ESC European Society of Cardiology

Europace (2021) **23**, 674–681
doi:10.1093/europace/euaa334

CLINICAL RESEARCH
Atrial fibrillation

Risk prediction of atrial fibrillation in the community combining biomarkers and genetics

Christin S. Börschel[1,2,†], Amelie H. Ohlrogge[1,†], Bastiaan Geelhoed[1,2], Teemu Niiranen[3], Aki S. Havulinna[3,4], Tarja Palosaari[3], Pekka Jousilahti[3], Michiel Rienstra[5], Pim van der Harst[5], Stefan Blankenberg[1,2], Tanja Zeller[1,2], Veikko Salomaa[3], and Renate B. Schnabel[1,2,*]

[1]Department of Cardiology, University Heart and Vascular Centre Hamburg-Eppendorf, Martinistr. 52, 20246 Hamburg, Germany; [2]German Centre for Cardiovascular Research (DZHK), Partner Site Hamburg/Kiel/Lübeck, 20246 Hamburg, Germany; [3]Department of Medicine, Turku University Hospital and University of Turku, Turku 20014, Finland; [4]Institute for Molecular Medicine Finland (FIMM), 00290 Helsinki, Finland; and [5]Department of Cardiology, University of Groningen, University Medical Centre Groningen, 9700RB Groningen, The Netherlands

Received 2 August 2020; editorial decision 2 October 2020; accepted after revision 8 October 2020; online publish-ahead-of-print 17 January 2021

Aims Classical cardiovascular risk factors (CVRFs), biomarkers, and common genetic variation have been suggested for risk assessment of atrial fibrillation (AF). To evaluate their clinical potential, we analysed their individual and combined ability of AF prediction.

Methods and results In $N = 6945$ individuals of the FINRISK 1997 cohort, we assessed the predictive value of CVRF, N-terminal pro B-type natriuretic peptide (NT-proBNP), and 145 recently identified single-nucleotide polymorphisms (SNPs) combined in a developed polygenic risk score (PRS) for incident AF. Over a median follow-up of 17.8 years, $n = 551$ participants (7.9%) developed AF. In multivariable-adjusted Cox proportional hazard models, NT-proBNP [hazard ratio (HR) of log transformed values 4.77; 95% confidence interval (CI) 3.66–6.22; $P < 0.001$] and the PRS (HR 2.18; 95% CI 1.88–2.53; $P < 0.001$) were significantly related to incident AF. The discriminatory ability improved asymptotically with increasing numbers of SNPs. Compared with a clinical model, AF risk prediction was significantly improved by addition of NT-proBNP and the PRS. The *C*-statistic for the combination of CVRF, NT-proBNP, and the PRS reached 0.83 compared with 0.79 for CVRF only ($P < 0.001$). A replication in the Dutch Prevention of REnal and Vascular ENd-stage Disease (PREVEND) cohort revealed similar results. Comparing the highest vs. lowest quartile, NT-proBNP and the PRS both showed a more than three-fold increased AF risk. Age remained the strongest risk factor with a 16.7-fold increased risk of AF in the highest quartile.

Conclusion The PRS and the established biomarker NT-proBNP showed comparable predictive ability. Both provided incremental predictive value over standard clinical variables. Further improvements for the PRS are likely with the discovery of additional SNPs.

Keywords Atrial fibrillation • Biomarkers • Genetics • Polygenic risk score • Community • Epidemiology • Risk Prediction

Introduction

Atrial fibrillation (AF) is a highly prevalent disease in aging populations worldwide with significant public health implications. Reliable risk prediction is needed to identify patients at increased risk of developing AF to possibly prevent AF, detect the disease earlier through intensified monitoring, and prevent complications such as stroke or heart failure.

* Corresponding author. Tel: +49 1522 2816064. *E-mail address*: r.schnabel@uke.de

† The first two authors contributed equally to the study.

What's new?

- Assessment of the predictive value of cardiovascular risk factor, N-terminal pro B-type natriuretic peptide (NT-proBNP), and 145 single-nucleotide polymorphisms (SNPs) combined in a polygenic risk score (PRS) for incident atrial fibrillation.
- The PRS and the established biomarker NT-proBNP showed comparable predictive ability.
- Both provided incremental predictive value over standard clinical variables.
- Further improvements in the PRS are likely with the discovery of additional SNPs.

