The Genetics of Cardiovascular Disease: From Epidemiology to Clinical Practice

Brenton R Swenson

A dissertation

Submitted in partial fulfillment of the

Requirements for the degree of

Doctor of Philosophy

University of Washington

2021

Reading Committee:

Nona Sotoodehnia, Chair

Wylie Burke

Susan Heckbert

Program Authorized to Offer Degree:

Public Health Genetics

©Copyright 2021

Brenton R Swenson

University of Washington

Abstract

The Genetics of Cardiovascular Disease: From Epidemiology to Clinical Practice

Brenton R Swenson

Chair of Supervisory Committee:

Nona Sotoodehnia

Department of Epidemiology – Public Health

Chapter 1 contains a GWAS of QRS duration, a measure of ventricular depolarization and conduction, in 15,124 Hispanic/Latino individuals from four studies. We identified six loci associated with QRS ($P<5x10^{-8}$), including two novel loci: *MYOCD*, a nuclear protein expressed in the heart, and *SYT1*, an integral membrane protein. **Chapter 2** demonstrates the strong association of a genetic risk score (GRS) with ischemic stroke in European-ancestry participants with atrial fibrillation from the UK Biobank study (n=21,185). The GRS was independent of clinical risk factors. Including the GRS in clinical risk models may better identify individuals who could benefit from anticoagulation. **Chapter 3** uses the example of the clinical practice guidelines concerning competitive sports participation of young athletes with long QT syndrome to demonstrate the challenges in using genetic data to deliver effective precision medicine.

Preface

My dissertation contains three distinct chapters united by the theme of cardiovascular genetics. The chapters follow a rough progression from early-stage population health and epidemiology towards implementation in clinical practice. The first chapter uses population genetics to better understand physiology and pathophysiology by employing a genome-wide association study (GWAS) to understand the genetic underpinnings of cardiac ventricular conduction (QRS interval duration) in Hispanics/Latinos. The second chapter explores the potential of a genetic risk score (GRS) to improve stroke prediction in individuals with atrial fibrillation to identify those who may benefit from anticoagulation therapy for stroke prophylaxis. Finally, the last chapter explores the biological and institutional challenges faced when implementing scientific discoveries in genetics into clinical practice. The clinical practice guidelines surrounding the participation of young athletes with long QT syndrome in competitive sports is used as an example.

Chapter 1: GWAS of QRS Duration Identifies New Loci Specific to Hispanic/Latino Populations

Introduction

The duration of the QRS complex on a resting, standard 12-lead electrocardiogram (ECG) represents the electrical depolarization of the ventricles as an impulse travels through the cardiac conduction system and the ventricular myocardium. Delay in cardiac ventricular conduction results in increased QRS durations, and has been shown to predict heart failure prognosis,[1, 2] sudden death,[3] and cardiovascular (CV) mortality in patients with and without left ventricular dysfunction, independent of traditional CV risk factors.[4] In turn, shortening of the QRS duration with the use of cardiac-resynchronization therapy (CRT) has been shown to decrease heart-failure related events in patients with QRS prolongation.[5]

To date, heritability estimates of QRS duration have varied, with up to ~40% heritability found in more recent studies.[6-9] While previous genome-wide association studies have focused predominantly on European populations,[10-12] there have been several smaller studies of Asian,[13-15] Pacific Islander,[16] and African American populations.[17, 18] Collectively, these GWAS analyses have identified 32 SNPs across 26 loci associated with QRS duration. These loci harbor ion channel and transcription factor genes involved in cardiac conduction, including *SCN5A, SCN10A, TBX3, TBX5, TBX20,* and *HAND1.*[10-14, 16-18] To our knowledge, there has been no GWAS

performed to study the genetics of QRS duration in Hispanic/Latino ancestry populations. We therefore performed a GWAS of QRS duration in four Hispanic/Latino study populations: the Hispanic Community Health Study/Study of Latinos (HCHS/SOL), the Multi-Ethnic Study of Atherosclerosis (MESA), the Starr County Study (Starr), and the Women's Health Initiative (WHI).

Results

Our GWAS included 15,124 Hispanic/Latino individuals from four contributing cohorts. Baseline characteristics varied substantially across the four cohorts. For example, WHI is a study of women only. The average age of the participants across the four cohorts ranged from 45 to 61 years. The prevalence of hypertension and diabetes across the four cohorts ranged from 26% to 42% for hypertension and 8% to 46% for diabetes (S1 Table).

Genome-wide Association Analysis

Following quality control (see Materials and Methods), individual studies contributed between 5.8 M and 20.2 M individual SNPs, yielding 21.1 M combined, unique SNPs overall. Neither the individual studies (λ range=0.96-1.03) nor the combined meta-analysis (λ=1.03) exhibited evidence of test-statistic inflation (S1-S2 Figs). SNPs in six loci exceeded the genome-wide threshold for significance (Table 1, Fig 1). Two of the loci (*MYOCD* and *SYT1*) were novel, whereas the remaining four (*SCN5A-SCN10A, HAND1, CDKN1A,* and *VTI1A*) were previously identified in other ethnic groups. There was no evidence of heterogeneity (Cochran's Q test *P*-value>0.05, S2 Table), and effect

direction was consistent across all contributing studies for all index SNPs (S3 Fig). All index SNPs were either directly genotyped or were imputed with high quality (S3 Table).

Table 1. Genome-wide significant SNPs identified in a GWAS meta-analysis of n=15,124 participants of Hispanic/Latino ancestry from four studies.

Locus	Nearest Gene[a][19]	Index SNP	Chr[b]	A1/A2[c]	Function	CAF[d]	β (SE)[e]	*P*	Multi *P*[f]
1	*SCN5A*	rs3922844	3	C/T	Intronic	0.63	1.03 (0.10)	1.19e-24	3.36e-15
1	**SCN5A*	rs62241190	3	G/A	Intronic	0.04	2.46 (0.27)	5.82e-20	1.09e-12
1	**SCN10A*	rs10428132	3	T/G	Intronic	0.38	0.79 (0.10)	1.43e-15	3.57e-11
1	**SCN5A*	rs9856387	3	C/T	Intronic	0.73	0.76 (0.11)	2.12e-12	1.68e-08
2	*HAND1*	rs13165478	5	G/A	Intergenic	0.67	0.68 (0.10)	2.69e-11	--
3	*CDKN1A*	rs3176326	6	A/G	Intronic	0.17	1.15 (0.13)	1.54e-19	4.26e-13
3	***SPRK1*	rs2395642	6	T/C	Intronic	0.16	0.93 (0.13)	4.71e-09	3.94e-06
4	*VTI1A*	rs7906312	10	A/C	Intronic	0.19	0.77 (0.13)	8.14e-10	--
5	*SYT1*	rs4842438	12	C/A	Intronic	0.93	1.04 (0.19)	4.24e-08	--
6	*MYOCD*	rs16946539	17	T/C	Intronic	0.06	1.28 (0.22)	1.74e-09	--

[a]* Denotes a secondary signal. ** Denotes a secondary SNP no longer genome-wide significant after multi-SNP testing at the locus. For the intergenic SNP (rs13165478), the nearest gene was determined by the nearest protein coding gene in base pairs from the National Center for Biotechnology Information RefSeq database.
[b]Chr: Chromosome
[c]A1/A2: Coded/non-coded alleles
[d]CAF: Coded allele frequency
[e]β: Effect estimate measured in milliseconds.
[f]*P* for the association of index SNP with QRS duration, upon fixed effects meta-analysis using a model which includes all index SNPs in the same locus.

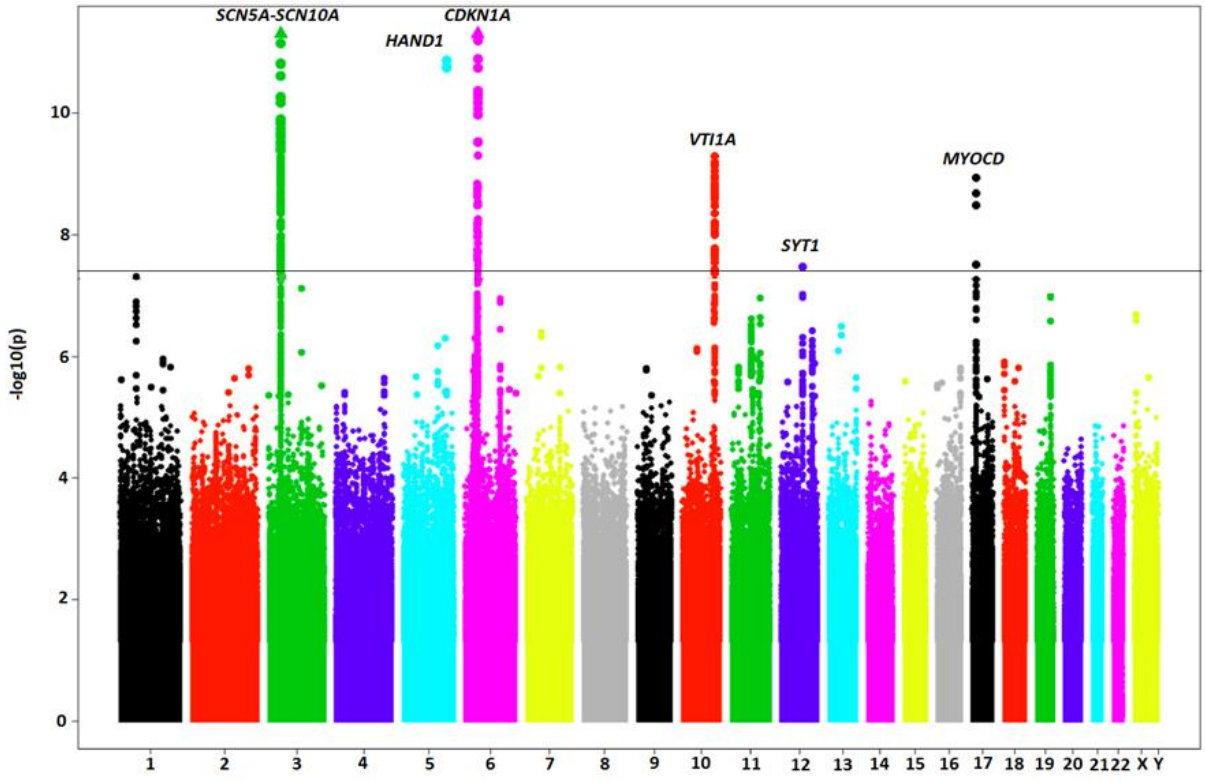


Fig 1. Manhattan plot of SNP-QRS associations. Manhattan plot showing the association of SNPs with QRS duration in the GWAS meta-analysis containing 15,124 individuals of Hispanic/Latino ethnicity. The horizontal line represents the genome-wide significance threshold of (*P*=5E-08). SNPs mapping to 6 loci exceeded the GWAS threshold for significance.

Novel Associations

The meta-analysis identified two novel loci (*MYOCD* and *SYT1)* associated with QRS duration. The index SNP (rs16946539) in *MYOCD (*myocardin, a nuclear protein found in cardiac and smooth muscle), as well as the only SNP (rs139859815) in high LD

(r^2>0.5) with it, are monomorphic in the European-descent and African-descent 1000 Genomes super populations (S4 Table).

The second novel locus was on chromosome 12 near *SYT1 (*synaptotagmin-1), an integral membrane protein of synaptic vesicles that responds to calcium signaling. The index SNP in *SYT1* (rs4842438) was examined in both European and African ancestry GWAS efforts, but failed to reach nominal significance (*P*>0.05; Fig 2). Indeed, the effect size of rs4842438 in Hispanics/Latinos (beta=1.04 ms) is statistically larger than in European-descent (beta=0.13 ms) and African-descent (beta=0.00 ms) individuals (*P* for difference=5.69×10^{-5} and 8.65×10^{-6}, respectively, Fig 2, S5-S7 Tables). Moreover, the broad LD pattern seen in Hispanics/Latinos is entirely absent in Europeans and African Americans (S4E Fig). The lack of associations among European and African descent individuals is not explained by lack of power from a smaller sample size or lower MAF (European MAF=0.06; African MAF=0.20; Hispanic/Latino MAF=0.07).

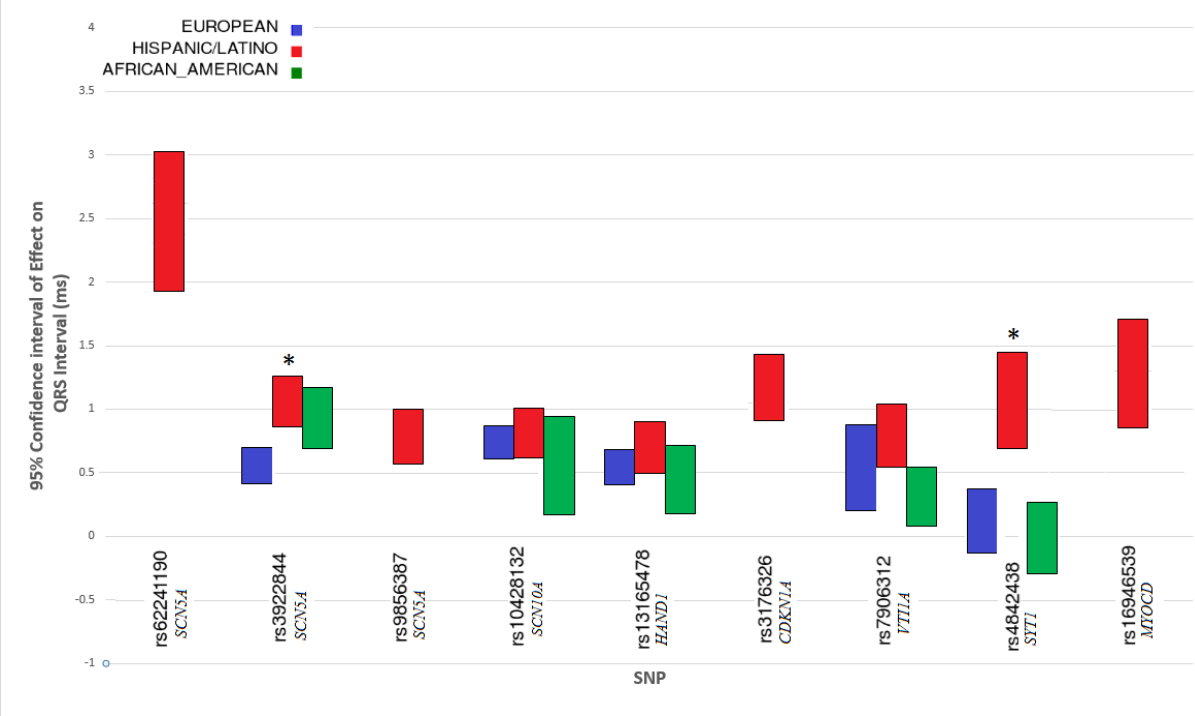


Fig 2. Transethnic comparison of QRS effect sizes for Hispanic/Latino SNPs.

Comparison of effect sizes and 95% confidence intervals for the 9 index SNPs that were genome-wide significant in the Hispanic/Latino QRS duration GWAS, and the effect sizes for corresponding SNPs in the European and African American GWAS. SNP rs10428132 was not directly measured in the European or African American studies, but instead a SNP in perfect LD (r^2=1) was used (rs6800541). Other SNPs that were not directly measured in Europeans or African Americans are not presented (rs62241190, rs9856387, rs3176326, and rs16946539). * Refers to SNPs where the difference in effect size between two ethnic groups was significant at the Bonferroni corrected *P*-value. Two SNPs showed larger effects in Hispanics/Latinos than in European-descent individuals: rs3922844 in *SCN5A* and rs4842438 in *SYT1.* See S7 Table for additional details.

