## Original Article

# Evaluating Polygenic Risk Scores in "Lone" Atrial Fibrillation

Julieta Lazarte, MSc,[a] Jacqueline S. Dron, BSc,[a] Adam D. McIntyre, BSc,[a] Allan C. Skanes, MD,[b]
Lorne J. Gula, MD,[b] Anthony S. Tang, MD,[b] Rafik Tadros, MD, PhD,[d]
Zachary W. Laksman, MD, MSc,[c] Robert A. Hegele, MD,[a] and Jason D. Roberts, MD, MAS[b]

[a] Department of Medicine and Robarts Research Institute, Schulich School of Medicine and Dentistry, Western University, London, Ontario, Canada
[b] Section of Cardiac Electrophysiology, Division of Cardiology, Department of Medicine, Western University, London, Ontario, Canada
[c] Department of Medicine, Faculty of Medicine, University of British Columbia, Vancouver, British Columbia, Canada
[d] Division of Cardiology, Montréal Heart Institute, University of Montréal, Montréal, Québec, Canada

## ABSTRACT

**Background:** Polygenic scores incorporating varying numbers of single nucleotide polymorphisms (SNPs) have been demonstrated to exert a prominent role in atrial fibrillation (AF). We sought to compare the relative discriminatory capacities of 2 previously validated polygenic scores in "lone" AF.

**Methods:** A total of 186 lone AF cases of European ancestry underwent SNP genotyping. A genome-wide polygenic score (GPS) and polygenic risk score (PRS) involving 6,730,541 and 1168 SNPs, respectively, were calculated for 186 cases and 423 controls of

## RÉSUMÉ

**Contexte :** Il a été démontré que des scores polygéniques intégrant un nombre variable de polymorphismes mononucléotidiques (PMN) jouent un rôle important en ce qui concerne la fibrillation auriculaire (FA). Nous avons comparé le potentiel discriminatoire relatif de deux scores polygéniques déjà validés dans la FA idiopathique.

**Méthodologie :** Au total, 186 sujets d'ascendance européenne atteints de FA idiopathique ont été soumis à un génotypage des PMN. Un score polygénique génomique (SPG) et un score de risque polygénique (SRP) comprenant respectivement 6 730 541 et 1 168 PMN ont été calculés

Rare and common genetic variants have been shown to affect the risk of developing atrial fibrillation (AF).[1] Notably, 22% of AF heritability has been suggested to be explained primarily by the cumulative effects of common variants or single nucleotide polymorphisms (SNPs), highlighting the significant polygenic burden of AF, with a minimal, yet important, contribution of rare variants with large effect size such as those found in the structural gene titin (TTN).[1-3] To date, genome-wide association studies (GWAS) have identified more than 100 SNPs that significantly associate with AF susceptibility.[4,5] In this context, polygenic scores that incorporate multiple small-effect SNPs and identify a proportion of subjects from the general population at an increased risk for AF, are viewed as harbouring great clinical utility, given the potential to

enhance screening and prevention therapies.[6] Notably, however, three-quarters of AF genetic risk has been identified to be driven by common variants in regions that have yet to satisfy the stringent thresholds for GWAS statistical significance.[2] Recognizing that sizeable amounts of genomic data may not be adequately leveraged at present, contemporary genetic risk scores have begun to include loci that have yet to reach GWAS levels of significance, in an effort to better estimate AF susceptibility.[1,6,7]

"Lone" AF represents a subtype of the arrhythmia that develops in the absence of identifiable clinical risk factors.[8] Although previously criticized because of its heterogeneous definitions and our improved ability to identify predisposing clinical risk factors, such cases are considered to harbour a greater genetic contribution relative to more common forms of the arrhythmia that develop in the setting of structural heart disease or other established risk factors.[9] We previously explored the role of genetic risk scores in a lone AF cohort and a locally procured healthy control set.[10] We found that genetic risk scores developed by both Weng et al. (2018; an ~1000 SNP polygenic risk score [PRS]) and Khera et al. (2018; ~6 million genome-wide polygenic risk score [GPS]), identified a significant portion of patients with lone AF with an elevated polygenic score compared with healthy controls.[6,7,10] However, no difference was observed when comparing the

European ancestry from the 1000 Genomes (1KG) Project. The distribution of the polygenic scores was compared between the cases and controls and their discriminatory capacities were evaluated using receiver operating characteristic (ROC) curves.