Classical cardiovascular risk factors (CVRFs) such as obesity, hypertension, and prevalent cardiovascular disease are strong predictors of incident AF. However, CVRF only explain ~50% of the population attributable risk. Thus, novel risk indicators have been suggested. Blood biomarkers may improve the discriminatory ability of risk prediction. Two biomarkers have consistently been related to AF: C-reactive protein (CRP) as an indicator of inflammatory activity and natriuretic peptides represented by N-terminal pro B-type natriuretic peptide (NT-proBNP). In addition, the heritable component of AF appears to be substantial. Over the past years, genome-wide association studies (GWAS) have revealed a polygenic basis with common genetic variations in >25 loci associated with a modification of AF risk.[1–3] Based on these findings, different polygenic risk scores (PRS) have been published previously using increasing numbers of single-nucleotide polymorphisms (SNPs) in addition to CVRF.[4–6] To date, PRS marginally improved risk prediction for AF beyond CVRF.[5] However, it has been suggested that with more detailed information on genetic variation, genetic risk prediction will improve. Two recent GWAS analyses updated the number of replicated genetic loci to $N = 97$ and $N = 111$, respectively.[7,8]

In the current study, our aim was to use a systematic approach by combining CVRF, the protein biomarkers CRP, and NT-proBNP as well as currently available GWAS information as long-term predictors of AF in two European population-based cohorts. We further performed comparative analyses of the strength of the different types of risk indicators separately and in combination to understand their potential clinical ability.

Methods

Study cohort and data collection

The FINRISK study is a population-based, prospective cohort study conducted in Finland every 5 years since 1972. FINRISK uses a new, independent study sample on each study cycle. We used the data collected in 1997. Participants between the age of 25 and 74 years were randomly selected from five regions of Finland. The FINRISK study was approved by the local ethics committee. Informed written consent was obtained from all participants. A more detailed description of the study has been published previously.[9]

In total, the cohort comprised 8387 individuals. Individuals with a history of AF or missing genetic data were excluded. In addition, we excluded individuals with heart failure and cardiovascular disease including prior myocardial infarction because they usually are more closely monitored for AF and our focus was on risk prediction in primary prevention in individuals in the general population. In total, 6945 (*N*) individuals remained for analysis. All participants underwent physical examination, completed a questionnaire, and provided a blood sample. The following information on known risk factors was collected at baseline: age, sex, body mass index (BMI), systolic and diastolic blood pressure, antihypertensive medication, smoking status, average alcohol consumption per week, and diabetes. While BMI and blood pressure values were measured in the examination, other variables were attained by history. Sensitive CRP was determined by latex immunoassay (Abbott, Architect c8000; detection level 0.06 mg/L) and NT-proBNP by an electrochemiluminescence immunoassay (ECLIA, Roche Diagnostics) using the ELECSYS 2010 platform. Comorbidities and outcomes were identified by data of the National Hospital Discharge Register, the National Causes of Death Register, and the National Drug Reimbursement Register.

Validation cohort and data collection

The Dutch Prevention of REnal and Vascular ENd-stage Disease (PREVEND) study is a prospective, observational cohort study investigating the natural course of microalbuminuria and its relation to renal and cardiovascular disease. Details of the protocol and covariate definitions have been described elsewhere (www.prevend.org). The PREVEND study was approved by the local ethics committee and conducted in accordance with the Declaration of Helsinki. Informed written consent was obtained from all participants. In total, the cohort comprised 8592 individuals.[10] For our analyses, 3245 individuals were available after exclusion of individuals with prevalent AF as well as individuals with missing genetic data. A detailed description of the inclusion and validation of individuals developing AF has been published before.[10] At baseline the following information was collected: demographics, health behaviours, anthropometric measurements, cardiovascular, and metabolic risk factors. Furthermore, all participants provided a blood sample and two 24-h urine samples on two consecutive days were collected. N-terminal pro B-type natriuretic peptide was measured on the Roche Modular E170 (Roche Diagnostics, Mannheim, Germany) with commercially available kits. High-sensitivity CRP was determined by nephelometry (BN II, Dade Behring, Marburg, Germany; detection level 0.16 mg/L).

Polygenic risk score

A weighted PRS was calculated as sum of risk allele counts, weighted with β coefficients which represent the increase of the logarithm of hazard or odds per risk allele. To avoid overfitting, these β coefficients were taken from earlier GWAS studies and we only included independent SNPs that reached genome wide significance in these previous GWAS studies.

We combined the findings of three recent GWAS[7,8] to achieve the best possible PRS for our testing purposes. Twenty-three (*N*) SNPs and their β coefficients identified in the main analysis (*Table 1*) by Christophersen *et al.*[1] were used together with 67 SNPs and their β coefficients from the main analysis (*Table 1*) of Roselli *et al.*[8] These $23 + 67 = 90$ SNPs were independent [distance ≥500 kb or low linkage disequilibrium (correlation ≤ 0.8)] from each other. To these 90, we added 55 further independent SNPs and their β coefficients out of the 111 SNPs identified in the main analysis by Nielsen *et al.*[7] Fifty-six (*N*) of the Nielsen SNPs were either identical to or not independent from 1 of the 90 Christophersen/Roselli *et al.* SNPs. In total, 145 ($N=90+55$) independent and genome-wide significant SNPs were included in the PRS. For the PRS, we used a linear model not considering interactions between SNPs. A non-linear genetic model with interactions might have been more appropriate, but the computing of interactions at such a large scale has remained a challenge until today. Data for three SNPs (rs465276,