SCN5A-SCN10A

The most significant association with QRS duration was found in chromosome region 3p22 (rs3922844) at locus 1, bridging *SCN5A* and *SCN10A*, two adjacent cardiac sodium-channel genes (Fig 3). This SNP had previously been found to be genome-wide significant among European and African American descent individuals (S5-S6 Tables). Similar to our findings in Hispanics/Latinos, rs3922844 is the most significantly associated QRS SNP in African Americans (Fig 3). In contrast, among European-descent individuals, the strongest SNP association (rs1601957) resides within an intron of the *SCN10A* gene. The effect size of rs3922844 in Hispanics/Latinos is statistically larger than in Europeans (1.03 ms vs 0.56 ms decrease in QRS duration, respectively, *P* for difference=5.8×10^{-5}), and is closer to the effect size seen among African Americans (0.94 ms, Fig 3 and S5-S7 Tables). Conditional analyses at the *SCN5A-SCN10A* locus among Hispanics/Latinos revealed three additional independent genome-wide significant secondary signals (Table 1). Two of the secondary index SNPs (rs62241190 and rs9856387) were not tested in either the European or African American GWAS, because those analyses were based on HapMap rather than 1000 Genomes imputation. There are no SNPs in these two populations that are in high LD ($r^2>0.75$) with rs62241190 and rs9856387. Therefore, whether these SNPs are also significantly associated with QRS duration among non-Hispanic/Latino groups is unknown.

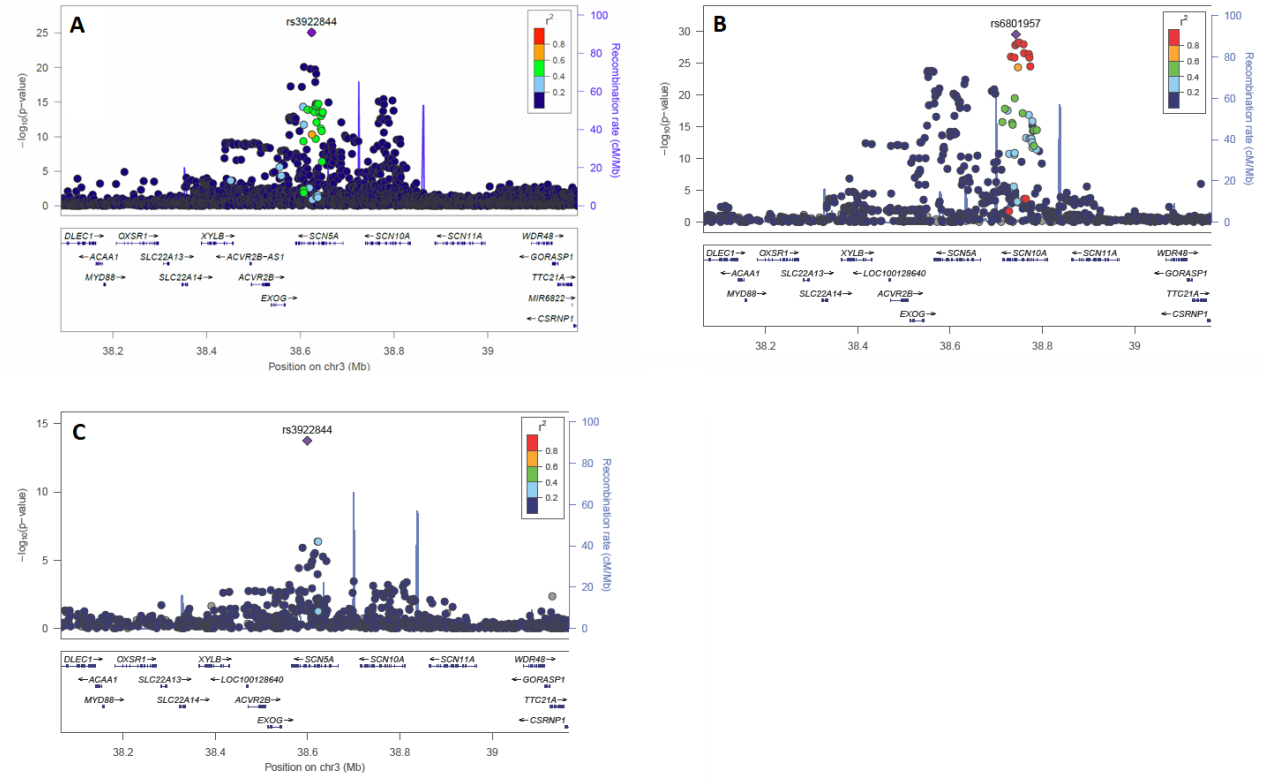


Fig 3. Regional association plots of *SCN5A-SCN10A*. A (top-left) – Hispanic/Latino GWAS. B (top right) – European GWAS. C (bottom left) – African American GWAS. Plots created with LocusZoom software.[20] The most significant SNP identified in the Hispanic Latino and African American GWAS was rs3922844 in *SCN5A*. The most significant SNP in the European GWAS was rs6801957 in *SCN10A*.

Additional Hispanic/Latino Associations in Known Loci

Three additional previously discovered SNP-QRS associations were also found in Hispanics/Latinos. These include intronic SNPs within *CDKN1A* and *VTI1A*, and an intergenic SNP near *HAND1.* While the Hispanic/Latino index SNP in *CDKN1A*

(rs3176326) was not directly evaluated in the European or African American GWAS, it was in high LD with a SNP that had been found to be highly significant in the previous European ancestry GWAS (r^2=0.71 with rs9462210). Interestingly, a conditional analysis of the *CDKN1A* locus revealed a suggestive secondary signal located in an intron of *SPRK1* approximately 1 Mb upstream from the primary *CDKN1A* signal (rs2395642, *P*=4.58x10^{-6} in conditional analyses; Table 1 and S4C Fig). The index SNP in *VTI1A* (rs7906312) is a novel SNP within a known locus, but it is not in high LD with the previously known SNP associated with QRS duration in Europeans (S5 Table and S4D Fig). Findings for *HAND1* (rs13165478) show that five SNPs in the region (all in very high LD with each other) were significantly associated with QRS duration in this 1000 Genomes imputation analysis. No other SNPs had near significant associations. This same haplotype was also identified as significant among European descent individuals (S4B Fig).

Transethnic Analyses

Generalization of previously known SNPs to Hispanics/Latinos

We examined 32 index SNPs from published GWAS analyses (27 European,1 African-American, 2 East Asian, and 2 from a meta-analysis of the European and African American GWAS results) for association with QRS duration among Hispanics/Latinos[11, 13, 18] (S8 Table, Fig 4). Of the 32 previously identified independent SNPs, 27 generalized in Hispanics/Latinos (r-value <0.05). These included 26 of the 27 SNPs from the European GWAS, with rs1362212 as the exception (S8 Table).

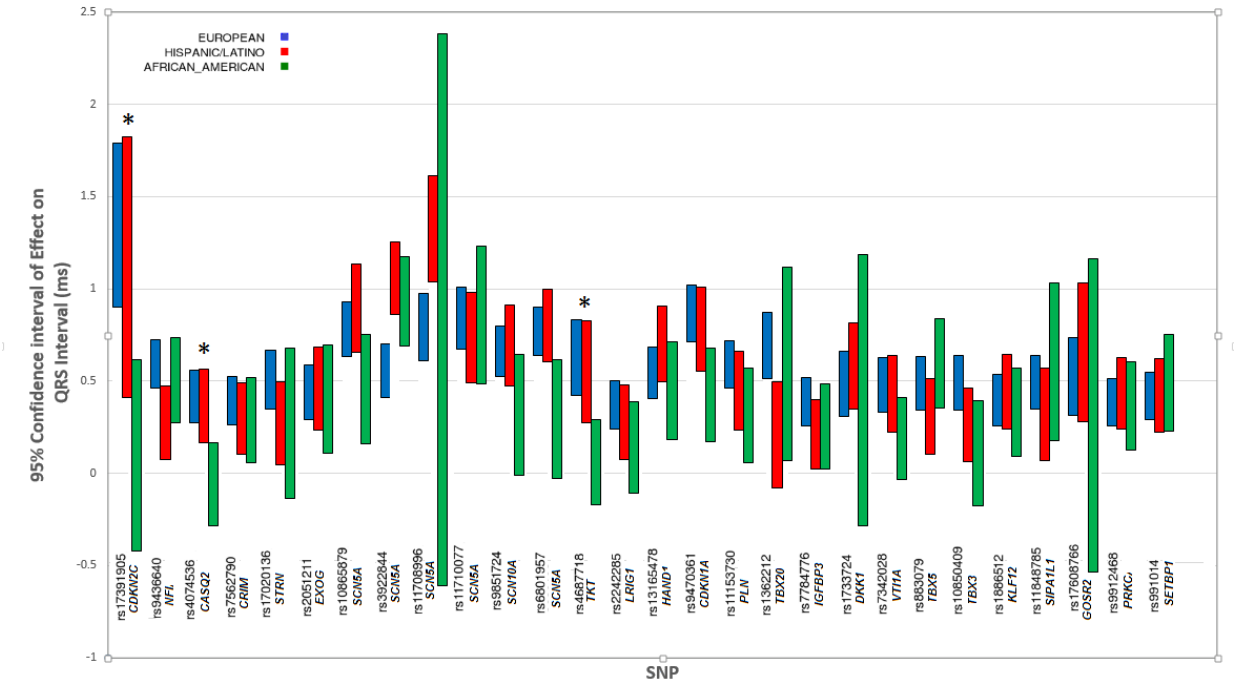


Fig 4. Transethnic comparison of QRS effect sizes for previously known SNPs. Comparison of effect sizes and 95% confidence intervals for 28 previously discovered SNPs for QRS duration across European, Hispanic/Latino, and African American GWAS results. Findings are largely similar across ethnic groups. * Refers to SNPs where the difference in effect size between two ethnic groups was significant at the Bonferroni corrected *P*-value. See S7 Table for additional details.

Effect Size Comparisons

We compared the effect sizes of the independent index SNPs identified in Hispanic/Latino, European descent, and African descent populations (using the current analysis for Hispanic/Latinos, and published meta-analyses for European and African descent individuals; Figs 2 and 4). Not all Hispanic/Latino SNPs were available in European and African-descent individuals, due to differences in imputation sets

(HapMap imputation in those of European and African descent versus 1000 genomes imputation in the current analysis). Hence, there are missing European and African-descent data for 4 SNPs, because 3 had no proxy SNP in high LD (r^2>0.9) with the index Hispanic/Latino SNP, and one was monomorphic in European and African-descent populations. After Bonferroni correction, 4 of the 29 independent SNPs showed evidence of significant differences in genotype-phenotype effect sizes (Fig 2 and 4, S7 Table).

Cross Phenotype Analyses

The 9 index SNPs identified in the Hispanic/Latino QRS duration GWAS were also examined for their association with other ECG phenotypes: QT duration,[21] PR duration,[22] heart rate,[23] and SDNN[23] (the standard deviation of normal to normal R-R intervals, a measure of heart rate variability, Table 2, S9 Table). The 4 SNPs in the *SCN5A-SCN10A* locus were genome-wide significantly associated with both QT and PR interval duration. SNPs that prolong the QRS interval also prolong the PR interval, but conversely shorten the QT interval. This pattern was also observed in the *CDKN1A* locus. The index SNPs met the Bonferroni corrected significance level for QT duration, but fell just short of significance for PR duration (*P*=1.44E-03). Intriguingly, the novel *MYOCD* locus was significantly associated with PR duration at the Bonferroni corrected significance level, with the SNP that prolongs QRS duration but shortens PR duration. None of the QRS index SNPs were significantly associated with heart rate or SDNN.

Table 2. **Pleiotropic Analyses.** Comparison of the effect size and significance level of QRS prolonging index SNPs with QT and PR duration in Hispanics/Latinos. There was no association of QRS SNPs with heart rate or heart rate variability (SDNN). Only significant results (*P*<0.05) are shown.

Nearest Gene[a][19]	SNP	QRS β (ms)	QRS *P* [b]	PR β (ms)	PR *P* [b]	QT β (ms)	QT *P* [b]
SCN5A	rs62241190	2.46	5.82E-20	4.72	2.90E-12	-3.58	1.83E-09
SCN5A	rs3922844	1.03	1.19E-24	3.39	3.57E-11	-1.77	9.52E-16
SCN5A	rs9856387	0.76	2.12E-12	1.88	8.90E-12	-1.42	3.43E-09
SCN10A	rs10428132	0.79	1.43E-15	3.81	4.89E-53	-1.29	3.04E-09
HAND1	rs13165478	0.68	2.69E-11	-		-	
CDKN1A	rs3176326	1.15	1.54E-19	1.02	1.44E-03	-1.1	8.59E-05
VTI1A	rs7906312	0.77	8.14E-10	-		-	
SYT1	rs4842438	1.04	4.24E-08	-		-	
MYOCD	rs16946539	1.28	1.74E-09	-2.28	1.82E-05	-	

[a]For the intergenic SNP (rs13165478), the nearest gene was determined by the nearest protein coding gene in base pairs from the National Center for Biotechnology Information RefSeq database.
[b]Bonferroni corrected significance: *P*<1.39E-03 for 36 tests

Functional Annotation

The function of all 9 index SNPs was investigated using the HaploReg 4.1 web server.[24] Functional information was obtained for 3 loci: *SCN5A-SCN10A, CDKN1A,* and *MYOCD.* SNPs in *CDKN1A* showed evidence for activating transcription in heart tissues, including fetal heart tissue, the right atrium, the right ventricle, and the left ventricle. *SCN5A-SCN10A* and *MYOCD* SNPs were identified as possible enhancers of transcription in the same heart tissues (S10 Table).

Discussion

Our GWAS meta-analysis of four cohorts (15,124 individuals) of Hispanic/Latino ethnicity found 9 index SNPs across 6 loci with genome-wide significant associations with QRS duration. Two loci were novel (*MYOCD* and *SYT1),* and four loci (*SCN5A-SCN10A, CDKN1A, HAND1*, and *VTI1A*) were previously identified in QRS GWAS analyses of European[11] and African-descent[18] individuals. This is the first GWAS of QRS duration in Hispanics/Latinos, a genetically admixed group, composed of European, African, and Native American ancestry populations, coming from what is now Mexico, Central America, South America, and the Caribbean islands. Hispanics/Latinos were excluded in prior published QRS duration GWAS, and the genetic as well as non-genetic determinants of cardiovascular risk remain under-examined in this population.[25, 26]

We identified two novel loci associated with QRS duration among Hispanic descent individuals: *MYOCD* and *SYT1*. *MYOCD* encodes a nuclear protein (myocardin) expressed in cardiomyocytes and smooth muscle cell-containing tissues. *MYOCD* has been shown to be essential for maintaining adult heart function.[27] Mice in which *MYOCD* is postnatally knocked down develop dilated cardiomyopathy and fatal heart failure.[28] A genetic study of Dominican families found evidence that *MYOCD* was associated with left atrial size.[29] Recent GWAS studies have found SNPs in *MYOCD* associated with PR interval duration[30] and atrial fibrillation[31] in European populations. The index SNP associated with QRS duration in *MYOCD* (rs16946539) is more common among Hispanics/Latinos (defined by the Ad Mixed American 1000 Genomes population, MAF=0.04) than among any other 1000 Genomes super

population, and is monomorphic among European- and African-descent individuals (European MAF=0.0; Asian MAF=0.02; African MAF=0.0). Therefore, a QRS GWAS in Hispanics/Latinos is uniquely advantageous in uncovering this genotype-phenotype association. The finding illustrates one of the main imperatives for conducting genetic studies in diverse and under-examined populations, both for ECG traits, and for other traits. Because underlying genetic architecture can differ across racial and ethnic populations, examining new populations may uncover novel genotype-phenotype associations.

The second novel locus found in this study involves an intronic SNP in *SYT1*. Interestingly, a different synaptotagmin gene, *SYT10*, was previously associated with heart rate.[32] The *SYT1* intronic SNP was not associated with heart rate in our study. It is noteworthy that this *SYT1* SNP showed no evidence of association with QRS among European and African-descent individuals. Further studies are warranted to validate this novel association in additional Hispanic/Latinos populations.