**Results:** A total of 34.4% of patients with lone AF had GPS scores greater than the top 10th percentile of 1KG controls, corresponding to a 4.64-fold increased odds (95% confidence interval [CI], 2.99-7.18; $P < 0.001$) for AF. A PRS score in the top 10th percentile of 1KG controls was observed in 26.3% of cases, which equated to a 3.16-fold increased odds (95% CI, 2.01-4.98; $P < 0.001$) for AF. Comparison of C-statistics from ROC curves indicated improved discriminatory capacity of the GPS (0.76) relative to the PRS (0.70) ($P = 0.002$).

**Conclusions:** Our study evaluating 2 polygenic scores for AF suggests that the GPS, containing more than 6.7 million SNPs, exhibits an improved discriminatory capacity in lone AF compared with a PRS possessing 1168 SNPs. Our findings suggest that genetic risk scores for AF that maximally leverage genomic data may provide improved predictive power.

pour les 186 sujets et pour 423 témoins d'ascendance européenne dont les données sont tirées du projet 1000 Genomes (1KG). Les distributions des scores polygéniques des sujets et des témoins ont été comparées, et leur potentiel discriminatoire a été évalué au moyen des courbes caractéristiques de la performance d'un test (courbes ROC, de l'anglais *Receiver Operating Characteristic*).

**Résultats :** Au total, 34,4 % des patients atteints de FA idiopathique avaient un SPG supérieur à celui des témoins du 10$^e$ centile supérieur du projet 1KG, ce qui représente une probabilité de FA 4,64 fois plus élevée (intervalle de confiance [IC] à 95 % : 2,99 à 7,18; $p < 0,001$). Un SRP situé dans le 10$^e$ centile supérieur des témoins du projet 1KG a été observé chez 26,3 % des patients atteints de FA, soit une probabilité de FA 3,16 fois plus élevée (IC à 95 % : 2,01 à 4,98; $p < 0,001$). Les résultats de la comparaison des statistiques C des courbes ROC indiquent que le SPG (0,76) a un potentiel discriminatoire supérieur à celui du SRP (0,70) ($p = 0,002$).

**Conclusions :** Les résultats de notre étude de deux scores polygéniques relatifs à la FA indiquent que le potentiel discriminatoire du SPG, qui comprend plus de 6,7 millions de PMN, pour prédire une FA idiopathique est supérieur à celui du SRP, qui comprend 1 168 PMN. Ces résultats indiquent que les scores de risque génétique de FA qui exploitent pleinement les données génomiques pourraient avoir un pouvoir prédictif supérieur.

discriminatory capacity of each score, perhaps secondary to inadequate statistical power. In an effort to address this limitation and attempt to discern which harbours the greatest predictive power—a concept that has yet to be evaluated in the AF literature—we sought to evaluate the behaviour of these 2 validated risk scores in lone AF, using a large publicly available dataset as a control cohort.

## Methods

### AF study cohort

Patients referred for AF management at the London Health Sciences Centre, London, Ontario, Canada, and St. Paul's Hospital, Vancouver, British Columbia, Canada, with AF, in the absence of known clinical risk factors, before 60 years of age, defined as lone AF, were recruited to the study. At least 1 episode of electrocardiographically documented AF, characterized by erratic atrial activity without distinct P waves and irregularly irregular QRS intervals lasting > 30 seconds, was required per patient. Exclusion criteria consisted of known risk factors for AF, including hypertension, coronary artery disease, left-ventricular ejection fraction < 50% or a history of clinical heart failure, moderate to severe valvular heart disease, hyperthyroidism, obstructive sleep apnea, and presence of inherited cardiomyopathy. All participants had a clinical history, physical examination, 12-lead electrocardiogram (ECG), and echocardiogram. A positive family history of AF was defined as presence of the arrhythmia in a first- or second-degree relative.

