be presented with a corresponding confidence interval estimate. Time-to-diagnosis for specific diseases between treatment arms are reported separately and considered descriptive in nature. Similarly, for secondary and other pre-specified outcomes an ITT approach is used to make outcomes comparisons across treatment groups. Additional exploratory ITT analyses make outcomes comparisons between the PRS-average and UC-average arms, between all PRS and UC participants, adjusting for risk group, and between all high-risk and average-risk participants, adjusting for randomization arm. Subset analyses, including sensitivity and group analyses by disease, randomization stratum, or other study features or patient characteristics, are also considered exploratory and will be described in detail if reported.

5.2 Interim analyses

No formal interim hypothesis testing for superiority is planned; however, data will be reviewed by an independent monitoring committee to determine whether the trial is on pace to observe an adequate number of events by the end of the study and to ensure participant safety. Interim assessment by the independent monitoring committee will occur after half (~23) of the minimum

number of events needed (46) is observed via assessment of electronic health record data. In the case of fewer than required observed events among high-risk participants, a recalculation of required sample size using unblinded efficacy data may be conducted and recommended by the independent monitoring committee. Moreover, in the presence of fewer than expected high genetic risk participants (based on OR_{PRS}>2.0), the PRS threshold defining high-risk may be lowered, to ensure enrollment of the target sample size.

5.3 Timing of final analyses

Final analyses of the GenoVA Study data are conducted upon the conclusion of the final participant's 24-month enrollment and completion of requisite data collection via CDW, the EHR, and participant surveys.

5.4 Confidence intervals, *P* values, and multiple testing

All statistical testing is reported with an effect, a two-sided 95% confidence interval, and *P* value, unless otherwise specified. *P* values less than 0.05 will be reported as significant for the primary outcome. *P* values reported as significant for secondary outcomes will undergo Bonferroni correction for multiple hypothesis-testing. Other pre-specified and post-hoc analyses are considered exploratory.

5.5 Missing data

Prior to statistical analysis, outcomes data are reviewed for amount and pattern of data missingness (e.g. missing at random) using standard statistical software and methods. For outcomes analysis, partially observed covariates and outcomes are imputed using fully conditional specification, or comparable methods, as appropriate.[27-29] Any necessary imputation will be conducted separately within each treatment arm. Imputation models will include the same covariates (i.e. demographic factors such as age, gender, socioeconomic status, and baseline health) as those used in final outcomes analysis. Proportions of data missingness, reasons for missingness (if known), and methods used for data imputation if required, including number of imputations and sensitivity analyses performed, will be reported in the final manuscript(s).

5.6 Statistical assumptions and issues

Prior to analysis and hypothesis testing, statistical assumptions are evaluated for each proposed outcome assessment. The presence of distributional assumptions, influential outliers, multicollinearity, and proportionality of hazards, among other common assumptions related to the analyses described here, are assessed. Methods used and results of the assessment of statistical assumptions (e.g. Schoenfeld residuals) will be acknowledged in the final study manuscript(s). In addition, issues related to significant differences between withdrawn, censored, and remaining cases as well as changes in study methods over time (e.g. change in study procedures that result in materially different patient outcomes) will be considered. A

description of unusual outliers, violated assumptions, or other issues that may impact the integrity of the analyses, and any corrective action (e.g. assumptions tested, removal of outliers, variable transformations, etc.) will be described in the final study manuscript(s) as applicable.

5.7 Clustered data

There is potential for small clustering effect among patients receiving care from the same PCPs[2,30] at VABHS with respect to select study outcomes, specifically disease diagnoses, diagnostic testing, and medication prescriptions. As a result, provider clustering is taken into account when analyzing these study outcomes using the method for clustered survival data proposed by Lee et al.,[31] which are detailed in Section 7.0. Given the similarities in clinical operations across all VABHS clinics and limited geographic scope, no clinic effect will be included in the statistical models. Participant-level self-report measures and other routinely collected clinical data will be assumed independent between study participants.

5.8 Statistical software

Statistical analysis will be conducted using appropriate and validated software, including SAS, STATA, R, or other comparable statistical programs. The applicable software(s), package(s), and version(s) used for the analyses of study data will be reported in the final manuscript(s).