The most significant association signal was found in and surrounding two voltage-gated sodium channel genes: *SCN5A* and *SCN10A*. Whereas *SCN5A* is the canonical cardiac sodium channel responsible for cellular depolarization and enables conduction of the electrical signal, *SCN10A* appears to be particularly enriched in the specialized *His Purkinje* conduction fibers.[11] Transethnic analyses intriguingly show that the regional association plot for QRS in Hispanics/Latinos more closely resembles African

Americans plots than European-descent plots, with rs3922844 being the most significant SNP in the region among Hispanics/Latinos and African Americans.

The generalization study of the 28 independent index SNPs from QRS GWAS analyses in Hispanic/Latino, European, and African ancestry individuals shows remarkable consistency in the magnitude and direction of the effects overall. Although the genetic architecture of QRS duration in the three ethnic groups is largely comparable, differences are also present, such as the larger effect size at rs3922844 among Hispanics/Latinos than among European descent individuals.

Several limitations deserve consideration. First, our GWAS represents the largest performed in Hispanics/Latinos, but our sample size was nonetheless small. Additional Hispanic/Latino cohort studies are needed to extend these findings. However, novel associations were identified, and known associations were confirmed in a new ethnic population, despite the relatively small sample size. Larger sample sizes may reveal more genome-wide significant loci. For example, a SNP in the previously identified *NFIA* locus fell just short of genome-wide significance in Hispanics/Latinos (P=5.1x10^{-8}). Furthermore, while we excluded individuals with QRS durations longer than 120 ms in order to exclude individuals who have conduction defects and/or bundle branch blocks due to acquired heart disease, interesting genetic associations may be missed by this approach.

In conclusion, our findings indicate that the genetics of QRS duration are largely similar among ethnic groups. However, important differences do exist, illustrated by the novel genome-wide significant SNP in *MYOCD* that is monomorphic in both Europeans and African Americans. Our study underscores the importance of conducting genetic studies in diverse and under-examined populations, such as Hispanics/Latinos, to uncover novel loci.

Materials and Methods

Study Populations

Primary meta-analysis of individuals with Hispanic/Latino ancestry

Our meta-analysis included 15,124 participants of self-identified Hispanic/Latino descent from the following four studies: the HCHS/SOL (n=11,566), the Multi-Ethnic Study of Atherosclerosis (MESA, n=1431), the Starr County Study (n=582), and the WHI (n=1545) (see Supplementary Material for cohort descriptions and S1 Table for baseline characteristics by cohort). Ancestry was confirmed through principal components analysis, and a small number of genetic outliers (individuals determined to be of primarily Asian ancestry) were excluded. All participants consented to the use of their genetic information for health-related research purposes.

Comparison with meta-analyses of individuals with European and African ancestry

Comparisons of results were made between the Hispanic/Latino meta-analysis, and two published meta-analyses of QRS duration in individuals of European (n=40,407) and African American (n=13,301) ancestries. Details of these studies can be obtained from their original publications.[11, 18]

Electrocardiography

Participants in each of the four cohorts underwent a standard 12-lead ECG by a certified technician (see S11 Table). Participants were excluded from further analysis if they had any of the following: poor quality ECGs, atrial flutter or fibrillation, a ventricular paced rhythm, QRS duration ≥120 ms, Wolff-Parkinson-White on ECG, a history of previous myocardial infarction or heart failure, or were taking class I or class III antiarrhythmic medications.

Genotyping and imputation

HCHS/SOL participants were genotyped on an Illumina custom array that included the Illumina Omni 2.5M array (HumanOmmni2.5-8v1-1) and an additional ~150,000 SNPs. The additional SNPs were chosen to contain markers relevant to Hispanic/Latino ancestry, markers informative of Native American ethnicity, and significant loci from previous association studies.[26] MESA, WHI, and Starr County participants were genotyped using the Affymetrix Genome-Wide Human SNP Array 6.0. SNP genotyping inclusion criteria varied slightly across studies. After individual cohort genotype QC, imputation based on the 1000 Genomes phase 1 reference panel[33] was performed resulting in roughly 38 million SNPs (S11 Table).

Statistical Analysis

Genome Wide Association

To assess the association between genotype and QRS duration, individual cohort studies used additive genetic linear regression models, either in a regression model (MESA, WHI, Starr County) or a mixed model (HCHS/SOL, to account for relatedness and shared environment between individuals). The two methods estimate the same effect. Models were adjusted for age, sex, heart rate, systolic blood pressure, BMI, height, study site/region, and principal components of genetic ancestry. After we received the results from the studies, we applied an individual cohort QC filter, which excluded SNPs with low imputation quality (<0.30) or small effective sample sizes for each individual SNP (*effN* <30), with *effN* = 2 x *MAF* x (*MAF*-1) x *N* x *Imputation Quality;* where *N* is the number of participants. Each cohort contributed between ~6M and ~20M imputed SNPs, after applying this filter.

Meta-analysis

Results were combined using fixed-effects inverse variance meta-analysis using the METAL software package,[34] using genomic control for summary statistics to reduce test-statistic inflation. Study heterogeneity was evaluated using the Cochran Q test. Approximately 21M unique SNPs were contained in the meta-analysis. Results were considered genome-wide significant for *P*-values < $5x10^{-8}$. Secondary signals were identified using iterative rounds of conditional analysis, with adjustment for additional

Hispanic/Latino index SNPs in the model, until there were no SNPs found to be genome-wide significant.

Transethnic Generalization Analysis

Previous meta-analyses in European, East Asian, African American ancestry populations identified a total of 32 genome-wide significant SNPs. To assess whether these significant findings generalize to Hispanics/Latinos, we used the method of Sofer *et al.*[35] An association is considered generalized if a significant effect in the same direction exists in both the non-Hispanic/Latino discovery population as well as the Hispanic/Latino population. This method controls the false discovery rate (FDR) of the generalization null hypotheses, and generates an r-value for each SNP (with r-values < 0.05 showing evidence that an association is generalizable under FDR control at the α=0.05 level.)

Transethnic Effect Size Analysis

Comparisons of effect size differences of SNPs on QRS duration across the Hispanic/Latino, European, and African American GWAS were done using a procedure analogous to Welch's *t*-tests for each of the 33 independent SNPs that were identified as having a primary or secondary independent association in any of those studies. For the purposes of Bonferroni correction, there were 29 independent SNPs available for testing in all 3 cohorts. Some SNPs were not available in all cohorts due to use of different imputation panels. Comparisons for 29 SNPs among 3 studies resulted in 87

tests. Therefore, differences in effect size were determined to be significant when *P*<5.75E-04.

Cross Phenotype Analysis

Index SNPs discovered in the QRS duration analysis were also examined for associations with other ECG phenotypes, including QT, PR, heart rate, and SDNN. These GWAS efforts were based upon the same underlying Hispanic/Latino cohorts as the QRS duration GWAS. However, due to different inclusion/exclusion criteria, differences in the study samples do exist between studies. Full details of these studies can be obtained from their original publications.[21-23] Significance for these other traits exceeded either: a genome-wide significance threshold (*P<5.0E-08)*; Bonferroni corrected significance (*P*<1.39E-03 for 36 tests); or nominal significance (*P<0.05).*

Functional Annotation

The HaploReg v4.1 online web resource was used to functionally annotate genome-wide significant SNPs.[24] HaploReg utilizes data obtained from the ENCODE[36] and RoadMap projects,[37] to give information on how SNPs might alter gene expression in diverse tissue types. We restricted our analysis to only specific heart tissues -- namely, fetal heart tissue, the right atrium, the right ventricle, and the left ventricle. Based on the chromatin-15 state model, we summarized the potential function of SNPs in the genome-wide significant loci in each of the different heart tissues. For each locus, we examined primary SNPs, secondary signal SNPs, and all other SNPs in high LD

(r^2>0.80) with these SNPs. The LD structure pattern used for this analysis was the 1000 Genomes AMR Phase-1 super-population.

Acknowledgements

We acknowledge the work of the CHARGE QRS GWAS Consortium and the CARe-COGENT African-American QRS Consortium for their work on the respective European and African American GWAS studies referenced here. Full membership for these two groups is provided in S1 Appendix.

Hispanic Community Health Study/Study of Latinos (HCHS/SOL): We thank the participants and staff of the HCHS/SOL study for their contributions to this study.

Multi-Ethnic Study of Atherosclerosis (MESA): We also thank the other investigators, the staff, and the participants of MESA for their valuable contributions. A full list of participating MESA investigators and institutions can be found at http://www.mesa-nhlbi.org.

Starr County Study: We thank the field staff in Starr County for collection of these data and are grateful to the study participants who gave their time and contributed to the study. Genotyping services were provided by the Center for Inherited Disease Research (CIDR).

Women's Health Initiative: We thank the study participants and staff for contributions to the study. All contributors to WHI science are listed at https://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Long%20List.pdf.

This chapter was published previously under a creative commons license:
Swenson BR, Louie T, Lin HJ, Méndez-Giráldez R, Below JE, Laurie CC, et al. GWAS of QRS duration identifies new loci specific to Hispanic/Latino populations. PLoS One. 2019;14(6):e0217796.

References

1. Sandhu R, Bahler RC. Prevalence of QRS prolongation in a community hospital cohort of patients with heart failure and its relation to left ventricular systolic dysfunction. The American Journal of Cardiology. 2004;93(2):244-6.

2. Iuliano S, Fisher SG, Karasik PE, Fletcher RD, Singh SN, Failure DoVASToATiCH. QRS duration and mortality in patients with congestive heart failure. American Heart Journal. 2002;143(6):1085-91.

3. Gatzoulis MA, Till JA, Somerville J, Redington AN. Mechanoelectrical interaction in tetralogy of Fallot QRS prolongation relates to right ventricular size and predicts malignant ventricular arrhythmias and sudden death. Circulation. 1995;92(2):231-7.

4. Desai AD, Yaw TS, Yamazaki T, Kaykha A, Chun S, Froelicher VF. Prognostic significance of quantitative QRS duration. The American Journal of Medicine. 2006;119(7):600-6.

5. Moss AJ, Hall WJ, Cannom DS, Klein H, Brown MW, Daubert JP, et al. Cardiac-resynchronization therapy for the prevention of heart-failure events. New England Journal of Medicine. 2009;361(14):1329-38.

6. Russell MW, Law I, Sholinsky P, Fabsitz RR. Heritability of ECG measurements in adult male twins. Journal of Electrocardiology. 1998;30:64-8.

7. Havlik RJ, Garrison RJ, Fabsitz R, Feinleib M. Variability of heart rate, PR, QRS and QT durations in twins. Journal of Electrocardiology. 1980;13(1):45-8.

8. Li J, Huo Y, Zhang Y, Fang Z, Yang J, Zang T, et al. Familial aggregation and heritability of electrocardiographic intervals and heart rate in a rural Chinese population. Annals of Noninvasive Electrocardiology. 2009;14(2):147-52.

9. Mutikainen S, Ortega-Alonso A, Alén M, Kaprio J, Karjalainen J, Rantanen T, et al. Genetic influences on resting electrocardiographic variables in older women: a twin study. Annals of Noninvasive Electrocardiology. 2009;14(1):57-64.

10. Ritchie MD, Denny JC, Zuvich RL, Crawford DC, Schildcrout JS, Bastarache L, et al. Genome-and phenome-wide analysis of cardiac conduction identifies markers of arrhythmia risk. Circulation. 2013;127:1377-85.

11. Sotoodehnia N, Isaacs A, de Bakker PI, Dörr M, Newton-Cheh C, Nolte IM, et al. Common variants in 22 loci are associated with QRS duration and cardiac ventricular conduction. Nature Genetics. 2010;42(12):1068-76.

12. Holm H, Gudbjartsson DF, Arnar DO, Thorleifsson G, Thorgeirsson G, Stefansdottir H, et al. Several common variants modulate heart rate, PR interval and QRS duration. Nature Genetics. 2010;42(2):117-22.

13. Hong K-W, Lim JE, Kim JW, Tabara Y, Ueshima H, Miki T, et al. Identification of three novel genetic variations associated with electrocardiographic traits (QRS duration and PR interval) in East Asians. Human Molecular Genetics. 2014;23(24):6659-67.

14. Chambers JC, Zhao J, Terracciano CM, Bezzina CR, Zhang W, Kaba R, et al. Genetic variation in SCN10A influences cardiac conduction. Nature Genetics. 2010;42(2):149-52.

15. Sano M, Kamitsuji S, Kamatani N, Hong K-W, Han B-G, Kim Y, et al. Genome-wide association study of electrocardiographic parameters identifies a new association for PR interval and confirms previously reported associations. Human molecular genetics. 2014;23(24):6668-76.

16. Smith JG, Lowe JK, Kovvali S, Maller JB, Salit J, Daly MJ, et al. Genome-wide association study of electrocardiographic conduction measures in an isolated founder population: Kosrae. Heart Rhythm. 2009;6(5):634-41.

17. Jeff JM, Ritchie MD, Denny JC, Kho AN, Ramirez AH, Crosslin D, et al. Generalization of Variants Identified by Genome-Wide Association Studies for Electrocardiographic Traits in African Americans. Annals of Human Genetics. 2013;77(4):321-32.

18. Evans DS, Avery CL, Nalls MA, Li G, Barnard J, Smith EN, et al. Fine-mapping, novel loci identification, and SNP association transferability in a genome-wide association study of QRS duration in African Americans. Human Molecular Genetics. 2016:;ddw284.

19. Pruitt KD, Tatusova T, Maglott DR. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic acids research. 2006;35(suppl_1):D61-D5.

20. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics. 2010;26(18):2336-7.

21. Méndez-Giráldez R, Gogarten SM, Below JE, Yao J, Seyerle AA, Highland HM, et al. GWAS of the electrocardiographic QT interval in Hispanics/Latinos generalizes previously identified loci and identifies population-specific signals. Scientific Reports. 2017;7(1):17075.

22. Seyerle AA, Lin HJ, Gogarten SM, Stilp A, Giráldez RM, Soliman E, et al. Genome-wide association study of PR interval in Hispanics/Latinos identifies novel locus at ID2. Heart. 2018;104(11):904-11.

23. Kerr KF, Avery CL, Lin HJ, Raffield LM, Zhang QS, Browning BL, et al. Genome-wide association study of heart rate and its variability in Hispanic/Latino cohorts. Heart Rhythm. 2017;14(11):1675-84.

24. Ward LD, Kellis M. HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. Nucleic Acids Research. 2016;44(D1):D877-D81.

25. Popejoy AB, Fullerton SM. Genomics is failing on diversity. Nature. 2016;538(7624):161.

26. Conomos MP, Laurie CA, Stilp AM, Gogarten SM, McHugh CP, Nelson SC, et al. Genetic diversity and association studies in US Hispanic/Latino populations:

applications in the Hispanic Community Health Study/Study of Latinos. The American Journal of Human Genetics. 2016;98(1):165-84.

27. Mikhailov AT, Torrado M. Myocardial transcription factors in diastolic dysfunction: clues for model systems and disease. Heart Failure Reviews. 2016;21(6):783-94.

28. Huang J, Lu MM, Cheng L, Yuan L-J, Zhu X, Stout AL, et al. Myocardin is required for cardiomyocyte survival and maintenance of heart function. Proceedings of the National Academy of Sciences. 2009;106(44):18734-9.

29. Wang L, Di Tullio MR, Beecham A, Slifer S, Rundek T, Homma S, et al. A comprehensive genetic study on left atrium size in Caribbean Hispanics identifies potential candidate genes in 17p10. Circulation: Cardiovascular Genetics. 2010;3(4):386-92.