Table 1 Baseline characteristics of the FINRISK cohort

	All (*N* = 6945)	No AF (*n* = 6394)	AF (*n* = 551)
Age (years)	47.3 (36.5–58.2)	45.9 (35.7–57.0)	59.9 (52.6–66.0)
Men, no. (%)	3330 (47.9)	3003 (47.0)	327 (59.3)
Cardiovascular risk factors			
BMI (kg/m^2)	25.9 (23.4–28.9)	25.8 (23.2–28.6)	28.1 (25.2–31.1)
Systolic blood pressure (mmHg)	134 (122–150)	134 (122–148)	146 (132–160)
Antihypertensive medication, no. (%)	951 (38.2)	776 (35.2)	175 (60.6)
Smoking, no. (%)	1667 (24.3)	1570 (24.9)	97 (18.2)
Alcohol consumption per week (g)	26 (4–76)	26 (4–75)	25 (3–96)
Diabetes, no. (%)	321 (4.7)	278 (4.4)	43 (8.1)
Biomarkers and genetics			
C-reactive protein (mg/L)	0.98 (0.49–2.22)	0.96 (0.48–2.17)	1.30 (0.67–2.73)
NT-proBNP (ng/L)	45.1 (23.9–81.6)	43.9 (23.1–76.7)	84.4 (41.7–159.4)
PRS, 145 SNPs	8.94 (8.50–9.42)	8.91 (8.49–9.39)	9.23 (8.77–9.74)

Provided are median, 25th and 75th percentiles for continuous variables. Number and percentage are shown for categorical variables.
BMI, body mass index; NT-proBNP, N-terminal pro B-type natriuretic peptide; PRS, polygenic risk score; SNP, single-nucleotide polymorphisms.

rs12648245, and rs2012809) were not available in the PREVEND study. Therefore, 142 (*N*) of these SNPs were used to validate the PRS. A full list of all SNPs can be viewed in Supplementary material online, *Table S1*.

Statistical analysis

For continuous variables, medians (25th percentile, 75th percentile), for binary variables, absolute, and relative frequencies are given. Using time since baseline as a time scale, cause-specific multivariable Cox regressions with incident AF as an outcome and death as competing event were performed. Baseline age, male sex, BMI, systolic blood pressure, current smoking, average alcohol consumption per week, and diabetes were included as covariates. Additional multivariable Cox regressions were performed where (i) biomarkers (CRP and NT-proBNP), (ii) the PRS (based on 145 SNPs), and (iii) both biomarkers and the PRS were included as additional covariates. To avoid non-linearity, the statistical significances of all possible 2nd-order interactions (including self-interactions) were assessed and interactions of baseline age and male sex with baseline age, and NT-proBNP with diabetes were found to be needed as additional covariates. The interactions of BMI and smoking with time since baseline were needed to avoid violations of the proportional hazards assumption.

The non-normally distributed biomarkers CRP and NT-proBNP were log-transformed for analysis. Continuous variables were centred on their mean when included in an interaction. The hazard ratio (HR) of the interaction between the PRS and time since baseline was separately investigated and found to be significant and therefore this interaction was also included. Its HR was 0.85 per 5 years, which can be interpreted as follows: the here reported HR (2.18) for the PRS is for mean time since baseline and the HR decreases slowly over time (85% each 5 years). Hazard ratios for quartiles are based on refitting the full model with biomarkers and PRS with the tetrachotomized covariate. The logarithms of the multivariable adjusted HRs were used as weights in the weighted risk scores based on (i) CVRF and (ii) CVRF + log(NT-proBNP) + log(CRP). The two other risk scores (iii) CVRF + PRS and (iv) CVRF + log(NT-proBNP) + log(CRP) + PRS were calculated by adding PRS to the risk scores (i) and (ii), respectively. Concordancies of these risk scores were assessed and compared with each other using Harrel's *C*-statistic for survival data. Net reclassification improvements (NRIs) were calculated using risk categories based on the quartiles of the risk scores.

As an external validation, the above four risk scores were calculated in the external PREVEND cohort in exactly the same way as was done in the FINRISK cohort, with weight factors used in these calculations not refitted in PREVEND, but taken exactly equal. Also, *C*-statistics were calculated in PREVEND according to the same methods as we used for FINRISK.

Subdistribution hazard ratio (SHR) were computed using Fine and Gray regression models as implemented with the crr() function of the R package cmprsk. Analyses were performed with R v. 3.5.3. Cox.zph function was used for checking the proportional hazards assumptions, and Hosmer–Lemeshow tests were used for checking that the prediction models were adequately calibrated. A two-sided *P*-value of <0.05 was considered statistically significant.

Results

Baseline characteristics

The median age of the overall study cohort was 47.3 [95% confidence interval (CI) 36.5–58.2] years, 47.9% were men (*n* = 3330). Of the 6945 participants, 551 (*n*, 7.9%) developed AF during a median follow-up time of 17.8 years. Baseline characteristics are shown in *Table 1*. Atrial fibrillation incidence rates per 1000 person years for quartiles of age, CVRF, NT-proBNP, and PRS are presented in Supplementary material online, *Table S2*. The highest incidence rate was noted in individuals in the highest quartile of CVRF (13.29/1000 person years). In the highest quartile of PRS, incidence rate was 7.43/1000 person years. Time to AF curves for quartiles of NT-proBNP and PRS are displayed in *Figure 1*.