6.0 Outcome definitions and timing

6.1 Primary outcome (Time-to-diagnosis of common complex disease))

The primary outcome of the trial is time-to-diagnosis of at least one of 6 common complex diseases. Time-to-diagnosis is assessed at 24 months after randomization for each participant in the high-risk stratum (PRS-high versus UC-high). If a participant is diagnosed with more than one of these diseases during the observation period, all relevant diagnoses are counted separately toward the primary outcome, with analysis accounting for correlated time-to-event data within an individual. Time-to-diagnosis is measured in total days from the participant's date of randomization to an initial date of diagnosis occurring within the participant's 24-month observation period. Chart review is done independently in duplicate by clinical experts blinded to participant PRS results and randomization status. Diagnosis and date of diagnosis are abstracted for all participants by expert clinical chart review using gold-standard diagnostic criteria and clinical judgment for each of the 6 common diseases. If a participant reports a new diagnosis of one of the 6 target diseases on the end-of-study survey, additional medical records will be requested from the participant for the clinical chart reviewers, if needed, to substantiate the diagnosis and date.

Time-to-diagnosis of any of the 6 target diseases will be collected for all patients, including those at average genetic risk for all conditions and those receiving a confirmed ACMG variant result. Although not a part of the primary outcome, these diagnoses will be used in exploratory analyses.

For all time-to-diagnosis outcomes, the date of randomization (or the date an ACMG variant result is reported) is considered the time origin ('Day 0') and time-to-diagnosis is derived by:

[total days to diagnosis] = ([date of initial diagnosis] – [date of randomization]) + 1

Cases with negative total days to diagnosis are considered diagnosed with disease at time origin.

In the event an exact date of diagnosis is unable to be determined, a date of diagnosis will be derived as follows:

1. If only a month and year of diagnosis is identified, day will be derived as the 15th day of the observed month.
2. If a month and year of diagnosis cannot be determined, but diagnosis is substantiated from the medical record, a date of diagnosis will be derived from the first date in which the patient's provider documents the presence of or initiates treatment specific to a disease of interest as determined by expert clinical chart review.

For each participant, diagnosis of each of 5 target diseases, relative to designated sex (only men assessed for prostate cancer and women assessed for breast cancer), that occurs within the 24-month observation period are coded as either a '1' (disease diagnosed) or '0' (disease not diagnosed). A date of diagnosis and derivation of time-to-diagnosis (e.g. days) accompany each participant's confirmed disease diagnosis.

Cases with no documented diagnosis events (i.e. no diagnoses of one of 6 common complex diseases) on or before the date 24 months from the date of randomization will be considered terminally censored. Patients who experience an AE or SAE rendering them unable to continue study participation or who voluntarily withdraw from the study are considered censored at the time of event, unless the event is associated with an outcome of interest as determined by expert clinical chart review (e.g. AE or SAE or withdraw occurs in conjunction with the diagnosis of one of 6 common complex diseases). In this case, patients will be considered to have experienced a disease diagnosis. We assume censoring to be independent and non-informative, unless otherwise determined. The number of events, description of events (e.g. disease diagnosis), censored subjects, and reasons for censoring (if known) will be reported in the final manuscript(s).

6.2 Secondary outcomes

6.2.1. Diagnostic testing of common complex disease

Evidence of diagnostic testing for any of the 6 target diseases are assessed at 24 months after randomization for all patient-participants. Evidence of diagnostic testing is identified through both CDW structured data collection and expert clinical chart review of VA and external medical records.

For each participant, diagnostic testing for each of 5 target diseases, relative to their designated sex (only men assessed for prostate cancer and women assessed for breast cancer), that occurs within the 24-month observation period are coded as either a '1' (diagnostic testing undertaken) or '0' (no diagnostic testing undertaken). A date of diagnostic testing and description of the type of diagnostic test used accompany each observation.

The following procedures are considered diagnostic for the purposes of the GenoVA Study:

a) CAD: Stress testing, cardiac CT for coronary artery calcium (CAC), coronary angiography

b) AFib: Electrocardiogram (ECG), heart rhythm monitoring by a provider-ordered modality (i.e. not with personal smartwatch or other consumer wearable)

c) T2D: Hemoglobin A1c, glucose tolerance test

d) Colorectal cancer: Colonoscopy, sigmoidoscopy, fecal blood testing, CT colonography

e) Breast cancer: Mammography, breast MRI, breast ultrasound, breast biopsy

f) Prostate cancer: PSA testing, prostate ultrasound, prostate biopsy

In the high-risk stratum (PRS-high and UC-high), diagnostic testing for any of the 6 target diseases will be considered a diagnostic test for the main analysis of this outcome. Exploratory pre-specified analyses will examine any diagnostic testing related to any of the six diseases in high-risk stratum and in the average-risk stratum.