30. van Setten J, Brody JA, Jamshidi Y, Swenson BR, Butler AM, Campbell H, et al. PR interval genome-wide association meta-analysis identifies 50 loci associated with atrial and atrioventricular electrical activity. Nature Communications. 2018;9(1):2904.

31. Nielsen JB, Thorolfsdottir RB, Fritsche LG, Zhou W, Skov MW, Graham SE, et al. Biobank-driven genomic discovery yields new insight into atrial fibrillation biology. Nature Genetics. 2018;50(9):1234-9.

32. Den Hoed M, Eijgelsheim M, Esko T, Brundel BJ, Peal DS, Evans DM, et al. Identification of heart rate-associated loci and their effects on cardiac conduction and rhythm disorders. Nature Genetics. 2013;45(6):621-31.

33. Howie B, Marchini J, Stephens M. Genotype imputation with thousands of genomes. G3: Genes, Genomes, Genetics. 2011;1(6):457-70.

34. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010;26(17):2190-1.

35. Sofer T, Heller R, Bogomolov M, Avery CL, Graff M, E. NK, et al. A powerful statistical framework for generalization testing in GWAS, with application to the HCHS/SOL. Genetic Epidemiology. 2017;41(3):251-8.

36. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489(7414):57.

37. Romanoski CE, Glass CK, Stunnenberg HG, Wilson L, Almouzni G. Epigenomics: Roadmap for regulation. Nature. 2015;518(7539):314-6.

Supporting information

S1 Fig. QQ plot of Hispanic/Latino QRS duration GWAS meta-analysis

S2 Fig. Individual Cohort QQ plots

S3 Fig. Forest plots showing 95% confidence intervals of the effect size of each Hispanic/Latino index SNP on QRS duration in milliseconds across each contributing study and the combined meta-analysis.

S4 Fig. Regional association plots showing all results in the European, Hispanic/Latino, and African American QRS duration GWAS surrounding each of the Hispanic/Latino significant loci. Plots created with LocusZoom software.[21] The index SNP in each figure is labeled and colored purple. All other SNPs in the region are plotted at their significance levels. The color of each SNP corresponds to the linkage disequilibrium (r^2) between the plotted SNP and the index SNP.

S5 Fig. Regional association plots for the HCHS/SOL cohort for each of the index SNPs which was imputed rather than being directly genotyped in the cohort. Plots created with LocusZoom software.[21] Each plot shows all GWAS results surrounding the index SNP. The index SNP in each figure is labeled and colored purple. All other SNPs in the region are plotted at their significance levels. The color of each SNP corresponds to the linkage disequilibrium (r^2) between the plotted SNP and the index SNP. SNPs plotted as circles were directly genotyped, and SNPs plotted as X's were imputed.

S1 Table. Participant characteristics from the studies contributing to the meta-analysis.

S2 Table. Heterogeneity tests for index SNPs across the participating cohorts.

S3 Table. Summary of genetic imputation of index SNPs across the participating cohorts.

S4 Table. Coded allele frequencies for index SNPs significantly associated with QRS duration among participants of Hispanic/Latino ancestry (n=15,124).

S5 Table. Index SNPs in Hispanic/Latino QRS duration GWAS (n=15,124) and corresponding SNPs in the European QRS duration GWAS (n=40,407).

S6 Table. Index SNPs in Hispanic/Latino QRS duration GWAS (n=15,124) and corresponding SNPs in the African American QRS duration GWAS (n=13,301).

S7 Table. Significant results for Welch's t-tests of differences in effect sizes at Hispanic/Latino index SNPs across Hispanic/Latino, European, and African American GWAS results.

S8 Table. Generalization of the associations in Hispanic/Latino QRS duration GWAS meta-analysis (n=15,124) for previously discovered loci from QRS duration GWAS among European (n=40,407), African American, (n=13,301), and East Asian (n=6805) populations, and a European-African America meta-analysis (n=53,708).

S9 Table. Associations of QRS duration index SNPs with other ECG phenotypes (QT duration, PR duration, Heart Rate, and Heart Rate Variability) in the same Hispanic/Latino study population.

S10 Table. Summary of functional annotations in heart tissues for QRS duration significant loci in the HaploReg v4.1 database.

S11 Table. ECG and genotype measurement methods for the participating cohorts.

S1 Appendix. Descriptions of cohort studies used in analysis and references for supporting information.

Chapter 2: Genetic Risk Score is Associated with Ischemic Stroke among UK Biobank Participants with Atrial Fibrillation

Background

Atrial fibrillation (AF), the most common sustained cardiac arrhythmia, is associated with 4-6 fold increased risk of stroke often due to thromboembolism.(1, 2) To reduce risk of stroke, patients with AF at particularly high stroke risk are given anticoagulation therapy. The CHA_2DS_2-VAS_C score has been the primary metric used to assess stroke risk for AF patients since 2010. The CHA_2DS_2-VAS_C score stratifies AF patients on a scale from 0-9.(3) Current guidelines from the AHA, updated in 2019, generally recommend that oral anticoagulants be prescribed to men with AF and a CHA_2DS_2-VAS_C score of ≥2, and to women with AF and a score ≥3, as female sex itself contributes 1 CHA_2DS_2-VAS_C points to women, but many studies do not support female sex as a contributor to increased risk of stroke.(4) The CHA_2DS_2-VAS_C score has other limitations, and the individual CHA_2DS_2-VAS_C components are not consistently associated with increased stroke risk in multivariable models. Moreover, CHA_2DS_2-VAS_C does not take into account the known genetic risk of stroke.

Genetic risk factors for ischemic stroke have been identified through genome wide association studies (GWAS)(5-11). The largest of these identified 32 single nucleotide

polymorphisms (SNPs) that reached GWAS significance for all stroke, ischemic stroke, or a subtype of ischemic stroke (cardioembolic, large vessel, or small vessel subtypes) among a multi-ethnic cohort.(6) In a recent analysis of European-descent participants enrolled in several cardiometabolic diseases trials, Marston et al, found individuals in the top tertile of a GRS constructed using these 32 multi-ethnic SNPs from the reference stroke GWAS had a 24% increased risk of ischemic stroke compared to those in the lowest tertile.(12)

We sought to extend these findings focusing on ischemic stroke risk among participants with AF using the UK Biobank study, a prospective cohort study of approximately 500,000 individuals from the United Kingdom.(13)

Methods

Study Population

All participants included in the study were from the UK Biobank cohort.(13) Participants were recruited between 2006 and 2010, and were between the ages of 40 and 69 at the time of recruitment. At an initial visit, researchers conducted phenotypic measurements, collected blood samples, and obtained permission to access past and future medical records from the National Health Service. A small percentage of participants (n=19,687) had up to three follow-up visits, but primarily participants were followed through medical record review and periodic questionnaires. Medical history prior to 2006 was available through medical record review.

We examined a total of 458,671 UK Biobank participants who self-identified as having European ancestry. From this group, 22,768 participants had been diagnosed with AF, either before they were enrolled in the study (prevalent AF), or after enrollment (incident AF). After excluding participants who had any type of stroke prior to their AF diagnosis (n=1166), had any type of stroke in the first week following AF diagnosis, n=415), or were diagnosed with AF before age 18 (n=2), 21,185 participants of European descent with AF were available for analyses. Because medical record data was available on all participants retrospectively, all participants were followed prospectively after their AF diagnosis. In one sensitivity analysis we included only participants who were diagnosed with AF following their enrollment in UK Biobank (n=13,717). In another sensitivity analyses, we excluded individuals who were ever documented to be treated with warfarin at any point following AF diagnosis, leaving n=18,185 in the analysis.

Genotyping

The details of genotyping and imputation in the UK Biobank study have been described elsewhere.(13) Briefly, genotyping was performed using both Axiom and custom arrays from Affymetrix. Genotyping was performed examining ~800,000 SNPs and imputed to 90 million SNPs. IMPUTE2 (14) was used for genotype imputation, and was based on reference panels from the 1000 Genomes Project and UK10K sequencing data. Both imputation and the calculation of principal components of ancestry were also computed locally by UK BioBank.

Genetic Risk Score

A total of 32 GWAS significant SNPs were identified to be associated with all stroke, ischemic stroke, or a subtype of ischemic stroke (cardioembolic, large vessel, or small vessel subtypes) in the stroke GWAS, a meta-analysis of 40,585 cases of stroke among 446,696 multi-ethnic participants.(6) Because we are examining individuals of European descent, we limited our GRS to SNPs found to have compelling association with stroke among the European descent individuals. Specifically, of the 32 SNPs, we only included in our genetic risk score (GRS) the following SNPs: (1) those which were genome-wide significant at P<5e-8 in the European-ancestry analyses for all stroke, or ischemic stroke subtypes (12 SNPs); or (2) those that were genome-wide significant in the multi-ethnic analyses and were significant at P<9.9e-07 in the European-only analysis (21 SNPs). Because these were overlapping sets, the total SNPs included in the GRS was 27.

The GRS was calculated for each participant by multiplying each variant's weight (the effect size [logged odds ratio] from the reference ischemic stroke GWAS among individuals of European descent) by the patient's imputed allelic dosage, and summing across all variants. Cutoff points between GRS tertiles were calculated based on the distribution of the GRS in the 458,671 UK Biobank participants with self-identified European ancestry. The GRS distribution had a mean of 21.72 and standard deviation of 3.40.

In sensitivity analyses, we created two additional risk scores. First, we limited the SNPs included in the GRS to those with a more stringent p-value cut-off among those of

European descent (P<9.9e-08 n=16 SNPs). Second, we excluded the 5 SNPs associated with AF (n=22 SNPs).

Outcome of interest and censoring

Ischemic stroke was the outcome of interest in all analyses. Participants who had a stroke that was not identified as an ischemic stroke were censored from further follow-up from that point in time. Follow-up data was available through March 31, 2017. Participants with AF were followed for an average of 6.72 years. Participants who had an ischemic stroke within 1 week of AF diagnosis, before or after, were analyzed separately.

Statistical Analysis

Hazard ratios for ischemic stroke were calculated using mixed effects Cox proportional hazards models, using the coxme R package (15) to adjust for relatedness in the UK Biobank cohort using random effects. All analyses were adjusted for the first 10 principal components of ancestry and relatedness. Individuals were entered into the Cox model on the day they were diagnosed with AF (both for incident and prevalent AF cases). CHA_2DS_2-VASc score on the date of AF diagnosis was calculated using prior health history. Because transient ischemic attacks were not adjudicated in UK Biobank, and only the date of the first incidence of stroke, as opposed to all stroke events, was available in UK Biobank, these factors did not contribute to the CHA_2DS_2-VASc score. In the analysis examining the effect of the GRS on ischemic stroke stratified by

CHA_2DS_2-VASc score, the contribution of female sex was not included in the calculation of CHA_2DS_2-VASc.

The average GRS of the 389 participants who had an ischemic stroke within the same week, before or after, as their diagnosis of AF were compared to the average GRS of the 20,204 participants with AF who did not have any type of stroke during the follow-up period using a two-sided T-test, adjusting for age of AF diagnosis, sex, 10 principal components of ancestry and relatedness.

The CHA_2DS_2-VASc-G score was created by including all the components of the CHA_2DS_2-VASc score, using the same point formula for all components, and adding one point for being in the top tertile of the GRS. The same high-risk threshold of ≥2 for men and ≥3 for women was used for CHA_2DS_2-VASc-G. Improvement of the CHA_2DS_2-VASc-G score to reclassify risk for ischemic stroke was calculated using net reclassification improvement at 10 years following diagnosis of AF.

Table 1: Baseline Characteristics of European-descent UK Biobank participants at the time of their AF diagnosis

Characteristics	
Participants	21,185
Age at AF Diagnosis	63.4 (8.6)
Female	7245 (34.2%)
Hypertension	8355 (39.4%)
Diabetes	1629 (7.7%)
Heart Failure	1681 (7.9%)
Vascular Disease	4518 (21.3%)
Genetic Risk Score	21.72 (3.40)
CHA_2DS_2-VASc Score	1.65 (1.22)

Table values indicate n (%) or mean (standard deviation)

Results

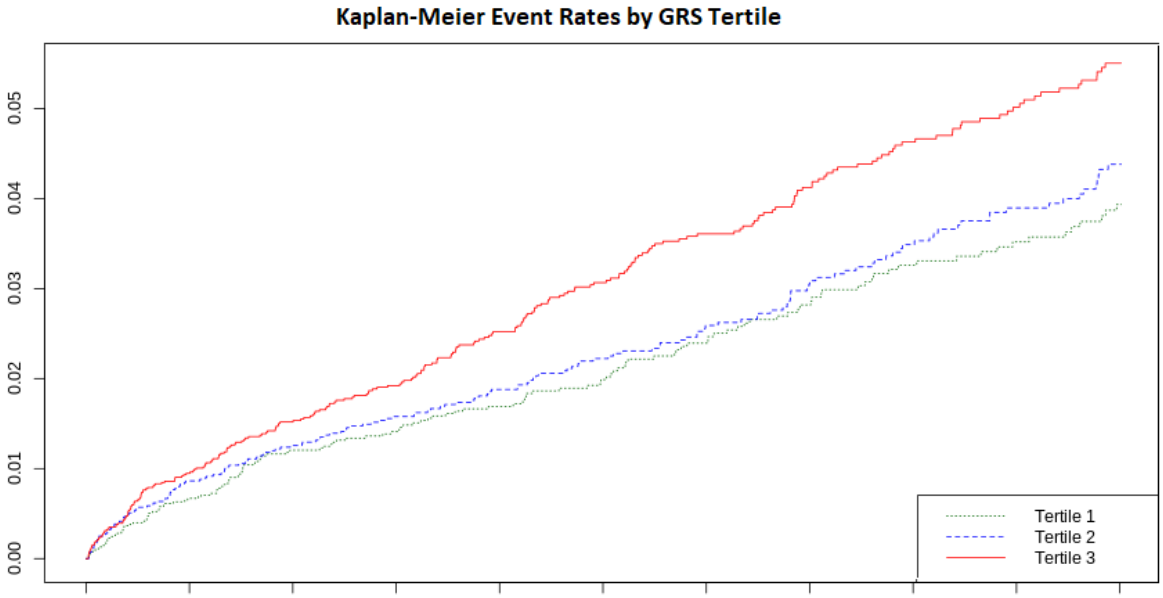
A total of 21,185 European-descent participants with AF were included in the analysis. 7468 had prevalent AF before being enrolled in the UK Biobank study, and 13,717 developed AF after enrollment. The average age of the cohort at the time of their AF diagnosis was 63 and 34% were women. Comorbidities which compose the CHA_2DS_2-VASc risk score were common among the cohort at time of AF diagnosis, including hypertension (39.4%) and vascular disease (21.3%, Table 1). 3000 (14.2%) participants reported using warfarin at some point during the follow-up period. Over an average follow-up period of 6.7 years, 981 participants with AF later went on to develop a stroke more than one week after their AF diagnosis, and 702 of these strokes were identified as ischemic (71.5%).

Of the 27 SNPs included in the GRS, 26 were positively associated with ischemic stroke in UK Biobank participants with AF, though not all reached nominal significance (Supplementary Table 1). The mean GRS value among AF patients was 21.7. A one unit increase in the GRS was found to be significantly associated with a 3.5% increased risk ischemic stroke in the full UKBB cohort who self-identified as having European ancestry (HR=1.035, 95% CI, 1.027-1.043), after adjusting for sex (Supplementary Table 2).