### Control cohort

The control cohort was derived from the 1000 Genomes (1KG) Project, a publicly available multiancestry cohort of 1756 persons above the age of 18 who self-reported as healthy.[11] SNP genotyping of the 1KG cohort was performed using the Illumina Omni 2.5 M DNA microarray (Illumina, San Diego, CA). Analyses were restricted to 423 persons of European (non-Finnish) ancestry following principal component analysis, described as follows.[11,12]

### DNA preparation and microarray genotyping of AF cases

Genomic DNA for lone AF cases was isolated using the Puregene DNA Blood Kit (Gentra Systems, Qiagen Inc., Mississauga, ON). Microarray analysis was performed with Infinium Global Screening Array-24 v2.0 (Illumina) at Genome Québec, Montréal, Québec, Canada. GenomeStudio software (Illumina) was used to retrieve and export the microarray data to SNP & Variant Suite (SVS) v8.8.3 (Golden Helix Inc, Bozeman, MT). To improve genotyping accuracy, data points were filtered if they had a GenCall score cutoff < 0.15. Microarray data was further cleaned by filtering samples with a < 95% rate of autosomal SNP calls over the total number of SNP calls in the dataset to avoid inappropriate results from faulty genotyping calls (n = 2).[13] Using X chromosome heterozygosity, samples were removed if positive for sex discordance between clinical data and genotype information (n = 4). Linkage disequilibrium (LD) pruning was applied to all autosomes to prepare the data for identity by descent estimation analysis; samples were removed if estimated Pi-hat (cryptic relatedness) for a sample pair was > = 0.5 consistent with first-degree relatives. One father and son pair was detected, and the father was removed from further analysis. Deviation from Hardy-Weinberg equilibrium was not used as a method to filter SNPs, given that it was unclear if various assumptions were met for both the lone AF and 1KG cohorts, including random mating and sufficiently large population sizes.[12]

Lazarte et al.
Polygenic Risk Scores in "Lone" AF

753

## Ancestry correction: European (non-Finnish) subgroup

Ancestry inference using principal component analysis was performed on SNP & Variant Suite (SVS) v8.8.3 (Golden Helix Inc) for the lone AF cases and controls post-LD pruning using EIGENSTRAT.[14] Among 5 formulated eigenvalues, the top 3 explained the majority of the stratification (Supplemental Fig. S1). Around the 1KG European (non-Finnish) population cluster, a centroid was mathematically identified, and any case sample that fell outside the 1.5 interquartile range (IQR) was excluded from the analysis, as it was deemed outside the European (non-Finnish) population cluster. Among a total of 240 lone AF cases that had undergone DNA microarray analysis, 54 were excluded on this basis.

## Imputation

Genotype imputation of DNA microarray data was performed for both cohorts to increase genomic coverage using the Michigan Imputation Server.[15] Imputation was performed using the Minimac4 1.2.4 imputation algorithm with Haplotype Reference Consortium r1.1 (cases) and 1KG Phase 3 v5 (1KG controls) reference set. Only SNPs that were genotyped or imputed with $r^2 > 0.3$ were used for score calculation.

## Polygenic score calculation

Two validated AF polygenic scores were calculated: a GPS developed by Khera et al. and a PRS developed by Weng et al. (both in 2018), hereafter referred to as "GPS" and "PRS," respectively.[6,7] The GPS was derived by Khera and colleagues, using the LDPred algorithm and association data from a previous genome-wide association study for AF, using separate testing and validation datasets from the UK Biobank.[6,16] The PRS developed by Weng and colleagues also used association data from the same AF genome-wide association study, however, used pruning and thresholding at various tuning parameters for its derivation.[7,16] A total of 30 candidate scores were developed and tested within the UK Biobank dataset and the optimal one was identified on the basis of its goodness-of-fit in accordance with the Akaike's Information Criterion.[7] Subjects were scored by counting genetic dosages of imputed variants using the –score option in PLINK2.0 and Wrapper Python script.[17] For each variant, the number of risk alleles present is multiplied by its respective weight, and the products for each variant are added to generate the final score. In total, 5,978,070 of 6,730,541 variants (88.82%) were available for the GPS and 872 of 1168 variants (74.66%) for the PRS in both cohorts.

## Statistical analysis

GPS/PRS distributions were assessed for normality using the D'Agostino-Pearson omnibus K2 test. Odds ratios (ORs) were calculated by comparing proportions of subjects in the top 10, 5, and 1 percentiles of the GPS/PRS using 2-by-2 contingency tables with Fisher's exact test. The performance of the GPS and PRS for discerning AF cases vs controls was assessed using receiver operating characteristic (ROC) curves with 1KG controls as the reference and compared in R version 3.5.1 (R Core Team, 2018) with pROC (version 1.7.2).[18] Impact of a high GPS/PRS score (in the top 10th percentile) on age at diagnosis (divided into quartiles: < 37, 37-45, 46-51, > 51) among lone AF cases was assessed using the $\chi^2$