6.2.2. Self-reported patient activation

Self-reported understanding, competence, and willingness to participate in health care decisions and processes are assessed via the baseline and end-of-study surveys, using the 13-item short form of the Patient Activation Measure (PAM-13).[21] Each PAM-13 item has four possible response options: "*Strongly disagree*" (1), "*Disagree*" (2), "*Agree*" (3), "*Strongly agree*" (4), as well as "*Does not apply*" (0). Response values are summed, divided by the total number of items responded to (excluding selections of non-applicable items), and multiplied by 13. The raw score is converted using a scoring table to derive both a linear score from 0 (no activation) to 100 (fully activated) and interval patient activation scores (1: activation not important, passive recipient of care; 2: lack of knowledge or confidence to take action; 3: beginning to take action; 4 taking action). Both a continuous (0-100) and interval score are derived for each patient at baseline and 24 months after randomization.

6.2.3. Self-reported medication adherence

Self-reported taking of medications as prescribed is assessed via the baseline and end-of-study surveys, using the 3-item Voils Medication Adherence Survey.[22] Each item has five possible response options: "*Never*" (1), "*Rarely*" (2), "*Sometimes*" (3), "*Somewhat often*" (4), and "*Very often*" (5). A total score reflecting nonadherence to medication taking is calculated as a numerical measure by averaging participant responses from item-1 (reverse-scored), item-2, and item-3. A measure of medication nonadherence is derived for each participant at baseline and 24 months after randomization.

6.2.4. Healthcare costs

A combination of administrative data and microcosting approaches are used to estimate costs over the 24 months after randomization. Estimates of the infrastructure and personnel needed to deliver the intervention are derived empirically from the study. Healthcare costs are abstracted from billing and administrative data from the CDW and CMS data.

6.3 Other pre-specified outcomes

6.3.1 Self-reported health status and quality of life

Self-reported health status and quality of life is assessed via the baseline and end-of-study surveys, using the Veterans RAND 12-Item Health Survey (VR-12).[18-20] The VR-12 computes two continuous composite scores, a physical component summary (PCS) and a mental component summary (MCS).

6.3.2 Blood pressure

The most recent systolic and diastolic blood pressure values are ascertained and recorded from the medical record for each participant on or prior to the date of randomization and on or prior to the date 24 months after randomization. Measures for systolic and diastolic blood pressure are presented as continuous values and units of measure are in millimeters of mercury (mmHg).

6.3.3 Body-mass index (BMI)

The most recent BMI values are ascertained and recorded from the medical record for each participant on or prior to the date of randomization and on or prior to the date 24 months after randomization. BMI is presented as a continuous value and units of measure are kilogram per meter squared (kg/m^2).

6.3.4 Self-reported aspirin use

Self-reported prescription or over-the-counter aspirin use is assessed via the baseline and end-of-study surveys, using a single-item question. Aspirin use is recorded and presented as a categorical value at baseline coded as '0' (not taking aspirin), '1' (taking aspirin), or '2' (unsure). End-of-study survey options reflect potential alterations to aspirin use since enrollment coded as

‘0’ (aspirin not recommended by healthcare provider), ‘1’ (aspirin recommended, but no longer taking), ‘2’ (aspirin recommended and still taking), and ‘3’ (unsure). For the purposes of pre-specified outcomes analyses, end-of-study aspirin use is collapsed into a dichotomous variable (not taking ‘0’ versus taking ‘1’).

6.3.5 Self-reported physical activity

Self-reported physical activity is assessed via the baseline and end-of-study surveys, using a single-item question “*How often would you say that you exercise?”* Response options are recorded as an ordinal Likert response ranging from *“Never”* (0) to *“Very often”* (4).

6.3.6 Self-reported alcohol intake

Self-reported alcohol intake is assessed via the baseline and end-of-study surveys, using a single-item question “*How often would you say that you consume alcohol?”* Response options are recorded as an ordinal Likert response ranging from *“Never”* (0) to *“Very often”* (4).

6.3.7 Self-reported processed meat consumption

Self-reported processed meat consumption is assessed via the baseline and end-of-study surveys, using a single-item question “*How often would you say that you consume processed meat?”* Response options are recorded as an ordinal Likert response ranging from *“Never”* (0) to *“Very often”* (4).

6.3.8 Low-density lipoprotein cholesterol (LDL-C)

The most recent LDL-C values are ascertained and recorded from the medical record for each participant on or prior to the date of randomization and on or prior to the date 24 months after randomization. LDL-C is presented as a continuous value and units of measure are milligrams per deciliter (mg/dL).