Atrial fibrillation prediction models

In multivariable-adjusted Cox proportional hazard models, NT-proBNP (HR of log transformed values 4.77; 95% CI 3.66–6.22; $P < 0.001$) and the PRS (HR 2.18, 95% CI 1.88–2.53; $P < 0.001$) were significantly related to incident AF. Because a significant interaction between time and PRS was found (HR per 5-year increase in time

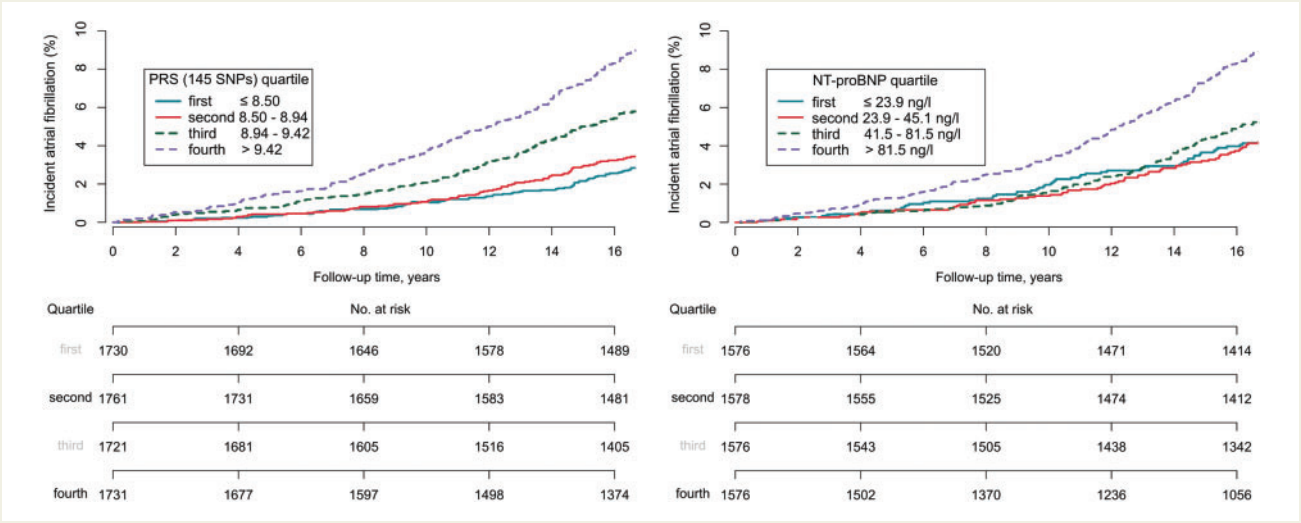

Figure 1 Time to atrial fibrillation (AF) curves for quartiles of N-terminal pro B-type natriuretic peptide (NT-proBNP) and the polygenic risk score (PRS). The respective numbers at risk are shown. AF, atrial fibrillation; NT-proBNP, N-terminal pro B-type natriuretic peptide; PRS, polygenic risk score.

Table 2 Multivariable Cox regression models for incident atrial fibrillation

	HR (95% CI) *P*-Value
Age (per 5 year increase)	1.74 (1.56–1.93)<0.001
Male sex	1.87 (1.38–2.54)<0.001
BMI (per 5 kg/m^2 increase)	1.38 (1.24–1.54)<0.001
Systolic blood pressure (per 10 mmHg increase)	1.03 (0.99–1.08)0.173
Smoking	1.09 (0.85–1.41)0.484
Average alcohol consumption per week (per 20 g increase)	1.02 (1.01–1.03)<0.001
Diabetes	1.23 (0.89–1.70)0.205
Biomarkers and genetics	
NT-proBNP per 10-fold increase	4.77 (3.66–6.22)<0.001
CRP per 10-fold increase	0.80 (0.64–1.00)0.053
PRS (145 SNPs) (per unit increase)	2.18 (1.88–2.53)<0.001
PRS (145 SNPs) × time (per 10-year time-increase per unit PRS-increase)	0.85 (0.73–0.97)0.021

Hazard ratios were adjusted for age, sex, BMI, systolic blood pressure, smoking, average alcohol consumption per week during the past 12 months, diabetes, BMI × time, smoking × time, age × age, and sex × age except for the associations of the respective variable itself. Log(NT-proBNP) and log(CRP) were additionally adjusted for each other and for log(NT-proBNP) × diabetes. The PRS and PRS × time hazard ratios had the same adjustments as log(NT-proBNP) and log(CRP) and were additionally adjusted for each other. Using 10 risk groups, all models were well calibrated with a Hosmer–Lemeshow χ^2 below the 95% confidence limit 16.9.