Table 2: Association of GRS with ischemic stroke, adjusted for sex, age, PCs and relatedness. (N=21,185, 702 events)

	Hazard Ratio (95% CI)	P-Value
GRS	1.046 (1.024-1.069)	4.8e-05
Female	0.883 (0.752-1.036)	0.13
Age at AF Diagnosis	1.078 (1.065-1.091)	2.3e-36

Among participants with AF, each unit increase in the GRS was associated with a 4.6% (HR=1.046, 95% CI, 1.024-1.069, Table 2) increased risk of ischemic stroke, after adjusting for sex, the age of AF diagnosis, principal components of ancestry (PCs), and relatedness. Performance of the GRS was also analyzed by tertiles (Figure 1, Supplementary Table 3). The highest risk tertile of GRS was associated with a 28% (HR=1.28, 95% CI, 10-48%) increased risk of ischemic stroke compared to the lower two tertiles combined. When adjusted for CHA_2DS_2-VASc score, the association of the GRS with ischemic stroke remained essentially unchanged, with a 1 unit increase in GRS being associated with a 4.5% (HR=1.045, 95% CI, 1.023-1.069) increased risk (Supplementary Table 4).



Number Remaining at Risk by Year

Tertile 1	6333	5270	4485	4026	3472	3050	2263	2319	2019	1740	1449
Tertile 2	6978	5844	5175	4482	3900	3429	2968	2590	2263	1885	1625
Tertile 3	7874	6759	5901	5196	4458	3927	3317	3009	2623	2289	1951

Figure 1: Kaplan-Meier event rates for ischemic stroke by genetic risk score tertile through 10 years following AF diagnosis. Hazard Ratio for GRS tertile 3 vs tertile 1=1.39 (95% CI=1.15-1.67, P-value adjusted for age and sex=0.00055) and tertile 3 vs tertiles 1 & 2 combined=1.28 (95% CI=1.10-1.48, adjusted P-value=0.0014).

Two sensitivity analyses were performed examining genetic risk scores composed of smaller numbers of SNPs. A GRS which excluded the 5 SNPs previously associated with AF was strongly associated with ischemic stroke (P=6.1e-04, Supplementary Table 5), as was a GRS which included just the 16 SNPs which reached a more stringent P<9.9e-08 threshold for inclusion as measured in the original stroke GWAS (P=0.005,

Supplementary Table 6). A sensitivity analysis which excluded 3000 participants who used warfarin at any point during the follow-up period, yielded a similar association between GRS and ischemic stroke, with a one point increase in GRS being associated with a 5.0% increase in ischemic stroke (HR=1.050, 95% CI, 1.024-1.077, Supplementary Table 7).

The GRS was also found to be strongly associated with ischemic stroke in a sensitivity analysis excluding participants with prevalent AF at enrollment into the UK Biobank study, with a one point increase in GRS being associated with a 6.1% increase in ischemic stroke (HR=1.061, 95% CI, 1.026-1.097, Supplementary Table 8).

The effect size of being in the highest GRS tertile was compared to the risk factors that compose the CHA_2-DS_2-VAS_c score in a multivariable model. When adjusting for each of the CHA_2-DS_2-VAS_c score components, being in highest tertile of the GRS was associated with a 27% increased risk of ischemic stroke (HR=1. 27; 95% CI, 1.10-1.48; Figure 2, Supplementary Table 9) compared to being in the combined lower two tertiles. The estimated effect of being in the highest GRS risk tertile had as large or larger effect on ischemic stroke risk than four clinically used components of the CHA_2-DS_2-VAS_c score: female sex (HR=0.95; 95% CI, 0.81-1.11), vascular disease (HR=1.11; 95% CI, 0.92-1.34), hypertension (HR=1.22; 95% CI, 1.02-1.45), and heart failure (HR=1.28; 95% CI, 0.97-1.70). Of the above four, only hypertension was nominally associated with stroke risk in UK Biobank participants with AF. The only components of CHA_2-DS_2-VAS_c which had substantially larger effect sizes than GRS were age (65 to <75 HR=1.82,

95% CI, 1.52-2.18; age 75+ HR=3.66, 95% CI, 2.30-5.83; both groups compared to those <65) and diabetes (HR=1.65, 95% CI, 1.29-2.12).

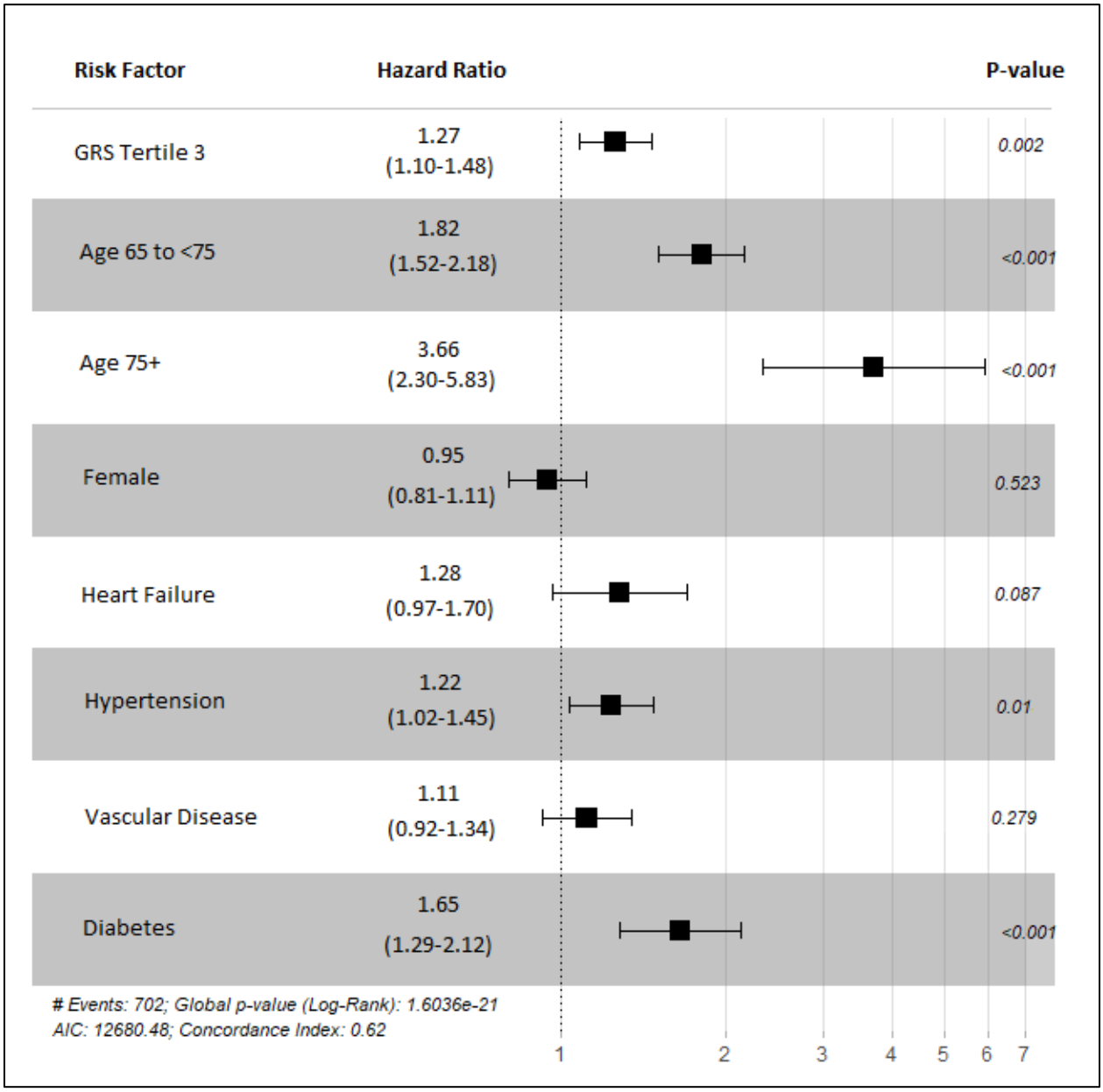


Figure 2: Association of GRS tertile 3 and CHA_2-DS_2-VAS_c score components with ischemic stroke adjusted for PCs and relatedness. Full results presented in Supplementary Table 9.

The GRS was positively associated with ischemic stroke across all strata of CHA_2-DS_2-VAS_c score (Supplementary Table 9), though these associations were not all significant at the 5% level. The interaction of GRS and CHA_2-DS_2-VAS_c score was not significant (*P*=0.19) when included in a Cox model adjusting for GRS, CHA_2DS_2-VASc, principal components of ancestry, and relatedness.

In secondary analyses, we examined the 389 individuals who experienced an ischemic stroke up to one week before or after their first AF diagnosis and therefore were excluded from the above analyses. These participants had a GRS 0.47 units higher (mean GRS=22.17, adjusted P=0.005) than the 20,204 participants with AF who did not go on to experience a stroke (mean GRS=21.70).

We examined incidence rate of stroke stratified by CHA_2-DS_2-VAS_c score (Supplementary table 10). We note that at a CHA_2-DS_2-VAS_c score of 2 (or 3 in women), the annual stroke incidence was only 0.0060. At a CHA_2-DS_2-VAS_c score of 4+, the annual stroke incidence was 0.015.

573 participants had an ischemic stroke within 10 years of their diagnosis with AF. Upon their diagnosis with AF, 344 participants met the threshold for anticoagulation using the CHA_2DS_2-VASc-G score who later had an ischemic stroke in the next 10 years, compared to 262 participants using the CHA_2DS_2-VASc score, resulting in a 14.3% increase in the true positive rate (Supplementary Table 11). In turn, using the CHA_2DS_2-VASc-G score, 12481 participants met the threshold for anticoagulation who did not go

on to have an ischemic stroke, compared to 10,227 participants using the CHA_2DS_2-VASc score, and a 10.9% increase in the false positive rate. The net reclassification index was +3.4%.

Discussion

We constructed a GRS associated with ischemic stroke based on the largest GWAS of stroke and stroke subtypes performed to date. We tested the performance of the GRS in European ancestry participants with AF in the UK Biobank cohort, a large prospective cohort study which incorporates National Health Service data. Our analysis included 21,185 participants with AF, 702 of whom went on to have an ischemic stroke.

Our results show that a GRS of stroke SNPs is strongly associated with incident ischemic stroke in an independent population of 21,185 European descent individuals with AF, Moreover, the magnitude of the effect on stroke risk for those in the highest tertile of genetic risk was found to be as large or larger than the effect on stroke risk of four components of the CHA_2-DS_2-VAS_c risk score: female sex, hypertension, vascular disease, and heart failure. We also found some evidence that adding the top tertile of genetic risk as a point within the CHA_2DS_2-VASc score has the potential to improve classification for ischemic stroke.

Examining a multi-ethnic set of SNPs, Marston, et al. found that the highest tertile of the GRS was associated with a 29% increased risk of stroke compared to the lowest tertile among those of European descent with AF and a CHA_2DS_2-VAS_C score≥2.(12)

Extending the work done by Marston, our study validates those findings in an independent and much larger sample of participants with AF. Moreover, by limiting SNPs included in the GRS to only those with evidence of association among European descent individuals, our GRS had a larger magnitude of effect in the UK Biobank European-descent AF population. In our study, the highest risk tertile was associated with a 39% increased risk of stroke compared with the lowest tertile, and a 28% increased risk of stroke in the top tertile compared to the lower two tertiles combined.

There seems to be mounting evidence that the construction of the CHA_2-DS_2-VAS_c score should be reassessed and potentially revised. This study reinforces findings from recently published studies which find the effect of GRS on ischemic stroke to be at least comparable to some components of the CHA_2-DS_2-VAS_c score.(12, 16) Additionally, the association of sex and ischemic stroke risk is increasingly dubious. Although sex is still included in the CHA_2-DS_2-VASc model, some practice guidelines are now recommending that men with AF receive anticoagulation therapy with a score ≥2, and women with a score ≥3, effectively ignoring sex in the decision.(4) Finally, many studies are finding that the components of CHADS VASC are not equal in effect. For instance, the association of age and diabetes with stroke risk among participants with AF is much larger in effect than that of the other components of CHA_2-DS_2-VAS_c.

Historically, prophylactic anticoagulation has been recommended to individuals with at least a ~2% annual risk of stroke, this corresponds to a CHA_2-DS_2-VAS_c score of 2.(3) For those will lower risk of stroke, the benefits of anticoagulation are not thought to

outweigh the risk of bleed events. Several population-based studies using medical record data suggest that annual stroke risk may be lower than that reported in the original CHA_2-DS_2-VAS_c study. Our data from the UK Biobank similarly suggests that at individuals with a CHA_2-DS_2-VAS_c score ≥2 (≥3 for women) have a much lower than 2% annual risk of stroke.

Creating a revised ischemic stroke risk score for individuals with AF that excludes sex, reweights age and diabetes, and incorporates genetics has the potential to improve outcomes by better identifying individuals who reach the 2% annual risk threshold, and therefore could benefit from anticoagulation therapy enough to theoretically outweigh the harm of bleeding risk due to anticoagulation. Hypothetically, individuals with low genetic risk could also be reclassified as not needing anticoagulation therapy as well. Some other attempts to add genetics to the CHA_2-DS_2-VAS_c model have been made, including a study by O'Sullivan that used a polygenic risk score of nearly 500,000 SNPs to reclassify patients from their initial CHA_2-DS_2-VAS_c score in accordance with their genetic risk in the UK Biobank.(16) O'Sullivan et al used data from a GWAS of cardioembolic stroke and not ischemic stroke as we have done in our study. Some improvement on the basic CHA_2-DS_2-VAS_c model was made by adding one point for being in the top quartile of genetic risk, and subtracting one point for being in the bottom quartile.

It should be noted that there have been major shifts in the rates of anticoagulation in individuals with AF, and the types of anticoagulants prescribed since the CHA_2-DS_2-

VAS_c model was developed. Since 2010, the rates of anticoagulation in AF patients with a CHA_2DS_2-VAS_C score ≥2 have increased the US, from about 50% to over 60% by 2017.(17) Prior to 2008, warfarin was the only oral anticoagulant being prescribed in the US. The first direct oral anticoagulants entered the market in 2008. By 2017 DOACs made up roughly 80% of all oral anticoagulant prescriptions written for AF patients.(18) Direct oral anticoagulants have fewer downsides than warfarin, and a lower risk of serious bleeding events. As such, the annual risk of ischemic stroke where the benefits of anticoagulation outweigh the risks of bleeding may now be lower than the ~2% annual risk that has been used historically.

Limitations

Several limitations exist for this study. First, our study was restricted to participants who self-identified as European ancestry as the sample size of non-European participants with AF was too small to examine. For this reason, we also created a GRS based on SNPs which were associated with stroke in a European-ancestry population. It is certainly possible that the genetic architecture of ischemic stroke differs somewhat by genetic ancestry. Future studies will be needed to examine more diverse populations to identify risk scores associated with increased stroke risk in other ethnic groups.

The UKBB study population has been criticized for suffering from healthy volunteer bias and not being representative of the sociodemographic or baseline health of the sample population. This may be one reason why the CHA_2DS_2-VAS_C score-specific incidence rates of ischemic stroke in our study are lower than have been reported in some clinical

trials, though incidence rates of stroke are similar to other population-based studies. It is also possible that medical records are incomplete, and we are missing potential stroke events.

It is possible that retrospectively assessing UK Biobank participants for AF and stroke prior to their enrollment in the study could introduce survivorship bias. However, a sensitivity analysis that including only those participants with incident AF yielded similar strongly significant associations between the GRS and ischemic stroke compared to our primary analysis.

Another possibility for the lower than expected a CHA_2DS_2-VAS_C score-specific incidence rates of ischemic stroke may be that anticoagulant use in UK Biobank is not well characterized. Only 14.2% of UK Biobank participants with AF were ever reported as having used warfarin. For participants with a CHA_2DS_2-VAS_C score≥2 for men or >3 for women, only 7.7% reported using warfarin. No data was available on the use of direct-oral anticoagulants in UK Biobank.

Because we only had access to the “first reported” incidence of a stroke with the UK Biobank data, it wasn’t possible to assess the effects of prior stroke on stroke recurrence among participants with AF. Neither a history of stroke nor transient ischemic attacks were included in the CHA_2DS_2-VAS_C score calculation. Marston et al found the GRS was more predictive in individuals with AF who had no history of stroke or transient ischemic attack, and wasn’t strongly predictive in those with such a

history.(12) However, because all individuals with AF who have a history of stroke and/or TIA meet the anticoagulation threshold in CHA_2DS_2-VAS_C model, these patients are less relevant to the effort to use a GRS to identify AF patients who are not using anticoagulants but may benefit from them.