6.3.9 Self-reported smoking status

Self-reported smoking status is assessed via the baseline and end-of-study surveys, using the 5-item “Tobacco Use” instrument from the Behavioral Risk Factor Surveillance System (BRFSS) (Core Section 9).[23]

6.3.10 Risk-reducing medication prescriptions

Relevant prescription medication changes during the 24-month observation period, including antihypertensives, cholesterol-lowering medications, anticoagulants, antiplatelet medications, diabetes medications, 5-alpha reductase inhibitors, selection estrogen receptor modulators, aromatase inhibitors, as collected from the CDW, the baseline and end-of-study surveys, and from clinical chart review.

6.3.11 Provider knowledge and beliefs about PRS

Semi-structured interviews assess participating providers' understanding of and perceived utility of PRS risk information.

7.0 Analysis methods

7.1 Covariate adjustment

For the purposes of primary, secondary, and pre-specified outcomes analysis, statistical models include sex as a covariate due its use as a stratification factor for randomization.[32-34] Disease categories are included as covariates in statistical models as post-randomization stratification factors where described.[25,26]

7.2 Primary outcome

7.2.1 High genetic risk group (primary research hypothesis)

ITT analyses compare the PRS-high and UC-high arms among all participants who undergo randomization in the high-risk group. The primary outcome is time-to-diagnosis of at least one of 6 common complex diseases, as described in Section 6.1. The analysis is based on the rate of new diagnoses (either undiagnosed prevalent or incident cases) at month 24 after randomization for the PRS-high and UC-high arms. Diagnosis among those at high genetic risk is characterized by a change in participant state from apparently non-diseased to diseased, as determined by expert clinical chart review during the 24-month observation period. Differences between the PRS-high and UC-high arms quantify the impact of telling high-risk patients and their PCPs about their high risk.

We use Cox modeling[31,35-38] and post-randomization stratification by disease categories to analyze time-to-diagnosis by treatment arm, specifying disease status (diseased versus not diseased) and time in days to disease diagnosis or other censoring as described in Section 6.1 (e.g. withdrawal, death, end of observation). Hazard ratio estimates, accounting for correlated data within participants with multiple new diagnoses and accounting for dependence among participants with the same primary care provider,[31] are obtained using treatment status, participant sex, and disease categories as covariates. Treatment effect is characterized by treatment versus control arm and is presented as a hazard ratio with accompanying estimates for robust standard error, 95% confidence interval, and *P* value.

Time-to-diagnosis outcomes between treatment groups are visualized using standard Kaplan-Meier curves (e.g. survival, cumulative incidence) and accompanying risk tables.

To assess the robustness of the primary endpoint in the ITT population, the approach described will be carried out within the high genetic risk complete case population.

7.2.2 Average genetic risk group

As a prespecified exploratory outcome, we use a similar Cox modeling approach as described for the high genetic risk group to compare PRS-average and UC-average arms among those participants at average genetic risk for all diseases. Here, the outcome is time-to-diagnosis of any of the 6 common complex diseases. The ITT analysis is based on the rate of new diagnoses (either undiagnosed prevalent or incident cases) at month 24 after randomization. Diagnosis is characterized by a change in participant state from apparently non-diseased to diseased for any of the 5 sex-specific diseases (only men assessed for prostate cancer and women assessed for breast cancer) of interest as determined by expert clinical chart review during the 24 month observation period.

To assess robustness of the ITT analysis for the average genetic risk time-to-diagnosis endpoint, the analysis will be replicated within the average genetic risk complete case population.

7.2.3 Subgroup analyses

Further analyses of the primary endpoint examine time-to-diagnosis outcomes for specific diseases separately and across randomization stratum. The following pre-specified exploratory analyses will be conducted:

1. Treatment arm comparison for time-to-diagnosis of common complex diseases among high-risk individuals (PRS-high versus UC-high) who have a corresponding disease diagnosis and disease-specific high-risk genetic result (e.g. an individual who is diagnosed with T2D and has a high-risk result for T2D). .
2. Treatment arm comparisons for time-to-diagnosis between the PRS-high/UC-high and PRS-average/UC-average stratum in order to quantify disease risk elevation among patients with $OR_{PRS}>2.0$ compared to those with all $OR_{PRS}<2.0$.

7.3 <u>Secondary and other pre-specified outcomes</u>

7.3.1. Diagnostic testing

One secondary outcome is the occurrence of diagnostic testing for any of the 6 common complex diseases among randomized participants that occurs within their 24-month observation period. For participants in the high-risk group, diagnostic testing is characterized by the administration of a diagnostic procedure for at least one of the common complex diseases. For participants in the average-risk group, diagnostic testing is characterized by the administration of a diagnostic procedure for any of 5 common diseases relevant to participant sex.