**Table 1.** Clinical Characteristics of the "Lone" AF Cohort

| Clinical variable | "Lone" AF cohort n = 186 |
|---|---|
| Age at diagnosis (years) | 44.26 ± 9.85 |
| Sex (male) | 151 (81.2%) |
| Body mass index (kg/m$^2$)* | 27.39 ± 3.92 |
| Type of AF at diagnosis | |
| Paroxysmal | 165 (89.7%) |
| Persistent | 19 (10.3%) |
| Family history of AF* | 57 (30.8%) |
| ECG values | |
| PR-interval* | 166.73 ± 25.95 |
| QRS duration | 95.27 ± 14.74 |
| QTc | 420.60 ± 26.70 |
| Left-atrial diameter (cm)* | 3.96 ± 0.55 |

Data are n (%) or mean ± standard deviation.

AF, atrial fibrillation; ECG, electrocardiogram.

*Values for body mass index, family history of AF, PR-interval, and left-atrial diameter were missing for 1, 1, 2, and 2 cases, respectively.

test and the $\chi^2$ test for trend. Evaluation for a different likelihood of possessing a high GPS/PRS score (in the top 10th percentile) in lone AF cases by sex and by body mass index (BMI) > 30 kg/m$^2$ was assessed using 2-by-2 contingency tables with Fisher's exact test. Unless otherwise stated, all statistical analyses were conducted using GraphPad Prism 8 for Windows (version 8.3.1; GraphPad Software, San Diego California). Statistical significance was defined as $P < 0.05$.

## Results

### Characteristics of lone AF cases

Demographic and clinical characteristics of 186 lone AF cases are described in Table 1. The mean age at AF diagnosis was $44.5 \pm 9.8$ years, and 157 (80.5%) patients were male. Twenty-two (11.3%) study participants had persistent AF at the time of diagnosis, whereas the remainder presented with paroxysmal AF. The mean left-atrial diameter on echocardiography at the time of presentation was $4.0 \pm 0.6$ cm.

### Polygenic scores

The distributions of the GPS (A) and PRS (B) across the cases and 1KG controls are shown in Figure 1. The GPS had a Gaussian distribution across the 2 groups, whereas the PRS was skewed to higher polygenic scores in the 1KG controls failing the normality test ($P = 0.02$).

In total, 34.4% (64 of 186) of patients with lone AF were in the top 10th percentile of the GPS distribution. Presence of a GPS score within the top 10th percentile of 1KG controls was associated with 4.64-fold increased odds (95% confidence interval [CI], 2.99-7.18; $P < 0.001$) for AF (Fig. 1A, Table 2). The odds of a score within the top 5% and 1% of the GPS distribution within 1KG controls were 4.22-fold (95% CI, 2.40-7.57; $P < 0.0001$) and 3.76-fold (95% CI, 1.18-10.30; $P = 0.02$) more likely among lone AF cases, respectively (Table 2).

Inheriting a PRS score in the top 10th percentile was seen in 26.3% (49 of 186) of patients with lone AF, which conferred a 3.16-fold increased odds (95% CI, 2.01-4.98; $P < 0.001$) for AF relative to 1KG controls (Fig. 1B, Table 2). The odds of a PRS score within the top 5% and 1%
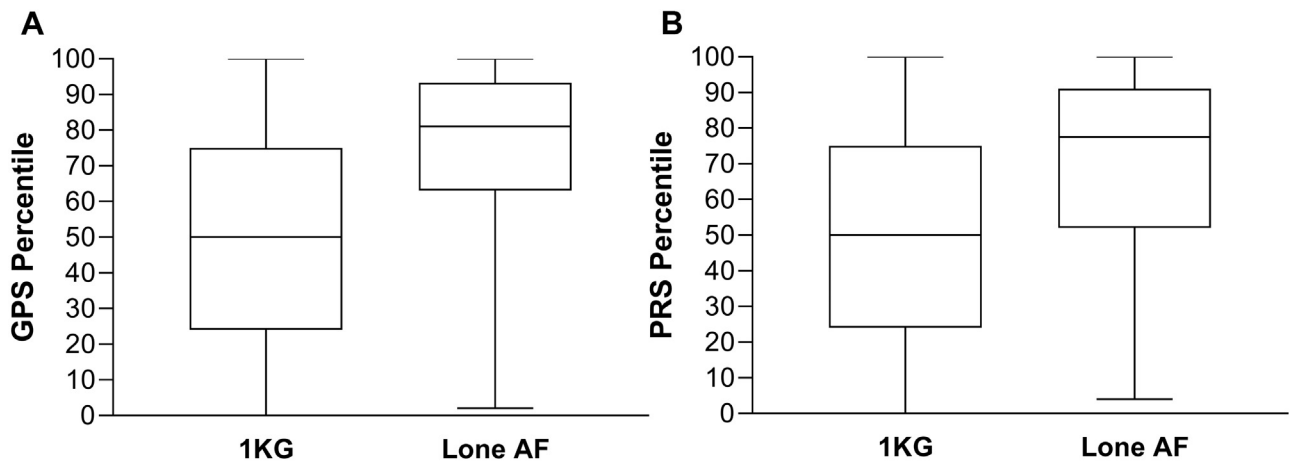
**Figure 1.** Distribution of the GPS (**A**) and PRS (**B**) percentiles among "lone" AF cases vs 1KG controls. For each boxplot, the horizontal lines represent the following: middle line = the median; the top and bottom line = interquartile range; and the whiskers = the maximum and minimum values within each group. AF, atrial fibrillation; GPS, genome-wide polygenic risk score; PRS, polygenic risk score; 1KG, 1000 genomes.