BMI, body mass index; CI, confidence interval; CRP, C-reactive protein; HR, hazard ratio; NT-proBNP, N-terminal pro B-type natriuretic peptide; PRS, polygenic risk score; SNP, single-nucleotide polymorphisms.

0.85; 95% CI 0.73–0.97; $P = 0.021$), the reported HR for the PRS is for the mean time since baseline (=8.2 years). Furthermore, there were significant associations between incident AF and age, male sex, BMI, as well as alcohol consumption. In multivariable-adjusted models, no association was seen between smoking, diabetes, systolic blood pressure, and AF (*Table 2*). Mortality in our cohort was 14.1%. Of the 6394, 818 (*n*, 12.8%) participants not developing AF died during follow-up. Analyses of competing risk of death using SHRs revealed only minor differences in the HRs (Supplementary material online, *Table* S3).

The predictive abilities for incident AF were analysed using four different models: CVRF alone; CVRF and biomarkers (CRP and NT-proBNP); CVRF and PRS; CVRF, biomarkers, and PRS, as presented in *Figure 2*. For CVRF alone, the *C*-index was 0.79. Addition of biomarkers or the PRS resulted in significantly improved risk prediction (*C*-index for improvement for biomarkers 0.81, ΔC biomarkers 0.014; 95% CI 0.0043–0.0238; $P = 0.004$ and *C*-index for improvement for PRS 0.82, ΔC PRS 0.022; 95% CI 0.012–0.032; $P < 0.001$). The model comprising CVRF, biomarkers, and the PRS showed the highest predictive ability. The *C*-index was moderately improved to 0.83 (ΔC 0.034; 95% CI 0.022–0.046; $P < 0.001$ for improvement). In accordance, NRI (13.3%) and integrated discrimination improvement (IDI) (0.681) were highest, when comparing the model with CVRF alone and the model with the combination of CVRF, biomarkers, and the PRS. Net reclassification improvements/IDIs are displayed in Supplementary material online, *Table S4*.

Further analysis of the PRS revealed most accurate AF prediction when all 145 SNPs were combined. As presented in *Figure 3*, addition of the first 10–20 SNPs accounted for the greatest impact on *C*-index. Including more SNPs resulted in a continuous further increase in the *C*-index.

As an additional analysis, HR quartiles were compared. The joint HR for individuals in the highest quartile of CVRF, PRS, and NT-proBNP compared with individuals in the respective lowest quartiles was 55.54 (95% CI 27.54–112.03; $P < 0.001$). Age stood out as the most important CVRF. Individuals in the oldest quartile had a 16.7-fold higher risk of AF than participants in the youngest quartile

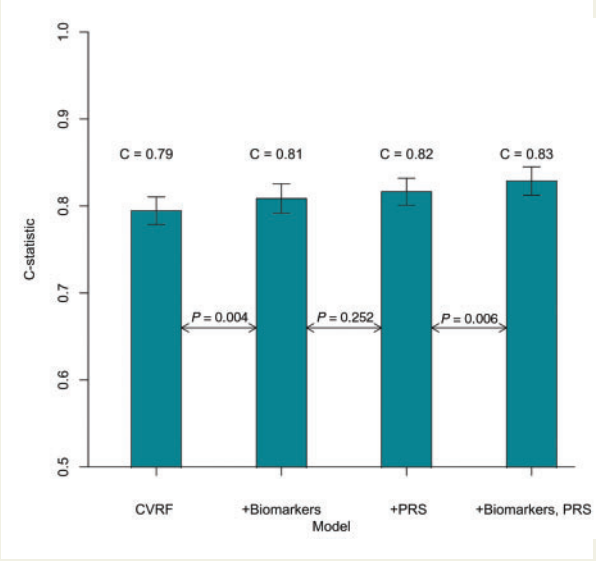

Figure 2 Bar graph of *C*-indices and 95% confidence intervals for AF prediction for classical cardiovascular risk factors, and additionally biomarkers, the PRS, and the combination of biomarkers and the PRS. AF, atrial fibrillation; PRS, polygenic risk score.