We use a generalized linear model, accounting for provider clustering by generalized estimating equations (GEE), to compare the binary outcome of diagnostic testing (e.g. 'Occurred' (1)

versus 'Did not occur' (0)) between treatment arms within the ITT populations.[30,39-42] Initial models assume a binomial distribution and use a logit link function and exchangeable correlation structure to derive estimates. Odds ratios are derived using treatment arm, participant sex, and disease categories as covariates. Treatment effect is characterized by an odds ratio estimate presented with robust standard error, 95% confidence interval, and *P* value.

An exploratory analysis will examine time-to-diagnostic testing, using the time in days from randomization to the date of diagnostic testing.

7.3.2 Risk reducing medications

Relevant prescription medication changes, including addition or dose adjustment, for any of the 6 common complex diseases among randomized participants that occurs within their 24-month observation period are analyzed. For participants in the high-risk group, risk reducing medication changes are characterized by the prescription of a new medication or change in dose of an existing medication for purposes of risk reduction. For participants in the average-risk group, risk reducing medication change is characterized by the prescription of a new medication or change in dose of an existing medication for the purposes of risk reduction for any of the 5 common diseases of interest, relative to participant sex.

Similar to the diagnostic testing outcome, we use a generalized linear model, accounting for provider clustering by generalized estimating equations (GEE), to compare the binary outcome of a new medication prescription or medication change (e.g. 'Occurred' (1) versus 'Did not occur' (0)) between treatment arms within the ITT populations.[30,39-41] Initial models assume a binomial distribution and use a logit link function and exchangeable correlation structure to derive estimates. Odds ratios are derived using treatment arm, participant sex, and disease categories as the only covariates. Treatment effect is characterized by an odds ratio estimate presented with robust standard error, 95% confidence interval, and *P* value. We also use a generalized linear model, accounting for provider clustering by GEE, to compare counts of risk reducing medication changes (either new prescriptions or dose adjustment) between treatment arms within the ITT populations.[30,39-41] Initial models assume a Poisson distribution and use a log link function and exchangeable correlation structure to derive estimates. Incident rate ratios are derived using treatment arm, participant sex, and disease categories as covariates. Treatment effect is characterized by an incident rate ratio estimate presented with robust standard error, 95% confidence interval, and *P* value.

7.3.3 Continuous outcome measures

The following continuous measures are compared between treatment groups among the ITT populations using standard linear methods[40,43-46]:

- Self-reported health status and quality of life
- Self-reported patient activation
- Self-reported medication adherence

- Blood pressure
- Body-mass index (BMI)
- Low-density lipoprotein cholesterol (LDL-C)

Analysis of covariance (ANCOVA) is used to compare continuous follow-up measures, including participant baseline measures, treatment group assignment, participant sex, and disease categories as covariates. Treatment effect is characterized by treatment versus control arm, presented as mean follow-up estimates with accompanying standard errors, 95% confidence intervals and *P* values. In the presence of substantial missing data, linear mixed modeling or other repeated measures designs may be implemented.

7.3.4 Categorical and ordinal outcome measures

The following measures are compared between treatment groups among the ITT populations using standard methods for categorical data analysis[41,47,48]:

- Self-reported aspirin use
- Self-reported smoking status
- Self-reported physical activity
- Self-reported alcohol intake
- Self-reported processed meat consumption

Frequency of end-of-study self-report responses to study surveys, including categorical, ordinal, and Likert items, are reported by treatment group (n, %). Binary logistic regression is used to compare end-of-study dichotomous outcomes between treatment groups. To assess post-treatment ordered outcomes between treatment arms, we use ordinal logistic regression (e.g. cumulative logit model). Initial models include participant baseline response, sex, and treatment group assignment as the only covariates. Treatment effect is characterized by an odds ratio estimate presented with standard error, 95% confidence interval, and *P* value.

7.2.5 Healthcare costs

Statistical and methodological considerations for healthcare cost analysis for the GenoVA Study is described in a future revision to this Statistical Analysis Plan.

7.2.6 Provider knowledge and beliefs about PRS

Statistical and methodological approaches related to the development and conduct of semi-structured provider interviews is described in a future revision to this Statistical Analysis Plan.

7.4 Additional analyses

Further exploratory analyses using the methods described may be conducted for all study outcomes between treatment arms (e.g. UC-high versus UC-average) and across relevant

subgroups (e.g. disease risk, age, etc.). Data from the 1-2% of participants excluded from the RCT because of a medically actionable finding will be analyzed separately in exploratory analyses.