distribution were 3.10-fold (95% CI, 1.71-5.49; $P = 0.0002$) and 7.33-fold (95% CI, 2.61-18.55; $P < 0.0001$) more likely among lone AF cases relative to 1KG controls, respectively (Table 2).

### Discriminative capacity of GPS and PRS

The ability of a high polygenic score to differentiate between lone AF cases and 1KG controls was assessed using ROC curve analysis. The C-statistic for the GPS was 0.76 (95% CI, 0.72-0.80) in comparison with a value of 0.70 (95% CI, 0.65-0.75) for the PRS. The GPS was noted to be superior relative to the PRS in discriminating lone AF cases vs 1KG controls ($P = 0.002$) (Fig. 2).

### Impact of age and sex on GPS and PRS

The likelihood of a high polygenic score did not differ on the basis of age at diagnosis ($P = 0.65$ [GPS], $P = 0.98$ [PRS]) or sex ($P = 0.17$ [GPS], $P = 0.56$ [PRS]) or BMI ($P = 0.55$ [GPS], $P = 0.30$ [PRS]) in the lone AF cohort and no statistical trend was identified for age ($P = 0.61$ [GPS], $P = 0.71$ [PRS]).

### Discussion

Our study evaluating the performance characteristics of a GPS containing ∼ 6 million SNPs and a PRS containing ∼ 1000 SNPs demonstrated a 6% improved discriminatory capacity of the GPS over the PRS for distinguishing lone AF cases from healthy controls. To our knowledge, this represents the first time a more comprehensive GPS has been found to exhibit superior predictive performance relative to a validated, but more parsimonious, polygenic score in the setting of AF. These findings highlight the value of maximizing the depth of genetic detail incorporated into polygenic scores designed to identify persons at increased risk of developing or possessing AF. Notably, high scores for both the GPS and PRS, defined as greater than the top 10th percentile for the control population, were present in upward of 25% of lone AF cases, highlighting their potential relevance to a large proportion of individuals affected by the arrhythmia.

Genome-wide polygenic scores, capturing millions of common variants from the entire genome, have previously been shown to provide a superior capacity to smaller polygenic scores in discerning affected patients from healthy controls across various disease entities.[6,19] For example, an ∼ 6 million-SNP score outperformed genetic risk scores possessing 50 and 49 thousand SNPs in head-to-head comparisons for prediction of coronary artery disease risk.[6] This is a notable departure from the initial strategy within the genetics field to keep polygenic scores with as few carefully selected SNPs as possible. Indeed, the incrementally improved performance of the ∼ 6 million SNP GPS relative to the ∼ 1000 SNP PRS in our lone AF cohort is novel in the field but consistent with previous work in other disease entities.