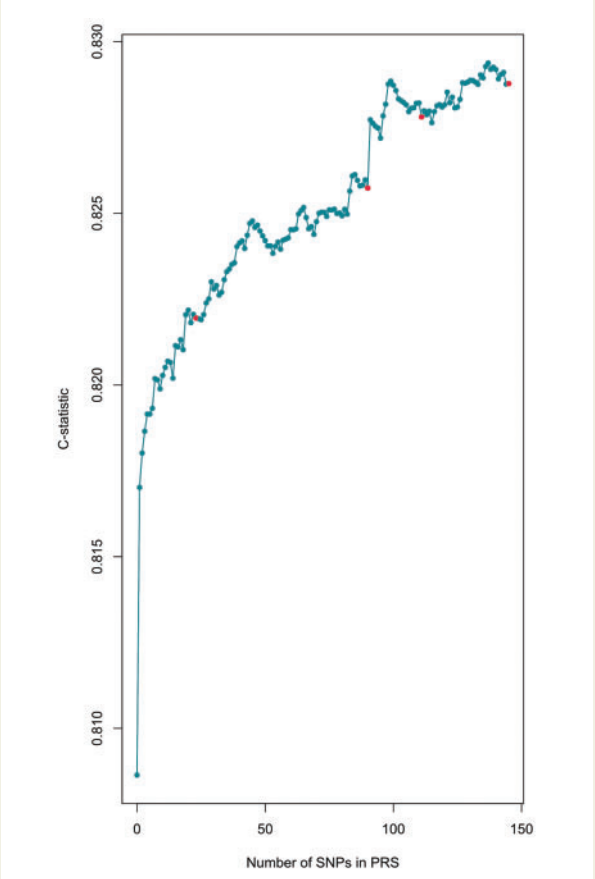

Figure 3 *C*-Index for increasing numbers of SNPs in the polygenic risk score (PRS) models for up to a total number of *N* = 145 SNPs. PRS, polygenic risk score; SNPs, single-nucleotide polymorphisms.

(*P* < 0.001). For both NT-proBNP and the PRS (independent of each other), the quartile with the highest values was associated with more than three-fold greater risk of AF than the lowest quartile (*P* < 0.001). Hazard ratios for age, NT-proBNP, and PRS in quartiles are presented in *Figure 4*.

Participants in the highest quartile of both PRS and CVRF had a 142-fold increased risk of developing AF compared with those in the lowest quartile of both. This risk increase was due to additive effects and no interaction of the PRS with CVRF was detected. Similar results were observed for NT-proBNP and CVRF (*Figure 5*).

Validation cohort

In the PREVEND cohort, the median age was 48.0 (39.0–59.0) years, 51% were men (*n* = 1654). Overall, 155 (*n*, 4.8%) participants developed AF during a median event-limited follow-up of 12.5 years.

Baseline characteristics are shown in Supplementary material online, *Table* S5. As for the FINRISK cohort, the predictive abilities for incident AF were analysed using the same four models. For CVRF alone, the *C*-index was 0.80. Addition of biomarkers resulted in significantly improved discriminatory ability (*C*-index 0.83, Δ*C* 0.032; 95% CI 0.015–0.049; *P* < 0.001) and the addition of the PRS reached borderline significance (*C*-index 0.82, Δ*C* 0.020; 95% CI 0.000–0.040; *P* = 0.051). The model comprising CVRF, biomarkers, and the PRS showed the highest predictive ability. The *C*-index was moderately improved to 0.85 (Δ*C* 0.052; 95% CI 0.032–0.072; *P* < 0.001).

Discussion

In our study, a model combining CVRF, the protein biomarker NT-proBNP, and common genetic polymorphisms moderately improved AF risk prediction compared with CVRF alone in two community-based cohorts, the FINRISK study as the discovery and the Dutch PREVEND cohort as the replication sample. The discriminatory ability of genetic risk prediction was enhanced with increasing numbers of SNPs. Although the PRS was a strong predictor of AF, its predictive performance was comparable with NT-proBNP. Both the PRS and NT-proBNP added predictive value over and above the model based on CVRF only. Age remained the strongest predictor of AF.

Our findings add to the rather limited number of pre-existing AF risk prediction studies including genetic information. Most of these were performed on selected cohorts, such as women only, postoperative or stroke patients[5,6,11] or were limited by a comparatively small sample size.[12] In contrast, our study is based on two large population-based cohorts. Only one of these previous studies had biomarker information available in their cohort but did not include it in their final risk prediction models.[6]

Massive genotyping efforts provide a large number of candidate loci for AF and lay the ground for further mechanistic investigation. However, the clinical implications have remained largely unclear. Recently, two studies could identify 145 different SNPs in total, associated with the development of AF.[7,8] We were able to prospectively validate these recent findings in our community-based cohort underlining their relevance for AF prediction.