Conclusion

A genetic risk score for ischemic stroke was found highly predictive of ischemic stroke in participants with AF and was independent of traditional clinical risk factors. The highest GRS risk tertile had a comparable effect on ischemic stroke risk as traditionally used clinical risk factors such as hypertension, heart failure, and vascular disease. Revising the CHA_2DS_2-VAS_C model to incorporate genetics has the potential to improve the precision of stroke prediction in AF patients, which could result in fewer strokes as well as fewer bleeding events associated with anticoagulation therapy.

References

1. Miyasaka Y, Barnes ME, Gersh BJ, Cha SS, Bailey KR, Abhayaratna WP, Seward JB, Tsang TS. Secular trends in incidence of atrial fibrillation in Olmsted County, Minnesota, 1980 to 2000, and implications on the projections for future prevalence. Circulation. 2006;114(2):119-25.

2. Morillo CA, Banerjee A, Perel P, Wood D, Jouven X. Atrial fibrillation: the current epidemic. Journal of geriatric cardiology: JGC. 2017;14(3):195.

3. Association DwtscotEHR, Surgery EbtEAfC-T, Members ATF, Camm AJ, Kirchhof P, Lip GY, Schotten U, Savelieva I, Ernst S, Van Gelder IC. Guidelines for the management of atrial fibrillation: the Task Force for the Management of Atrial Fibrillation of the European Society of Cardiology (ESC). European heart journal. 2010;31(19):2369-429.

4. January CT, Wann LS, Calkins H, Chen LY, Cigarroa JE, Cleveland JC, Ellinor PT, Ezekowitz MD, Field ME, Furie KL. 2019 AHA/ACC/HRS focused update of the 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Rhythm Society. Journal of the American College of Cardiology. 2019;74(1):104-32.

5. Traylor M, Farrall M, Holliday EG, Sudlow C, Hopewell JC, Cheng YC, Fornage M, Ikram MA, Malik R, Bevan S, Thorsteinsdottir U, Nalls MA, Longstreth W, Wiggins KL, Yadav S, Parati EA, Destefano AL, Worrall BB, Kittner SJ, Khan MS, Reiner AP, Helgadottir A, Achterberg S, Fernandez-Cadenas I, Abboud S, Schmidt R, Walters M, Chen WM, Ringelstein EB, O'Donnell M, Ho WK, Pera J, Lemmens R, Norrving B, Higgins P, Benn M, Sale M, Kuhlenbaumer G, Doney AS, Vicente AM, Delavaran H, Algra A, Davies G, Oliveira SA, Palmer CN, Deary I, Schmidt H, Pandolfo M, Montaner J, Carty C, de Bakker PI, Kostulas K, Ferro JM, van Zuydam NR, Valdimarsson E, Nordestgaard BG, Lindgren A, Thijs V, Slowik A, Saleheen D, Pare G, Berger K, Thorleifsson G, Hofman A, Mosley TH, Mitchell BD, Furie K, Clarke R, Levi C, Seshadri S, Gschwendtner A, Boncoraglio GB, Sharma P, Bis JC, Gretarsdottir S, Psaty BM, Rothwell PM, Rosand J, Meschia JF, Stefansson K, Dichgans M, Markus HS. Genetic

risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide association studies. Lancet Neurol. 2012;11(11):951-62. Epub 2012/10/09. doi: 10.1016/s1474-4422(12)70234-x. PubMed PMID: 23041239; PMCID: PMC3490334.

6. Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, Rutten-Jacobs L, Giese AK, van der Laan SW, Gretarsdottir S, Anderson CD, Chong M, Adams HHH, Ago T, Almgren P, Amouyel P, Ay H, Bartz TM, Benavente OR, Bevan S, Boncoraglio GB, Brown RD, Jr., Butterworth AS, Carrera C, Carty CL, Chasman DI, Chen WM, Cole JW, Correa A, Cotlarciuc I, Cruchaga C, Danesh J, de Bakker PIW, DeStefano AL, den Hoed M, Duan Q, Engelter ST, Falcone GJ, Gottesman RF, Grewal RP, Gudnason V, Gustafsson S, Haessler J, Harris TB, Hassan A, Havulinna AS, Heckbert SR, Holliday EG, Howard G, Hsu FC, Hyacinth HI, Ikram MA, Ingelsson E, Irvin MR, Jian X, Jimenez-Conde J, Johnson JA, Jukema JW, Kanai M, Keene KL, Kissela BM, Kleindorfer DO, Kooperberg C, Kubo M, Lange LA, Langefeld CD, Langenberg C, Launer LJ, Lee JM, Lemmens R, Leys D, Lewis CM, Lin WY, Lindgren AG, Lorentzen E, Magnusson PK, Maguire J, Manichaikul A, McArdle PF, Meschia JF, Mitchell BD, Mosley TH, Nalls MA, Ninomiya T, O'Donnell MJ, Psaty BM, Pulit SL, Rannikmae K, Reiner AP, Rexrode KM, Rice K, Rich SS, Ridker PM, Rost NS, Rothwell PM, Rotter JI, Rundek T, Sacco RL, Sakaue S, Sale MM, Salomaa V, Sapkota BR, Schmidt R, Schmidt CO, Schminke U, Sharma P, Slowik A, Sudlow CLM, Tanislav C, Tatlisumak T, Taylor KD, Thijs VNS, Thorleifsson G, Thorsteinsdottir U, Tiedt S, Trompet S, Tzourio C, van Duijn CM, Walters M, Wareham NJ, Wassertheil-Smoller S, Wilson JG, Wiggins KL, Yang Q, Yusuf S, Bis JC, Pastinen T, Ruusalepp A, Schadt EE,

Koplev S, Bjorkegren JLM, Codoni V, Civelek M, Smith NL, Tregouet DA, Christophersen IE, Roselli C, Lubitz SA, Ellinor PT, Tai ES, Kooner JS, Kato N, He J, van der Harst P, Elliott P, Chambers JC, Takeuchi F, Johnson AD, Sanghera DK, Melander O, Jern C, Strbian D, Fernandez-Cadenas I, Longstreth WT, Jr., Rolfs A, Hata J, Woo D, Rosand J, Pare G, Hopewell JC, Saleheen D, Stefansson K, Worrall BB, Kittner SJ, Seshadri S, Fornage M, Markus HS, Howson JMM, Kamatani Y, Debette S, Dichgans M. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. Nat Genet. 2018;50(4):524-37. Epub 2018/03/14. doi: 10.1038/s41588-018-0058-3. PubMed PMID: 29531354; PMCID: PMC5968830.

7. Loci associated with ischaemic stroke and its subtypes (SiGN): a genome-wide association study. Lancet Neurol. 2016;15(2):174-84. Epub 2015/12/29. doi: 10.1016/s1474-4422(15)00338-5. PubMed PMID: 26708676; PMCID: PMC4912948.

8. Nielsen JB, Thorolfsdottir RB, Fritsche LG, Zhou W, Skov MW, Graham SE, Herron TJ, McCarthy S, Schmidt EM, Sveinbjornsson G, Surakka I, Mathis MR, Yamazaki M, Crawford RD, Gabrielsen ME, Skogholt AH, Holmen OL, Lin M, Wolford BN, Dey R, Dalen H, Sulem P, Chung JH, Backman JD, Arnar DO, Thorsteinsdottir U, Baras A, O'Dushlaine C, Holst AG, Wen X, Hornsby W, Dewey FE, Boehnke M, Kheterpal S, Mukherjee B, Lee S, Kang HM, Holm H, Kitzman J, Shavit JA, Jalife J, Brummett CM, Teslovich TM, Carey DJ, Gudbjartsson DF, Stefansson K, Abecasis GR, Hveem K, Willer CJ. Biobank-driven genomic discovery yields new insight into atrial fibrillation biology. Nat Genet. 2018;50(9):1234-9. Epub 2018/08/01. doi: 10.1038/s41588-018-0171-3. PubMed PMID: 30061737; PMCID: PMC6530775.

9. Lubitz SA, Parsons OE, Anderson CD, Benjamin EJ, Malik R, Weng LC, Dichgans M, Sudlow CL, Rothwell PM, Rosand J, Ellinor PT, Markus HS, Traylor M. Atrial Fibrillation Genetic Risk and Ischemic Stroke Mechanisms. Stroke. 2017;48(6):1451-6. Epub 2017/05/05. doi: 10.1161/strokeaha.116.016198. PubMed PMID: 28468926; PMCID: PMC7438180.

10. Gretarsdottir S, Thorleifsson G, Manolescu A, Styrkarsdottir U, Helgadottir A, Gschwendtner A, Kostulas K, Kuhlenbäumer G, Bevan S, Jonsdottir T, Bjarnason H, Saemundsdottir J, Palsson S, Arnar DO, Holm H, Thorgeirsson G, Valdimarsson EM, Sveinbjörnsdottir S, Gieger C, Berger K, Wichmann HE, Hillert J, Markus H, Gulcher JR, Ringelstein EB, Kong A, Dichgans M, Gudbjartsson DF, Thorsteinsdottir U, Stefansson K. Risk variants for atrial fibrillation on chromosome 4q25 associate with ischemic stroke. Ann Neurol. 2008;64(4):402-9. Epub 2008/11/11. doi: 10.1002/ana.21480. PubMed PMID: 18991354.

11. Gudbjartsson DF, Holm H, Gretarsdottir S, Thorleifsson G, Walters GB, Thorgeirsson G, Gulcher J, Mathiesen EB, Njølstad I, Nyrnes A, Wilsgaard T, Hald EM, Hveem K, Stoltenberg C, Kucera G, Stubblefield T, Carter S, Roden D, Ng MC, Baum L, So WY, Wong KS, Chan JC, Gieger C, Wichmann HE, Gschwendtner A, Dichgans M, Kuhlenbäumer G, Berger K, Ringelstein EB, Bevan S, Markus HS, Kostulas K, Hillert J, Sveinbjörnsdóttir S, Valdimarsson EM, Løchen ML, Ma RC, Darbar D, Kong A, Arnar DO, Thorsteinsdottir U, Stefansson K. A sequence variant in ZFHX3 on 16q22 associates with atrial fibrillation and ischemic stroke. Nat Genet. 2009;41(8):876-8. Epub 2009/07/15. doi: 10.1038/ng.417. PubMed PMID: 19597491; PMCID: PMC2740741.

12. Marston NA, Patel PN, Kamanu FK, Nordio F, Melloni GM, Roselli C, Gurmu Y, Weng L-C, Bonaca MP, Giugliano RP. Clinical application of a novel genetic risk score for ischemic stroke in patients with cardiometabolic disease. Circulation. 2021;143(5):470-8.

13. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O'Connell J. The UK Biobank resource with deep phenotyping and genomic data. Nature. 2018;562(7726):203-9.

14. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS genetics. 2009;5(6):e1000529.

15. Therneau TM, Therneau MTM. Package 'coxme'. R package version. 2015;2(5).

16. O'Sullivan JW, Shcherbina A, Justesen JM, Turakhia M, Perez M, Wand H, Tcheandjieu C, Clarke SL, Rivas MA, Ashley EA. Combining clinical and polygenic risk improves stroke prediction among individuals with atrial fibrillation. Circulation: Genomic and Precision Medicine. 2020.

17. Marzec LN, Wang J, Shah ND, Chan PS, Ting HH, Gosch KL, Hsu JC, Maddox TM. Influence of direct oral anticoagulants on rates of oral anticoagulation for atrial fibrillation. Journal of the American College of Cardiology. 2017;69(20):2475-84.

18. Zhu J, Alexander GC, Nazarian S, Segal JB, Wu AW. Trends and variation in oral anticoagulant choice in patients with atrial fibrillation, 2010–2017. Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy. 2018;38(9):907-20.

Supporting information

Supplementary Table 1. Risk alleles, effect size, and frequency, and strength of association of the 27 SNPs included in the GRS.

Supplementary Table 2. Association of GRS with ischemic stroke in all white UKBB participants, also adjusted for sex, PCs and relatedness. (N=458,671, 5577 events)

Supplementary Table 3. Association of GRS tertiles with ischemic stroke, adjusted for sex, age, PCs and relatedness. (N=21,185, 702 events)

Supplementary Table 4. Association of GRS with ischemic stroke, adjusted for CHA_2-DS_2-VAS_c score, PCs and relatedness. (N=21,185, 702 events)

Supplementary Table 5. Association of GRS score with ischemic stroke, adjusted for PCs and relatedness. The 5 SNPs independently associated with AF were not included in the GRS calculation. (N=21,185, 702 events)

Supplementary Table 6. Association of 16 SNP GRS with ischemic stroke, adjusted for PCs and relatedness. (N=21,185, 702 events)

Supplementary Table 7. Association of GRS with ischemic stroke, adjusted for PCs and relatedness, and excluding participants who used warfarin at any time during follow-up. (N=18,185, 511 events)

Supplementary Table 8. Association of the highest GRS tertile (vs lower 2/3) with ischemic stroke, adjusted for CHA_2-DS_2-VAS_c score components, PCs and relatedness. (N=21,185, 702 events)

Supplementary Table 9. Association of the highest GRS tertile (vs lower 2/3) with ischemic stroke, stratified by CHA_2-DS_2-VAS_c score, and adjusted for PCs and relatedness.

Chapter 3: Competitive Sports Participation Guidelines for LQTS and the Challenges of Precision Medicine

Introduction

The completion of the Human Genome Project in 2003 held the promise of a new age of precision medicine, where an individual's unique biology could be used to guide clinical care. The cost of sequencing a human genome has dramatically fallen from $100,000,000 in 2001 to under $1000 today.(1) Technological and financial barriers to the incorporation of genetic data into clinical care have largely fallen away. However, while genetic knowledge has made notable contributions to clinical care, progress towards a paradigm-changing transformation of medical care has been slower than many envisioned 20 years ago. This slow rate of progress is partially due to the staggering genetic and environmental complexity revealed to underlie many diseases. As just one example, Type 2 Diabetes, the 9th leading cause of death worldwide, is influenced by both genetics and the social environment. Epidemiological genetic studies have found hundreds of genetic variants across dozens of loci to be associated with the condition, with each variant making a very small contribution to overall risk.(2) Social factors such as education, poverty, diet and level of physical activity exert a greater effect on risk, and interactions between social and genetic factors are likely.(3) As with Type 2 Diabetes, epidemiological studies have painted an increasingly complex picture

of the etiology of many common diseases, including not only diabetes but also other important public health burdens, such as cardiovascular diseases and most cancers.(4) Designing, testing, and implementing clinical interventions for complex diseases using genetic data will be a decades-long scientific challenge, and ultimately may prove less effective than measures addressing behavior or other aspects of the social environment.(3)

While most common diseases are influenced by multiple genetic and environmental factors, millions of individuals are afflicted with rare Mendelian diseases. Because these conditions have more straightforward underlying genetic causes, the incorporation of genetic data into clinical care holds great promise for improving the diagnosis and treatment of these conditions. Genetics can be a useful tool for dividing individuals with a genetic disease into subgroups of individuals who are similar in their disease presentation and response to treatment. For example, cystic fibrosis (CF) is a rare Mendelian disease characterized by thick secretions from mucous membranes. CF can be caused by several different genetic variants, which are associated with nuances in the presentation of CF. The use of genetic testing in clinical care has been useful in predicting disease severity. Genetic data has also been crucial in the recent development of drugs to treat CF which have been approved for use in patients with specific causal genetic variants.(5)

However, even for Mendelian conditions with proven medical interventions, successfully incorporating genetic data into clinical care can be surprisingly complex. Obtaining an

individual's genetic data is only the first step in delivering effective personalized care, and many barriers could remain. The scientific and clinical implications of the genetic data could be undeveloped, the clinical institutional framework to communicate what the genetic data means to the patient might not be in place, and there could be ethical disagreements about what should be the best course of action in a given situation.