**Table 2.** Proportion of "lone" AF cases and odds of possessing a GPS/PRS in the Top 10, 5, and 1 Percentiles*

| High GPS definition | "Lone" AF cases | 1KG controls | Odds ratio | 95% CI | P value |
|---|---|---|---|---|---|
| Top 10% of distribution | 64 (34.4%) | 43 (10.2%) | 4.64 | 2.99-7.18 | < 0.0001 |
| Top 5% of distribution | 35 (18.8%) | 22 (5.2%) | 4.22 | 2.40-7.57 | < 0.0001 |
| Top 1% of distribution | 8 (4.3%) | 5 (1.2%) | 3.76 | 1.18-10.30 | 0.02 |
| High PRS definition | | | | | |
| Top 10% of distribution | 49 (26.3%) | 43 (10.2%) | 3.16 | 2.00-4.96 | < 0.0001 |
| Top 5% of distribution | 27 (14.5%) | 22 (5.2%) | 3.10 | 1.71-5.49 | 0.0002 |
| Top 1% of distribution | 15 (8.1%) | 5 (1.2%) | 7.33 | 2.61-18.55 | < 0.0001 |

Data are n (%).

AF, atrial fibrillation; CI, confidence interval; GPS, genome-wide polygenic risk score containing 5,978,070 single nucleotide polymorphisms; PRS, polygenic risk score containing 872 single nucleotide polymorphisms; 1KG, 1000 genomes.

* Percentile of the polygenic score corresponds to the distribution of the 1KG controls.
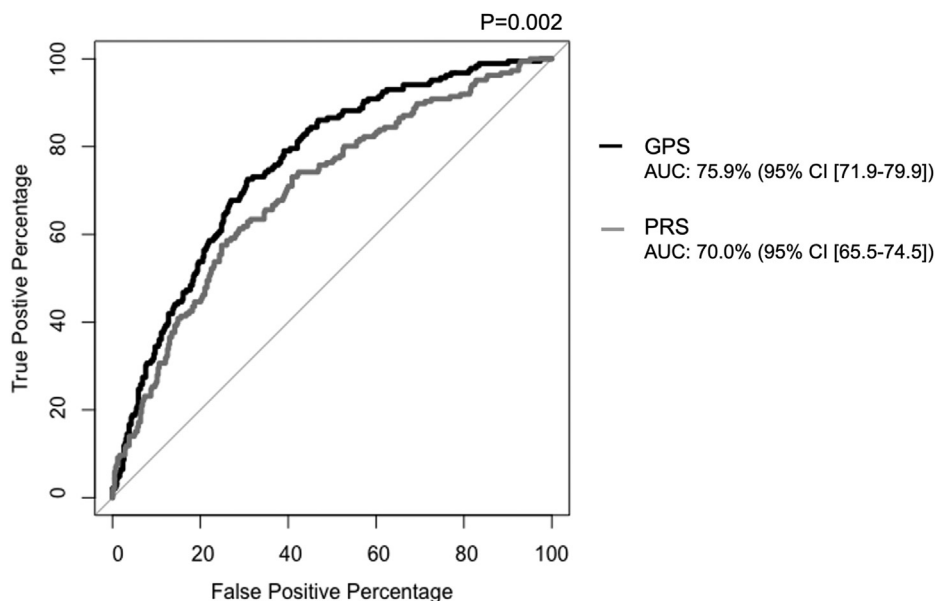
Lazarte et al.
Polygenic Risk Scores in "Lone" AF

755

**Figure 2.** Receiver operating characteristic curves for the GPS (**black line**) and PRS (**grey line**) with the 1KG control distribution as the reference. The area under the curve for the GPS (75.9%; 95% CI, 71.9-79.9) was consistent with improved discriminatory capacity relative to the PRS (70.0%; 95% CI, 65.5-74.5; $P = 0.002$). AUC, area under the curve; CI, confidence interval; GPS, genome-wide polygenic risk score; PRS, polygenic risk score.

We previously investigated the role of these genetic risk scores in lone AF using a locally sourced control set (controls = 86).[10] No difference in their discriminatory capacities for distinguishing lone AF cases from healthy controls was observed; however, the relatively small size of the control cohort may have resulted in limited the statistical power. Although use of publicly available datasets as control cohorts must be performed cautiously secondary to their generally stemming from different source populations, coupled with different genotyping platforms often being used, these drawbacks may be counterbalanced by their large size and the resulting improved statistical power provided.