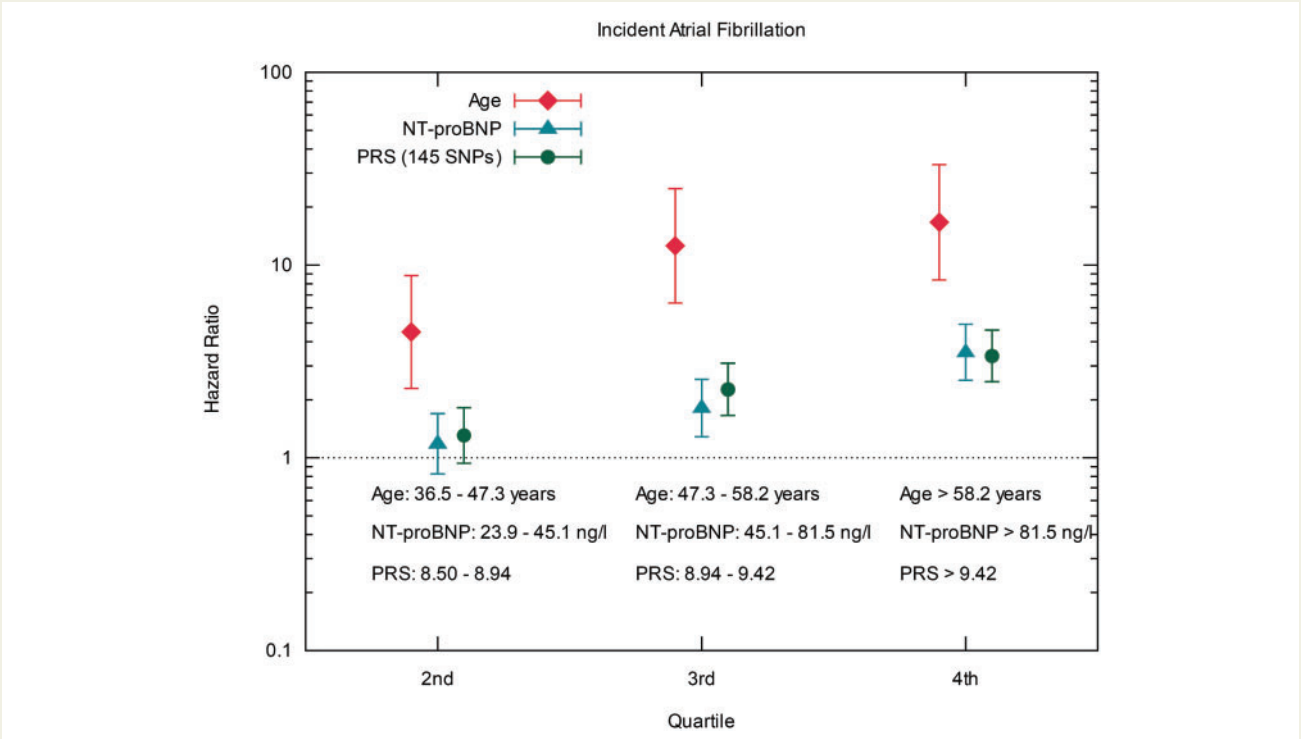

Figure 4 Hazard ratios and 95% confidence intervals for quartiles of age, NT-proBNP, and the PRS. NT-proBNP, N-terminal pro B-type natriuretic peptide; PRS, polygenic risk score.

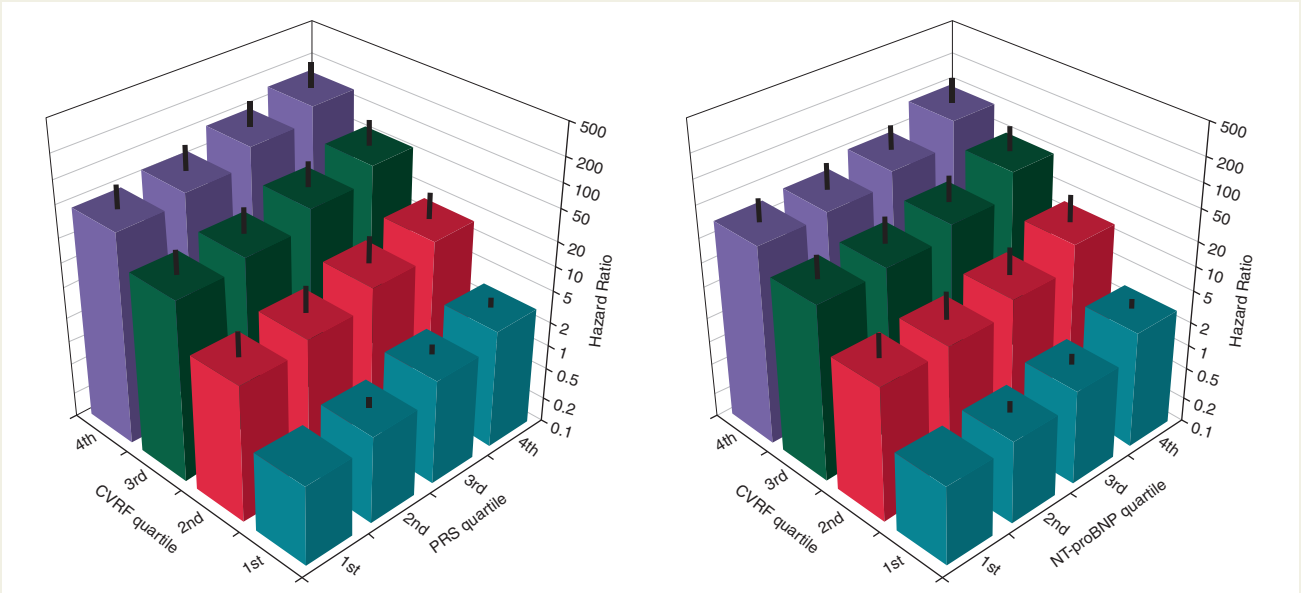

Figure 5 Association between quartiles of CVRF and PRS or NT-proBNP. CVRF, cardiovascular risk factors; NT-proBNP, N-terminal pro B-type natriuretic peptide; PRS, polygenic risk score.