Inclusion of additional covariates (e.g. summary PRS scores, age, baseline health status, race, socioeconomic status) in the models described or use of alternative statistical methods may be implemented to enhance model precision, to adjust for differences in baseline factors or multilevel characteristics, or to improve the integrity of the analyses (e.g. in the event of substantial missing data), among other reasons. To assess robustness of ITT analyses, analyses may be replicated within the relevant complete case populations. The addition of covariates or use of alternative methods (e.g. survival rate ratios, restricted mean survival time)[49,50] to assess primary, secondary, and exploratory outcomes may be considered and are supplemental to the pre-specified analyses.

Additional exploratory analyses may be conducted to further examine study data or address research questions that arise during the conduct of the study.

Any exploratory analyses or use of alternative methods will be justified and described in detail if reported.

8.0 References

1. Gamble C, Krishan A, Stocken D, Lewis S, Juszczak E, Dore C, et al. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. *JAMA.* 2017;318(23):2337-2343 DOI: 10.1001/jama.2017.1855610.1001/jama.2017.18556.
2. Glynn RJ, Brookhart MA, Stedman M, Avorn J, Solomon DH. Design of cluster-randomized trials of quality improvement interventions aimed at medical care providers. *Med Care.* 2007;45(10 Supl 2):S38-43 DOI: 10.1097/MLR.0b013e318070c0a0 10.1097/MLR.0b013e318070c0a0.
3. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet.* 2015;47(10):1121-1130 DOI: 10.1038/ng.339610.1038/ng.3396. Epub 2015 Sep 7.
4. Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet.* 2018;50(9):1219-1224 DOI: 10.1038/s41588-018-0183-z10.1038/s41588-018-0183-z. Epub 2018 Aug 13.
5. Roselli C, Chaffin MD, Weng LC, Aeschbacher S, Ahlberg G, Albert CM, et al. Multi-ethnic genome-wide association study for atrial fibrillation. *Nat Genet.* 2018;50(9):1225-1233 DOI: 10.1038/s41588-018-0133-910.1038/s41588-018-0133-9.
6. Mahajan A, Taliun D, Thurner M, Robertson NR, Torres JM, Rayner NW, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat Genet.* 2018;50(11):1505-1513 DOI: 10.1038/s41588-018-0241-610.1038/s41588-018-0241-6. Epub 2018 Oct 8.
7. Schmit SL, Edlund CK, Schumacher FR, Gong J, Harrison TA, Huyghe JR, et al. Novel Common Genetic Susceptibility Loci for Colorectal Cancer. *J Natl Cancer Inst.* 2019;111(2):146-157 DOI: 10.1093/jnci/djy09910.1093/jnci/djy099.