For the current analysis, we believed it was reasonable to use the 1KG cohort as a control dataset, given that we were evaluating previously validated genetic risk scores rather than deriving our own or attempting to identify novel loci, which should reduce the likelihood of false positive associations. Principal component analysis was used to restrict cases and controls to a uniform genetic ancestry, which should limit potential bias secondary to cohort-selection factors and population stratification. In addition, we applied several additional measures to further minimize biases in the analysis, including filtering out data points with low accuracy, removing low call SNPs before imputation, and only keeping imputed SNPs with high quality. Finally, we also compared the distribution of scores between the 1KG cohort and our locally sourced control set (derived from the same region as our cases and genotyped with the same technology) and identified no difference in the proportion of individuals in the top 10th percentile using 2-by-2 contingency tables with Fisher's exact test (Supplemental Fig. S2).

Beyond highlighting the improved utility of genome-wide polygenic scores containing millions of SNPs relative to smaller scale genetic risk scores for lone AF, their relevance to a large proportion of AF cases further alludes to their potential clinical utility. The rapidly expanding prevalence of AF, worldwide and in Canada, has led to a major impetus to try to curb incident cases. Although evidence for prevention of AF through upstream therapies has yet to be established with randomized trials, the use of Mendelian randomization studies has served to bolster the probable causal role of certain AF risk factors, including BMI and increased thyroid activity, suggesting that intervening on these factors may prevent AF.[20-22] Genetic risk scores may enable targeted delivery of therapies to individuals at greatest likelihood to benefit: indeed, in particular to those with substantial genetic burdens.[23] Moreover, pairing polygenic scores with data from new "wearable" device technology could potentially maximize the clinical utility of both technologies and improve early detection of AF.[24] Thus, identifying individuals from the general population who are at substantial increased risk of the arrhythmia before its onset may enable effective administration of primary prevention strategies that may allow for early intervention and potentially curb incidence of AF. Hence, in this context, polygenic scores may serve as valuable clinical tools.

### Limitations

Although our study provides important insight into the value of maximizing the depth of genetic detail polygenic scores in lone AF, it has several limitations. Our study was restricted to European ancestry, partially necessitated by the PRS and GPS scores having been derived in this ancestry, coupled with the allele frequency and effect size of common variants being ancestry specific.[6,7] In this context, our findings are not anticipated to be generalizable beyond cohorts of

European ancestry, and hence additional polygenic scores will need to be developed for this purpose. Use of the 1KG dataset was pursued to provide a larger control group; however, we acknowledge the significant potential for bias secondary to different genotyping methods and population substructures. Because of these concerns, we performed principal component analysis to minimize the potential impact of ancestry and batch-effect differences and included the top 3 principal components in the outlier analysis to formulate our final cohort of 186 subjects. Although our study sample size was an important limiting factor for a contemporary investigation into the complex genetics of a common disease, the statistically significant findings for our primary hypotheses highlight that our statistical power was adequate, although insufficient power likely precluded meaningful assessment for interactions between clinical risk factors and the polygenic scores in relation to AF risk. Indeed, 1KG control participants self-reported as healthy, but it is conceivable that some had undetected AF; however, the likelihood of AF under ascertainment would not be anticipated to be affected by GPS/PRS values. The corresponding nondifferential misclassification of the outcome among controls would only serve to reduce our statistical power secondary to bias toward the null rather than resulting in spurious false positive associations. Finally, the failure of the normality test for the 1KG controls PRS score distribution may have potentially biased the discriminatory capacity of the PRS score toward the null. Although no significant difference in subjects in the top 10th percentile was encountered between 1KG and locally sourced controls (Supplemental Fig. S2), the skew toward higher scores in the 1KG controls may have diminished the PRS discriminatory capacity. Given these collective limitations, future replication in an independent lone AF cohort of European ancestry will be critical for validation of our current findings.

## Conclusions

Our study findings suggest that genome-wide polygenic scores, capturing millions of common variants from the entire genome, provide a superior discriminatory capacity compared with smaller polygenic scores in lone AF. Given their relevance to a large proportion of lone AF cases, integration of genome-wide polygenic scores into clinical practice may facilitate identification of persons at risk of developing AF, potentially leading to improved care.

## Funding Sources

## Disclosures

R.A.H. has received consulting fees from Acasti, Aegerion, Akcea/Ionis, Amgen, HLS Therapeutics, and Sanofi. The other authors have no conflicts of interest to disclose.

## References

1. Choi SH, Jurgens SJ, Weng LC, et al. Monogenic and polygenic contributions to atrial fibrillation risk: results from a national biobank. Circ Res 2020;126:200-9.