Long QT syndrome (LQTS) is a rare genetic heart condition that is illustrative of many of the challenges in delivering precision medicine. LQTS is characterized by a prolonged QT interval (an ECG measurement measuring the time it takes from ventricular depolarization to complete repolarization), and is typically caused by rare variants in 1 of 17 known genes.(6) While broad similarities are found in the presentation of LQTS regardless the specific genetic cause, important clinically relevant distinctions exist between LQTS patients depending on which gene contains the implicated causal variant. As such LQTS is broken down into subtypes based on the underlying genetic causes.

Individuals with LQTS are at increased risk for syncope and sudden cardiac death when they engage in competitive sports.(7) Although all strenuous activity can carry an elevated risk for individuals with LQTS, competitive sports have historically been regarded as more dangerous than other forms of exercise such as jogging. This is because competitive sports can entail maximum cardiac exertion, and also because epinephrine released while playing sports, which normally shortens the QT interval, can prolong the QT interval for some individuals with LQTS.(8) Because of this risk, many

individuals with LQTS seek counsel from medical providers regarding their risk from participating in competitive sports. The patient's specific genetic LQTS subtype can be informative as to the patient's risk level. While helping a LQTS patient decide whether they should participate in sports with the aid of genetic information might seem like a relatively simple and straightforward example of precision medicine, many biological factors and ethical considerations can make arriving at the best answer challenging. To illustrate these challenges, we provide a background of the biology of LQTS, an examination of the scientific challenges in diagnosing and treating LQTS and explore the ethical policymaking dilemmas surrounding athletic sports participation by young athletes with LQTS.

Long QT Syndrome Biology

LQTS is a relatively rare heart condition in which the cardiac ventricular myocytes take longer to repolarize than normal. The QT interval is a measurement of the time it takes for the ventricles to depolarize and repolarize as measured using an ECG. The length of a "normal" QT interval changes with age, and women tend to have naturally longer QT than men. LQTS refers to a naturally prolonged QT interval, although importantly, environmental factors such as certain medications can also prolong the QT interval. LQTS puts one at a significantly increased risk for episodes of torsades de pointes, a distinct type of arrhythmia which occurs when the ventricular myocytes cannot repolarize quickly enough to contract in sync with the electrical wave signaling contraction through the ventricles. The results in rapid ventricular asynchronous contraction, and ineffective pumping of blood out of the heart. Episodes of torsades de

pointes can cause syncope and are potentially fatal if they extend beyond 60 seconds and the heart rhythm is not rapidly restored to normal. Death results from the brain becoming deprived of oxygen.

Rare genetic variants in 17 different genes have so far been shown to cause LQTS, and these subtypes are designated LQT1-LQT17.(9) The condition is inherited in a Mendelian pattern, with a single gene variant being capable of causing the disease. Many of the implicated genes code for cardiac ion channel proteins. These proteins are responsible for the movement of Na^+, K^+, and Ca^{2+} ions across the cell membrane, which is necessary for the generation of the cardiac action potential, leading to the depolarization and repolarization of the heart muscle. LQTS mutations can result in erroneous movement of ions across these channels (either too much or too little, depending on the channel), and a subsequent prolongation in the refractory period before the cell can generate the next heartbeat. Most commonly, LQTS is caused by rare variants in the *KCNQ1* (LQT1 subtype, 30-35% of cases), *KCNH2* (LQT2, 25-30%), and *SCN5A* (LQT3, 5-10%) genes.(10) Not all genetic variants which cause LQTS have yet been identified, and an estimated 20-25% of hereditary LQTS has an unknown genetic cause. A very rare recessive subtype of LQTS called Jervell and Lange-Nielson syndrome (JLNS) can be caused by two copies of variants in *KCNQ1* or *KCNE1*. JLNS is a severe type of LQTS and is also characterized by bilateral hearing loss and a severely prolonged QT interval.

LQTS Treatment

Treatment for LQTS typically involves pharmaceutical and lifestyle interventions. Beta-blockers are commonly recommended to patients with LQTS, although there is substantial evidence to suggest that their efficacy is largely dependent upon which type of LQTS a patient has.(11) Beta blockers are highly effective at suppressing episodes of torsades de pointes in patients with LQT1, moderately effective in patients with LQT2, and less effective in patients with LQT3, with some studies suggesting that beta-blockers may even be harmful in these patients. Pacemakers, external defibrillators, and implantable cardioverter-defibrillators (ICDs) may be recommended to patients who are resistant to beta blocker treatment or cannot take beta blockers, again depending on which LQT subtype the patient has.

For patients with rarer subtypes of LQTS, there is insufficient epidemiological evidence available to guide treatment. However, risk levels can be inferred from biological pathway analysis. For example, LQT1 results from variants in *KCNQ1* which encodes the α-subunit of the $K_V7.1$ potassium channel. The much rarer LQT5, in turn, is caused by variants in *KCNE1* which encodes the β-subunit of the same potassium channel. In the absence of clinical or epidemiological data on LQT5 patients, a reasonable assumption can be made that the LQT5 subtype should closely resemble LQT1, because they both are caused by molecular alterations in the same potassium channel.(6)

In addition to these direct medical interventions, lifestyle modifications and activity restrictions are often recommended to patients with LQTS. Episodes of torsades de

pointes in individuals with LQTS do not usually occur at random. Rather, they are typically precipitated by “trigger” events which lead to an onset of the episode. Common triggers are strenuous exercise, swimming, loud startling noises, intense emotional moments, and sleep. Interestingly, there seems to be strong correlation between trigger events and the patient’s genetic subtype of LQTS.(11) LQT1 patients usually have cardiac events triggered by strenuous exercise, with swimming being an especially common trigger. Highly emotional moments and loud auditory stimuli are much more common triggers for LQT2 patients (26% compared to 2% in LQT1 and 7% in LQT3). LQT3 patients are most likely to experience cardiac events during sleep or resting states. Clinical practice guidelines do not tend to differentiate which activities should be avoided based on genotype. Rather, LQTS patients are typically advised to, whenever possible, avoid common triggers, such as participating in competitive sports or extreme exercise, attending amusement parks, or watching frightening movies.

Precision medicine and LQTS

The goal of precision medicine as defined by the National Research Council “…is to classify individuals into subpopulations that differ in their susceptibility to a particular disease or their response to a specific treatment. Preventative or therapeutic interventions can then be concentrated on those who will benefit, sparing expense and side effects for those who will not.”(12) The clinician’s goal for treating patients with (or potentially with) LQTS should be to accurately inform them of the risks of LQTS, inform the patient as to what activities they should be cautious of, and what activities can be fully embraced, and to discuss the potential benefits and drawbacks of

pharmaceutical/surgical interventions. There is a careful balancing between helping patients avoid potentially fatal episodes and limiting activities that would otherwise be healthy and part of a full life. Through precision medicine, this process can be informed by the patient's health history, genetics, and their personal risk tolerance. Because these factors can vary greatly from patient to patient, this approach to medicine has the potential for improved health outcomes and patient satisfaction compared to a one-size-fits-all policy for how to treat LQTS patients.

For individuals who are healthy enough to engage, there is ample evidence that competitive sports participation has several benefits. Studies in both adults and children have linked sports participation with increased health-related quality of life.(13, 14) Sports are one of the primary ways in which children engage in physical activity, and children who play sports are more likely to be physically active during adulthood. High levels of sports participation have been associated with higher perceived vitality ("feeling good") and increased levels of energy.(15) Psychologically, sports participation has been associated with several positive outcomes, including higher self-esteem, less emotional distress, and fewer depressive symptoms.(16)

At the same time, there are wide differences in the importance individuals place on the value of competitive sports participation. Some people have little inclination towards sports, others value it as a leisure activity, and, for some highly trained athletes, sports could be the most dominant force in their lives. Because competitive sports can potentially be very dangerous for them, individuals with LQTS have historically been

cautioned against playing competitive sports if they have a sufficiently long QT interval. Following this counsel could be trivially easy for some but require career ending sacrifices for others. For athletes with LQTS whose sports performance is at the highest level, the stakes couldn't be higher. Of course, it makes sense to caution athletes who are putting themselves at extreme risk about the dangers of competitive sports. But counseling athletes to make extreme lifestyle changes if they could participate safely would be a tragedy.

Ideally genetic data could be used to better sort out the individuals with LQTS who are at highest risk from those who are not, and more precisely inform athletes of their level of risk in participating in competitive sports. Those at highest risk could be cautioned to not participate in sports, and to avoid dangerous triggering events, but patients deemed to be at sufficiently low risk could be allowed to participate with relative safety. This approach faces both scientific and institutional challenges.

Scientific Challenges

Diagnosis

There are two methods of diagnosing LQTS, and both have some degree of uncertainty. Typically, LQTS is diagnosed by a cardiologist based on clinical findings. The patient undergoes a resting ECG where the QT interval is measured. Through a formula that takes into account the QT interval, age, sex, family history, health history such as unexplained fainting or seizures, and ECG anomalies, a diagnostic score is calculated. Patients with high scores are deemed "highly likely" to have LQTS, while those with

lower scores are “moderately likely” or “unlikely” to have LQTS.(17) The diagnosis thus carries some degree of uncertainty, and is complicated by the fact that roughly 2.5% of the healthy population has prolonged QT intervals not due to genetic LQTS, and 10-15% of patients with known LQTS causal mutations either do not have an abnormally prolonged QT interval, or demonstrate a prolonged interval only under certain situations (e.g. exercise), and therefore a resting ECG may not detect the disease.(18) One study found that 40% of individuals who received a diagnosis of LQTS from a heart rhythm specialist were reclassified as having a normal rhythm after receiving a second opinion at a LQTS specialty clinic.(19)

Genetic testing for LQTS is available, but the above clinical approach remains the gold standard for diagnosis. Genetic testing may be employed as an adjunct to traditional diagnosis. Genetic testing can be employed to help diagnose cases with a borderline QTc interval, to diagnose asymptomatic family members after one family member has tested positive for the condition, and to determine specific LQTS genetic subtype. However, roughly 20-25% of inherited LQTS cases are not caused by variants in known genes, and therefore those cases cannot be identified through genetic testing.(10)

Incomplete genetic penetrance

In addition, many people who have genetic variants known to cause LQTS never develop signs or symptoms of the disorder. They do not have a prolonged QT interval, and never have episodes of syncope or torsades de pointes. Others who do have a

prolonged QT interval never develop clinical symptoms. For this reason, genetic testing alone cannot reliably identify which LQTS patients are at high risk of cardiac events.

Variants of Unknown Significance

While hundreds of variants within 17 genes have been shown to cause LQTS, not all variants found within these genes can cause the disease. Some variants are benign, and do not cause functional changes in the proteins that the gene they are contained within code for. Other variants could even have the opposing effect, and are associated with short QT syndrome.(20) It is often unclear what effect a novel rare variant in a LQTS gene might have if it is detected through genetic testing in the absence of clinical evidence for a functional effect. Techniques to predict the functional consequences of genetic variants are still in the early stages of development. As a result, genetic testing frequently identifies genetic variants of uncertain clinical significance.(21)

Lack of Epidemiological Risk Information

Rarer subtypes of LQTS are challenging to study, as traditional epidemiological approaches rely upon having a large sample of cases to compare against control groups. Because it is not possible to collect the underlying data about rare LQTS subtypes, clinicians simply cannot be as confident in determining what level of risk patients with rarer LQTS subtypes face when considering playing competitive sports, compared to the risk patients with more common subtypes face.

Institutional Challenges

Medical professional genetic competence

While the exact risk a LQTS patient faces by playing competitive sports may be challenging to assess scientifically, even under the best circumstances of sufficient data the information may be difficult to convey. Several competencies are needed by the clinician to understand and communicate the problem: an adequate understanding of cardiovascular biology, a firm grasp of medical genetics, a comprehension of statistical risk information, and the capacity to understand the role that athletics place in their patients' lives. Not all medical professionals are well equipped to understand the intricacies of the genetics, the variation in clinical presentation and the pros and cons of different treatment options involved in LQTS. A majority of primary care providers are not confident they can become experts in genetic testing, and feel uncertain about its clinical utility.(22) Conversely, genetic counselors are specialized medical professionals who are trained to understand and explain genetic information to patients, but they may not understand the cardiovascular nuances that differentiate the subtypes of LQTS or have familiarity with the range of treatment options. Furthermore, the limited number of genetic counselors mean they are not easily accessible for all patients.

Multidisciplinary approaches to treating common diseases are common. For example, it is not uncommon for diabetic patients to be seen by endocrinologists, dietitians, health educators, ophthalmologists, and other professionals in an integrated manor.(23) While LQTS patients, and patients with other rare diseases would undoubtedly benefit from this type of integrated approach, institutional networks which bring together experts from

relevant fields are often not feasible for such rare diseases, and they rarely exist outside of a small number of academic research clinics.

Problems in Communication

Even when clinicians are well equipped to understand the relevant genetic and cardiovascular information, they could have problems communicating these ideas with their patients. Not all clinicians possess the same level of interpersonal communication skills. Cultural differences between patients and providers can lead to misunderstandings. Perspectives on what it means to have a disease and what constitutes optimal treatment differ across cultures. Patients and clinicians can have different expectations on shared decision making, as well as what types of information would be most useful for treatment decisions.(24) Some patients struggle to grasp biological and genetic details which they have not been exposed to before. One study found that only about half of oncology patients understood the terms “genetic testing” and “gene”.(25) Many people struggle with the concepts of probabilities and risk. Some patients may prefer to defer to the physician’s judgment, while others will embrace being part of the decision-making process. Patients may also not have enough time to process the information they have been given, or to determine the acceptability of reconsidering major lifestyle choices.

Physician and patient goals misaligned

Clinicians can have their own opinions about what risk levels are acceptable for their patients, and these may differ from the patient’s own views. This can happen because

not all risks and benefits are considered in the same way by either the patient or provider. Patients have different levels of risk tolerance. Some LQTS patients will choose to participate in sports when others at the same risk level will decline. It is possible for both clinicians and patients to have difficulty balancing the acute risk of a deadly event with the less obvious potential harm caused by avoiding sports. Clinicians giving advice might feel compelled to caution the patient against participating in sports, believing that what they perceive as lifestyle inconveniences should be given little weight compared to the risk of serious injury, whereas a patient might view the tradeoffs differently. Clinicians also might fear they could be found legally responsible if they recommended a patient participate in sports, and that patient subsequently suffered an injury or a fatal cardiac event.

While the harms avoided by preventing a potentially fatal cardiac event by avoiding sports participation are obvious, the harms caused by being overly cautious might not be fully considered. For children especially, sports are likely to be a large component of their physical exercise. Insufficient exercise has large implications for overall lifetime health and chronic disease risk. Children who do not participate in sports may be at higher risk of obesity, diabetes, and numerous other chronic conditions. A relevant historical study evaluating the effectiveness of a screening intervention to detect heart murmurs at early age was later found to be harmful because children who had murmurs which were benign in reality were prevented from engaging in strenuous exercise.(26) Because so many individuals with LQTS mutations never experience any cardiac events, it is likely that a program which restricted exercise for LQTS patients would

harm individuals in a similar manner. Additionally, sports are part of a normal childhood for many people. Not being allowed to participate could leave a child feeling different or isolated.