8. Frampton MJ, Law P, Litchfield K, Morris EJ, Kerr D, Turnbull C, et al. Implications of polygenic risk for personalised colorectal cancer screening. *Ann Oncol.* 2016;27(3):429-434 DOI: 10.1093/annonc/mdv54010.1093/annonc/mdv540. Epub 2015 Nov 16.
9. Dunlop MG, Dobbins SE, Farrington SM, Jones AM, Palles C, Whiffin N, et al. Common variation near CDKN1A, POLD3 and SHROOM2 influences colorectal cancer risk. *Nat Genet.* 2012;44(7):770-776 DOI: 10.1038/ng.2293.
10. Michailidou K, Lindstrom S, Dennis J, Beesley J, Hui S, Kar S, et al. Association analysis identifies 65 new breast cancer risk loci. *Nature.* 2017;551(7678):92-94 DOI: 10.1038/nature2428410.1038/nature24284. Epub 2017 Oct 23.
11. Schumacher FR, Al Olama AA, Berndt SI, Benlloch S, Ahmed M, Saunders EJ, et al. Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. *Nat Genet.* 2018;50(7):928-936 DOI: 10.1038/s41588-018-0142-8.
12. Schumacher FR, Olama AAA, Berndt SI, Benlloch S, Ahmed M, Saunders EJ, et al. Author Correction: Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. *Nat Genet.* 2019;51(2):363 DOI: 10.1038/s41588-018-0330-6.
13. Szulkin R, Whitington T, Eklund M, Aly M, Eeles RA, Easton D, et al. Prediction of individual genetic risk to prostate cancer using a polygenic score. *Prostate.* 2015;75(13):1467-1474 DOI: 10.1002/pros.2303710.1002/pros.23037. Epub 2015 Jul 14.
14. Seibert TM, Fan CC, Wang Y, Zuber V, Karunamuni R, Parsons JK, et al. Polygenic hazard score to guide screening for aggressive prostate cancer: development and validation in large scale cohorts. *BMJ.* 2018;360:j5757 DOI: 10.1136/bmj.j5757.
15. Price LE, Shea K, Gephart S. The Veterans Affairs's Corporate Data Warehouse: Uses and Implications for Nursing Research and Practice. *Nurs Adm Q.* 2015;39(4):311-318 DOI: 10.1097/NAQ.000000000000011810.1097/NAQ.0000000000000118.
16. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ.* 2010;340:c869 DOI: 10.1136/bmj.c86910.1136/bmj.c869.
17. Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ.* 2008;337:a2390 DOI: 10.1136/bmj.a2390.
18. Jones D, Kazis L, Lee A, Rogers W, Skinner K, Cassar L, et al. Health Status Assessments Using the Veterans SF-12 and SF-36: Methods for Evaluating Outcomes in the Veterans Health Administration. *Journal of Ambulatory Care Management.* 2001;24(3):68-86 DOI: 10.1097/00004479-200107000-00011.
19. Kazis EL, Selim SA, Rogers RW, Ren RX, Lee RA, Miller RD. Dissemination of Methods and Results From the Veterans Health Study: Final Comments and Implications for Future Monitoring Strategies Within and Outside the Veterans Healthcare System. *Journal of Ambulatory Care Management.* 2006;29(4):310-319 DOI: 10.1097/00004479-200610000-00007.
20. Selim A, Rogers W, Fleishman J, Qian S, Fincke B, Rothendler J, et al. Updated U.S. population standard for the Veterans RAND 12-item Health Survey (VR-12). *Qual Life Res.* 2009;18(1):43-52 DOI: 10.1007/s11136-008-9418-2.
21. Hibbard JH, Mahoney ER, Stockard J, Tusler M. Development and testing of a short form of the patient activation measure. *Health Serv Res.* 2005;40(6 Pt 1):1918-1930 DOI: 10.1111/j.1475-6773.2005.00438.x10.1111/j.1475-6773.2005.00438.x.
22. Voils CI, Maciejewski ML, Hoyle RH, Reeve BB, Gallagher P, Bryson CL, et al. Initial validation of a self-report measure of the extent of and reasons for medication nonadherence. *Med Care.* 2012;50(12):1013-1019 DOI: 10.1097/MLR.0b013e318269e12110.1097/MLR.0b013e318269e121.

23. Centers for Disease Control and Prevention (CDC). Behavioral Risk Factor Surveillance System Survey Questionnaire. U.S Department of Health and Human Services: Atlanta, Georgia. 2019.
24. Lachin JM. Statistical considerations in the intent-to-treat principle. *Control Clin Trials.* 2000;21(3):167-189 DOI: 10.1016/s0197-2456(00)00046-510.1016/s0197-2456(00)00046-5.
25. Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. *Br J Cancer.* 1976;34(6):585-612 DOI: 10.1038/bjc.1976.22010.1038/bjc.1976.220.
26. McHugh R, Matts J. Post-stratification in the randomized clinical trial. *Biometrics.* 1983;39(1):217-225.
27. Spiro A, Rogers WH, Qian S, Kazis EL. Imputing Physical and Mental Summary Scores (PCS and MCS) for the Veterans SF-12 Health Survey in the Context of Missing Data. *Technical Report, scoring algorithms and users guide, submitted and approved by the Center for Medicare and Medicaid Services (CMS) and National Committee for Quality Assurance (NCQA).* 2004.
28. Lee KJ, Carlin JB. Multiple imputation for missing data: fully conditional specification versus multivariate normal imputation. *Am J Epidemiol.* 2010;171(5):624-632 DOI: 10.1093/aje/kwp42510.1093/aje/kwp425. Epub 2010 Jan 27.
29. Bartlett JW, Seaman SR, White IR, Carpenter JR, Alzheimer's Disease Neuroimaging I. Multiple imputation of covariates by fully conditional specification: Accommodating the substantive model. *Stat Methods Med Res.* 2015;24(4):462-487 DOI: 10.1177/096228021452134810.1177/0962280214521348. Epub 2014 Feb 12.
30. Stedman MR, Gagnon DR, Lew RA, Solomon DH, Brookhart MA. An evaluation of statistical approaches for analyzing physician-randomized quality improvement interventions. *Contemp Clin Trials.* 2008;29(5):687-695 DOI: 10.1016/j.cct.2008.04.00310.1016/j.cct.2008.04.003. Epub 2008 Apr 23.
31. Lee EW, Wei LJ, Amato DA. Cox-Type Regression Analysis for Large Numbers of Small Groups of Correlated Failure Time Observations. In: Klein JP, Goel PK, eds. *Survival Analysis: State of the Art.* Dordrecht, Netherlands: Kluwer Academic Publishers; 1992:237-247.
32. Kalish LA, Begg CB. Treatment allocation methods in clinical trials: a review. *Stat Med.* 1985;4(2):129-144 DOI: 10.1002/sim.4780040204.
33. Kahan BC, Morris TP. Reporting and analysis of trials using stratified randomisation in leading medical journals: review and reanalysis. *BMJ.* 2012;345:e5840 DOI: 10.1136/bmj.e5840.
34. Kahan BC, Morris TP. Improper analysis of trials randomised using stratified blocks or minimisation. *Stat Med.* 2012;31(4):328-340 DOI: 10.1002/sim.443110.1002/sim.4431. Epub 2011 Dec 4.
35. Lin DY, Wei LJ. The Robust Inference for the Cox Proportional Hazards Model. *Journal of the American Statistical Association.* 1989;84(408):1074-1078 DOI: 10.2307/2290085.
36. Klein JP, Goel PK, SpringerLink (Online service). Survival Analysis: State of the Art. In: *Nato Science, Series E: Applied Sciences,.* 1992: http://dx.doi.org/10.1007/978-94-015-7983-4.
37. Lin DY. Cox regression analysis of multivariate failure time data: the marginal approach. *Stat Med.* 1994;13(21):2233-2247 DOI:10.1002/sim.478013210510.1002/sim.4780132105.
38. Therneau TM, Grambsch PM. *Modeling survival data : extending the Cox model.* New York: New York : Springer; 2000.