2. Weng LC, Choi SH, Klarin D, et al. Heritability of atrial fibrillation. Circ Cardiovasc Genet 2017;10.

3. Ahlberg G, Refsgaard L, Lundegaard PR, et al. Rare truncating variants in the sarcomeric protein titin associate with familial and early-onset atrial fibrillation. Nat Commun 2018;9:4316.

4. Roselli C, Chaffin MD, Weng LC, et al. Multi-ethnic genome-wide association study for atrial fibrillation. Nat Genet 2018;50:1225-33.

5. Nielsen JB, Thorolfsdottir RB, Fritsche LG, et al. Biobank-driven genomic discovery yields new insight into atrial fibrillation biology. Nat Genet 2018;50:1234-9.

6. Khera AV, Chaffin M, Aragam KG, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nat Genet 2018;50:1219-24.

7. Weng LC, Preis SR, Hulme OL, et al. Genetic predisposition, clinical risk factor burden, and lifetime risk of atrial fibrillation. Circulation 2018;137:1027-38.

8. Fuster V, Ryden LE, Cannom DS, et al. ACC/AHA/ESC 2006 guidelines for the management of patients with atrial fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the European Society of Cardiology Committee for Practice Guidelines (Writing Committee to Revise the 2001 Guidelines for the Management of Patients With Atrial Fibrillation): developed in collaboration with the European Heart Rhythm Association and the Heart Rhythm Society. Circulation 2006;114:e257-354.

9. Wyse DG, Van Gelder IC, Ellinor PT, et al. Lone atrial fibrillation: does it exist? J Am Coll Cardiol 2014;63:1715-23.

10. Lazarte JD, McIntyre JS, Skanes AD, Gula AC, et al. The role of common genetic variation in "lone" atrial fibrillation. Circ Genom Precis Med 2021:e003179.

11. Genomes Project C, Auton A, Brooks LD, et al. A global reference for human genetic variation. Nature 2015;526:68-74.

12. Roslin NMLW, Paterson AD, Strug LJ. Quality control analysis of the 1000 Genomes Project Omni2.5 genotypes (Abstract/Program #576/F). Vancouver, British Columbia, Canada: 66th Annual Meeting of The American Society of Human Genetics, 2016.

13. Miyagawa T, Nishida N, Ohashi J, et al. Appropriate data cleaning methods for genome-wide association study. J Hum Genet 2008;53:886-93.

14. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 2006;38:904-9.

15. Das S, Forer L, Schonherr S, et al. Next-generation genotype imputation service and methods. Nat Genet 2016;48:1284-7.

16. Christophersen IE, Rienstra M, Roselli C, et al. Large-scale analyses of common and rare variants identify 12 new loci associated with atrial fibrillation. Nat Genet 2017;49:946-52.

Lazarte et al.
Polygenic Risk Scores in "Lone" AF

757

17. Khera AV, Chaffin M, Zekavat SM, et al. Whole-genome sequencing to characterize monogenic and polygenic contributions in patients hospitalized with early-onset myocardial infarction. Circulation 2019;139:1593-602.

18. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011;12:77.

19. Dron JS, Hegele RA. The evolution of genetic-based risk scores for lipids and cardiovascular disease. Curr Opin Lipidol 2019;30:71-81.

20. Chatterjee NA, Giulianini F, Geelhoed B, et al. Genetic obesity and the risk of atrial fibrillation: causal estimates from Mendelian randomization. Circulation 2017;135:741-54.

21. Salem JE, Shoemaker MB, Bastarache L, et al. Association of thyroid function genetic predictors with atrial fibrillation: a phenome-wide association study and inverse-variance weighted average meta-analysis. JAMA Cardiol 2019;4:136-43.

22. Ellervik C, Roselli C, Christophersen IE, et al. Assessment of the relationship between genetic determinants of thyroid function and atrial fibrillation: a Mendelian randomization study. JAMA Cardiol 2019;4:144-52.

23. Andrade JG, Aguilar M, Atzema C, et al. The 2020 Canadian Cardiovascular Society/Canadian Heart Rhythm Society comprehensive guidelines for the management of atrial fibrillation. Can J Cardiol 2020;36:1847-948.

24. Perez MV, Mahaffey KW, Hedlin H, et al. Large-scale assessment of a smartwatch to identify atrial fibrillation. N Engl J Med 2019;381:1909-17.

## Supplementary Material

To access the supplementary material accompanying this article, visit *CJC Open* at https://www.cjcopen.ca/ and at https://doi.org/10.1016/j.cjco.2021.02.001.