Special considerations for minors

In most situations, people are understood to have full autonomy over decisions regarding their healthcare. But in circumstances where the child is under the age of 18, this principle becomes more complicated. The child's parents or guardians clearly should have a large say in any decision making, but it could be argued that schools or sports leagues where the child would be competing may have some responsibility to watch out for the child's health and best interests. Many schools and sports leagues in the United States recommend or require health screening as a condition of participation. These screening programs are not universally required and vary widely across groups. Typically, screenings involve a brief physical exam, as well as obtaining a medical history. Many medical practitioners who conduct sports physicals follow the guidelines released by the American Heart Association (AHA). In 1996 and 2007 the AHA released a 14-element recommended cardiovascular screening for young athletes. Any abnormal outcomes could be flagged, and the athletes could be referred to a cardiologist for follow-up testing.(27) However, how an institution like a school should use such screening information is a debatable question. Should schools just try to make sure those who participate are adequately informed of the risks, or should they make the decisions themselves and restrict children from playing who they deem are too high risk for injury?

Imperfect practice guidelines

Practice guidelines are developed to assist medical professionals in offering the best possible clinical care recommendations. Thousands of randomized control trials are published each year. Individual clinicians cannot reasonably sift through the enormous amounts of information of different levels of quality being published in their field. Clinical practice guidelines are created by expert medical professionals who consider the current body of medical research to aid clinicians in delivering optimal care. Practice guidelines ideally give practitioners access to the most up-to-date methods for treating their patients, with evidence concerning their efficacy.

In order to be effective and trustworthy practice guidelines need to be based on rigorous scientific evidence, be transparent in the methodology used to create them, correlate the strength of the recommendations with the strength of the available evidence, and include a multi-disciplinary approach with input from multiple stakeholders.(28) Patients with the conditions that are the focus of the guidelines are important stakeholders, and their meaningful participation in the process is critical. The goal of practice guidelines is not to create a rigid one-size-fits-all approach that treats all patients the same, but to give clinicians the best guidance to use in different clinical situations. For rare diseases, such as LQTS, practice guidelines take on special importance, because the clinicians treating LQTS patients might not be as intimately familiar with the disease as, for example, a clinician treating a common disease like diabetes. Practice guidelines can be developed by different organizations, and there can occasionally be disagreements

between sets of guidelines on the best way to handle a given situation. Guidelines can and should be revised over time as more research data becomes available, or improved treatments are developed. Although practice guidelines do not carry the force of law, physicians who stray from practice guidelines might also worry about opening themselves up to malpractice lawsuits if their patient is injured after following advice which differed from the guidelines.(29)

In 2005, two sets of practice guidelines dealing with competitive sports participation for athletes with LQTS were published from the European Society of Cardiology (ESC) and the 36th Bethesda Conference of the American College of Cardiology (ACC). (30, 31) The Bethesda guidelines recommended that patients who have either 1) exhibited LQTS symptoms (torsades de pointes, syncope, seizure, etc.), 2) have a corrected QT interval (QTc) above 470 ms in males or 480 ms in females, or 3) have an implantable cardioverter-defibrillator (ICD) not play most competitive sports. Lower intensity sports, such as billards, bowling, or golf are not restricted. The ESC guidelines are more restrictive than the Bethesda guidelines stating that people with a QTc interval above 440 ms in males and 460 ms in females should be disqualified from participation. Although they used different thresholds to define LQTS, both guidelines stated that individuals with LQTS should be counseled against playing competitive sports. The European standards have not changed since their original publication but in 2015 the ACC guidelines were updated to reflect new study data, as well as the ability to use genetic data in clinical care.(32) The new standards differentiate between genotypic and phenotypic diagnosis. In this update, patients who have received a positive genotypic

diagnosis, but a negative phenotypic diagnosis are not recommended to be restricted from playing competitive sports. There have been no reported instances of individuals with this type of diagnosis ever experiencing a cardiac event during sports. However, these individuals are recommended to take basic precautions such as avoiding QT prolonging drugs, staying hydrated, and making sure an external defibrillator is on hand. Individuals who have had LQTS diagnosed on an ECG can still participate in sports if they undergo treatment (either surgical or by taking beta-blockers) and have been otherwise asymptomatic for 3 months. The guideline recommends that competitive swimming for individuals with phenotypically diagnosed LQT1 should not be allowed.

The data which initiated the change in guidelines came from a study which tracked 130 patients with LQTS who chose to remain in competitive sports after undergoing a medical evaluation at the Mayo Clinic.(33, 34) The study participants were seen between July 2000 and November 2010, and were between the ages of 6 to 40. Of the group studied 70 (54%) violated the strict European guidelines but were within the Bethesda guidelines. The remaining 60 athletes (46%) were participating against both the European and Bethesda guidelines. 20 participated with an ICD. Only 1 athlete experienced a sports-related cardiac event during the study period: a 9-year-old boy with LQT1, an extreme QT interval length (QTc>550 ms), and a previous history of cardiac arrest. The boy received shocks from the ICD which ended the ventricular fibrillation. The episode also occurred at a time when the patient was not adhering to his prescribed beta blocker therapy. With follow-up of more than 650 athlete-years, the

rate of cardiac events per athlete year was calculated to be 1 event for every 331 athlete-years (95% CI 1 in 92 to 1 in 2763 athlete-years).

Conclusion

Even for a Mendelian genetic condition like LQTS, there are many barriers to the delivery of optimal precision medicine. While some advancements in scientific knowledge may bring to light more accurate risk information for some LQTS patients, the realities of research on very rare conditions means for others this data might not be forthcoming. Investments in research to better predict the functional consequences of rare variants are needed, since traditional epidemiological techniques are less useful for these conditions. Improvements in training medical professionals in genetics, or to utilize more genetic counselors who are deeply trained to handle this information, may also improve care, but many patients with rare diseases will lack access to specialty clinicians in all the fields relevant to their disease management.

In light of these realities, the role of clinical practice guidelines is crucial, as well as a recognition of their limitations. When developed in a rigorous fashion, following accepted standards practice guidelines provide the best available advice to clinicians and patients.(28) However, if the evidence is limited, guidelines will be similarly limited. Practice guidelines should thus be revised as new evidence comes forward, and transparent in areas where the evidence behind the recommendations is weaker. Recent advances in genetic technology, and the option to utilize genetic data in treatment present opportunities to revisit guidelines. Guidelines for rare Mendelian

conditions should use a precision medicine approach when possible by reflecting the improved risk classification and treatment options available through genetic testing. Guideline creators should also take care to better reflect the choices that truly matter to patients. For many individuals with LQTS, participation in athletics is very important, and the noble goal of preventing dangerous episodes of torsades de pointes seems to have been given too much weight compared to the real harms caused to these individuals, both physically and emotionally, by an overly cautious approach. Guidelines could be improved if they provide clinicians with a framework to talk with their patients and learn how important potentially risky lifestyle activities are to them. Ideally, more patients will be able to live the lives they find most fulfilling, while still minimizing dangerous events.

The 2005 ACC guidelines, and the changes following the availability of new evidence illustrate the significance of these dynamics. The 2005 guideline was broadly prohibitive of athletic activity for patients with LQTS; the 2015 update was more permissive because new evidence demonstrated the low risk for many patients. It is worth noting that this evidence, which played a key role in modifying the guideline, became available for two reasons: (1) a number of patients chose not to comply with the guidelines, presumably reflecting the importance they placed on athletic participation; and (2) clinicians at the Mayo Clinic observed the noncompliance and sought to evaluate its effects systematically.

The 2015 update to the ACC guidelines for sports participation represents an improvement upon the 2005 recommendations. It differentiates between genotypic and

phenotypic positive LQTS cases, recognizes that different subtypes of LQTS convey different levels of risk, and scales back the overly stringent recommendation that all patients with LQTS should be cautioned against participating in sports. However, the updated guidelines are still written from an arguably myopic perspective that the only treatment goal should be the prevention of cardiac events. Clinicians also need to seriously consider the lifelong health benefits sports can potentially provide, and that patients who give up playing sports could be more likely to develop obesity and other health problems later in life. Perhaps more importantly, guideline development should be informed by the patients' own goals, values, and perspectives. The fact that many LQTS patients in the Mayo Clinic study chose to participate in sports against the medical advice of clinicians demonstrates just how important sports was to their lifestyle.

Making stronger efforts to include stakeholder input in the development of guidelines is needed, to better align clinician and patient goals. Patient perspectives should also be more thoughtfully included in the process of updating evidence-based guidelines. Because patients can have different goals in their treatment than clinicians, patients might value certain types of evidence and studies differently than clinicians. For example, if clinicians have a goal of minimizing episodes of torsades de pointes at all costs, but patients have a goal of minimizing risk while still participating in sports, the underlying research that is considered in guideline development might be different when informed by patient perspectives. If patient values help guide what research is

conducted and funded, the evidence is likely to inform treatment improvements in the ways patients most care about.

References

1. Wetterstrand KA. The Cost of Sequencing a Human Genome Genome.gov: National Human Genome Research Institute; 2020 [cited 2021 October 15]. Available from: https://www.genome.gov/about-genomics/fact-sheets/Sequencing-Human-Genome-cost.

2. Mahajan A, Go MJ, Zhang W, Below JE, Gaulton KJ, Ferreira T, Horikoshi M, Johnson AD, Ng MC, Prokopenko I. Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. Nature genetics. 2014;46(3):234-44.

3. Burke W, Trinidad SB, Schenck D. Can precision medicine reduce the burden of diabetes? Ethnicity & disease. 2019;29(Suppl 3):669.

4. Claussnitzer M, Cho JH, Collins R, Cox NJ, Dermitzakis ET, Hurles ME, Kathiresan S, Kenny EE, Lindgren CM, MacArthur DG. A brief history of human disease genetics. Nature. 2020;577(7789):179-89.

5. Rafeeq MM, Murad HAS. Cystic fibrosis: current therapeutic targets and future approaches. Journal of translational medicine. 2017;15(1):1-9.

6. Brewer KR, Kuenze G, Vanoye CG, George Jr AL, Meiler J, Sanders CR. Structures illuminate cardiac ion channel functions in health and in long QT syndrome. Frontiers in Pharmacology. 2020;11:550.

7. Schnell F, Behar N, Carré F. Long-QT syndrome and competitive sports. Arrhythmia & electrophysiology review. 2018;7(3):187.

8. Shimizu W, Noda T, Takaki H, Kurita T, Nagaya N, Satomi K, Suyama K, Aihara N, Kamakura S, Sunagawa K. Epinephrine unmasks latent mutation carriers with LQT1 form of congenital long-QT syndrome. Journal of the American College of Cardiology. 2003;41(4):633-42.

9. Nakano Y, Shimizu W. Genetics of long-QT syndrome. Journal of human genetics. 2016;61(1):51-5.

10. Kapa S, Tester DJ, Salisbury BA, Harris-Kerr C, Pungliya MS, Alders M, Wilde AA, Ackerman MJ. Genetic testing for Long QT Syndrome-Distinguishing pathogenic mutations from benign variants. Circulation. 2009;120(18):1752.

11. Schwartz PJ, Priori SG, Spazzolini C, Moss AJ, Vincent GM, Napolitano C, Denjoy I, Guicheney P, Breithardt Gn, Keating MT. Genotype-phenotype correlation in the long-QT syndrome: gene-specific triggers for life-threatening arrhythmias. Circulation. 2001;103(1):89-95.

12. Council NR. Committee on A Framework for Developing a New Taxonomy of Disease. Toward precision medicine: building a knowledge network for biomedical research and a new taxonomy of disease. Washington, DC: National Academies Press; 2011.

13. Vella SA, Cliff DP, Magee CA, Okely AD. Sports participation and parent-reported health-related quality of life in children: longitudinal associations. The Journal of pediatrics. 2014;164(6):1469-74.

14. Oja P, Titze S, Kokko S, Kujala UM, Heinonen A, Kelly P, Koski P, Foster C. Health benefits of different sport disciplines for adults: systematic review of observational and intervention studies with meta-analysis. British journal of sports medicine. 2015;49(7):434-40.

15. Puetz TW. Physical activity and feelings of energy and fatigue. Sports medicine. 2006;36(9):767-80.

16. Eime RM, Young JA, Harvey JT, Charity MJ, Payne WR. A systematic review of the psychological and social benefits of participation in sport for children and adolescents: informing development of a conceptual model of health through sport. International journal of behavioral nutrition and physical activity. 2013;10(1):1-21.

17. Khan IA. Long QT syndrome: diagnosis and management. American heart journal. 2002;143(1):7-14.

18. Schwartz PJ, Stramba-Badiale M, Crotti L, Pedrazzini M, Besana A, Bosi G, Gabbarini F, Goulene K, Insolia R, Mannarino S. Prevalence of the congenital long-QT syndrome. Circulation. 2009;120(18):1761-7.

19. Taggart NW, Haglund CM, Tester DJ, Ackerman MJ. Diagnostic miscues in congenital long-QT syndrome. Circulation. 2007;115(20):2613-20.

20. Schimpf R, Wolpert C, Gaita F, Giustetto C, Borggrefe M. Short QT syndrome. Cardiovascular research. 2005;67(3):357-66.

21. Ackerman MJ. Genetic purgatory and the cardiac channelopathies: exposing the variants of uncertain/unknown significance issue. Heart rhythm. 2015;12(11):2325-31.

22. Schmidlen TJ, Scheinfeldt L, Zhaoyang R, Kasper R, Sweet K, Gordon ES, Keller M, Stack C, Gharani N, Daly MB. Genetic knowledge among participants in the Coriell Personalized Medicine Collaborative. Journal of genetic counseling. 2016;25(2):385-94.

23. Bratcher CR, Bello E. Traditional or centralized models of diabetes care: the multidisciplinary diabetes team approach. The Journal of family practice. 2011;60(11 Suppl):S6-11.

24. Charles C, Gafni A, Whelan T, O'Brien MA. Cultural influences on the physician–patient encounter: the case of shared treatment decision-making. Patient education and counseling. 2006;63(3):262-7.

25. Blee SM, Shah RP, Pinheiro AP, Switchenko J, Dixon M, Owonikoko TK, Hill CE, Szabo SM, Pentz RD. Physician Communication and Patient Understanding of Molecular Testing Terminology. The Oncologist. 2021.

26. Association CM. Morbidity of cardiac" non-disease" in school children. CMAJ. 1967;97(24):1490-1.

27. Maron BJ, Friedman RA, Kligfield P, Levine BD, Viskin S, Chaitman BR, Okin PM, Saul JP, Salberg L, Van Hare GF. Assessment of the 12-lead ECG as a screening test for detection of cardiovascular disease in healthy general populations of young people (12–25 years of age) a scientific statement from the American Heart Association and the American College of Cardiology. Circulation. 2014;130(15):1303-34.

28. Steinberg E, Greenfield S, Wolman DM, Mancher M, Graham R. Clinical practice guidelines we can trust: National Academies Press; 2011.

29. Mackey TK, Liang BA. The role of practice guidelines in medical malpractice litigation. AMA Journal of Ethics. 2011;13(1):36-41.

30. Pelliccia A, Fagard R, Bjørnstad HH, Anastassakis A, Arbustini E, Assanelli D, Biffi A, Borjesson M, Carrè F, Corrado D. Recommendations for competitive sports participation in athletes with cardiovascular diseaseA consensus document from the Study Group of Sports Cardiology of the Working Group of Cardiac Rehabilitation and Exercise Physiology and the Working Group of Myocardial and Pericardial Diseases of the European Society of Cardiology. European heart journal. 2005;26(14):1422-45.

31. Zipes DP, Ackerman MJ, Estes NM, Grant AO, Myerburg RJ, Van Hare G. Task force 7: arrhythmias. Journal of the American College of Cardiology. 2005;45(8):1354-63.

32. Ackerman MJ, Zipes DP, Kovacs RJ, Maron BJ. Eligibility and disqualification recommendations for competitive athletes with cardiovascular abnormalities: task force 10: the cardiac channelopathies: a scientific statement from the American Heart Association and American College of Cardiology. Circulation. 2015;132(22):e326-e9.

33. Johnson JN, Ackerman MJ. Return to play? Athletes with congenital long QT syndrome. British journal of sports medicine. 2013;47(1):28-33.

34. Johnson JN, Ackerman MJ. Competitive sports participation in athletes with congenital long QT syndrome. Jama. 2012;308(8):764-5.