39. Hanley JA, Negassa A, Edwardes MD, Forrester JE. Statistical analysis of correlated data using generalized estimating equations: an orientation. *Am J Epidemiol.* 2003;157(4):364-375 DOI: 10.1093/aje/kwf215.
40. Vittinghoff E. *Regression Methods in Biostatistics Linear, Logistic, Survival, and Repeated Measures Models.* 2nd ed. 2012.. ed. New York, NY: New York, NY : Springer New York : Imprint: Springer; 2012.
41. Agresti A. *Categorical data analysis.* 3rd ed.. ed. Hoboken, New Jersey: Hoboken, New Jersey : Wiley-Interscience; 2013.
42. Ristl R, Urach S, Rosenkranz G, Posch M. Methods for the analysis of multiple endpoints in small populations: A review. *Journal of Biopharmaceutical Statistics.* 2019;29(1):1-29 DOI: 10.1080/10543406.2018.1489402.
43. Vickers AJ, Altman DG. Statistics notes: Analysing controlled trials with baseline and follow up measurements. *BMJ.* 2001;323(7321):1123-1124 DOI: 10.1136/bmj.323.7321.112310.1136/bmj.323.7321.1123.
44. Vickers AJ. Parametric versus non-parametric statistics in the analysis of randomized trials with non-normally distributed data. *BMC Med Res Methodol.* 2005;5:35 DOI: 10.1186/1471-2288-5-3510.1186/1471-2288-5-35.
45. O'Connell NS, Dai L, Jiang Y, Speiser JL, Ward R, Wei W, et al. Methods for Analysis of Pre-Post Data in Clinical Research: A Comparison of Five Common Methods. *J Biom Biostat.* 2017;8(1):1-8 DOI: 10.4172/2155-6180.100033410.4172/2155-6180.1000334.
46. Clifton L, Clifton DA. The correlation between baseline score and post-intervention score, and its implications for statistical analysis. *Trials.* 2019;20(1):43 DOI: 10.1186/s13063-018-3108-310.1186/s13063-018-3108-3.
47. Harrell JFE. *Regression Modeling Strategies With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis.* 2nd ed. 2015.. ed. Cham: Cham : Springer International Publishing : Imprint: Springer; 2015.
48. Agresti A. *An introduction to categorical data analysis.* Third edition. ed. Hoboken, NJ: Hoboken, NJ : John Wiley & Sons, Inc.; 2019.
49. Uno H, Claggett B, Tian L, Inoue E, Gallo P, Miyata T, et al. Moving beyond the hazard ratio in quantifying the between-group difference in survival analysis. *J Clin Oncol.* 2014;32(22):2380-2385 DOI: 10.1200/JCO.2014.55.2208.
50. Pak K, Uno H, Kim DH, Tian L, Kane RC, Takeuchi M, et al. Interpretability of Cancer Clinical Trial Results Using Restricted Mean Survival Time as an Alternative to the Hazard Ratio. *JAMA Oncol.* 2017;3(12):1692-1696 DOI: 10.1001/jamaoncol.2017.2797.