In our analyses, we compared several PRS models including the most relevant SNPs from recently identified susceptibility loci.[1,7,8] As expected, we observed an initial steep increase in predictive ability on top of CVRF for the first 10–20 SNPs that had been identified during the first GWAS studies 10 years ago. These genetic variants are located in a pleiotropy of genes enriched within the transcriptional

regulation, development and signalling pathways of electrophysiological, contractile, and structural characteristics of cardiomyocytes. Many of these genes have previously been associated with other medical conditions, such as cardiovascular or musculoskeletal diseases. Addition of more SNPs resulted in a weaker but steady model improvement. Based on these results, we can project a further improvement of risk prediction with new AF loci. In contrast to previous studies investigating genetic AF prediction,[4,5] we could include substantially more risk alleles based on the recent GWAS findings.

The combination of known CVRF had the strongest predictive value compared with genetics and biomarkers. As previously reported, age remained the single most important risk factor.[13] Furthermore, we could confirm that the biomarker NT-proBNP is a strong predictor of incident AF.[14] In our study, the predictive performance of the PRS was comparable with that observed for NT-proBNP. The combination of CVRF, NT-proBNP, and the PRS resulted in the most accurate risk prediction.

In contrast to unadjusted analyses, systolic blood pressure was not significantly associated with AF in multivariable analyses in both cohorts after adjusting for age, sex, and BMI indicating probable confounding or possible cross correlations.

While our results may help to better understand the strength of genetic risk prediction in AF, near-term clinical implications of genetic risk prediction remain unclear. N-terminal pro B-type natriuretic peptide was recognized as a predictor for incident AF over a decade ago and is already established in clinical routine. Its use in AF prediction has been validated by several studies[15,16] and is further supported by our results. It can currently be measured faster and cheaper than genotyping. More specific biomarkers for AF are not yet available but may improve risk prediction in the future. In contrast, the PRS is less dependent on comorbidities and other confounders such as age, sex, and kidney function.[17] Although one could have argued that the genetic component of AF risk may be stronger in younger age groups, the predictive ability of the PRS was not strongly age dependent. Further investigations are needed to determine which subgroups may benefit most from additional genotyping compared with the clinical model.[18] Moreover, rapidly declining costs for genotyping, including genome sequencing, render a future use of genetics in clinical routine more likely. Eventually, an improved risk prediction may allow primary preventive interventions in individuals at risk and early detection of the disease. For a risk score use in clinical routine, simplicity and practicality are essential criteria. For now, easily available CVRF should remain the basis for all risk prediction efforts.

Limitations and strengths

The FINRISK study includes only individuals of European ancestry and the population of Finland in 1997 likely was more homogeneous than most societies today. Similarly, the PREVEND cohort is mainly composed of participants of European ancestry. Therefore, our findings may not be generalizable to other ethnicities. Likewise, the exclusion of individuals aged 75 or older or with previously diagnosed heart failure and myocardial infarction may limit generalizability. The outcomes were defined using national hospital-based and drug reimbursement-based databases. However, AF is often an outpatient diagnosis which does not have specific medications that are pathognomonic for the diagnosis. Thus, we may have missed AF cases with outpatient diagnosis only. However, we would expect that the association of PRS and biomarkers may be even stronger than calculated. An overestimation of the effects is not likely. Furthermore, our analysis is based on a single time point assessment. The individual risk and biomarkers such as NT-proBNP and CRP can vary over time. Serial measurements at different time points might have provided more reliable results. However, there are multiple previous studies suggesting that intraindividual fluctuations of CRP and NT-proBNP are relatively small allowing their use in an epidemiological context.[19,20] In addition, chronic kidney disease was not included in the risk score even though it is associated with AF and may influence biomarker levels.

A major strength of the study is the large number of carefully phenotyped and genotyped individuals in combination with a long follow-up based on nationwide health registers which cover the whole population. Furthermore, we were able to replicate our results in an external validation cohort.

Conclusions

Our study demonstrates that genetics and biomarkers add independent risk information for AF incidence. A systematic approach combining CVRF, NT-proBNP, and genetics provided the most accurate risk prediction. This observation calls for in-depth examination of the genetic underpinnings of AF. The identification of additional genetic loci associated with AF will help to further improve risk prediction. At present, however, CVRF and routine biomarkers remain the most important tools in AF prediction in the clinical setting.

Supplementary material

Supplementary material is available at *Europace* online.

Acknowledgements

The samples and data used for the research were obtained from THL and PREVEND biobank. We thank all study participants for their generous participation in the FINRISK study and the PREVEND study.

Funding

The BiomarCaRE Project is funded by the European Union Seventh Framework Programme (FP7/2007–2013) under grant agreement no. HEALTH-F2-2011 to 278913. The activities of the MORGAM Data Centre have been sustained by recent funding from European Union FP 7 project CHANCES (HEALTHF3-2010–242244). A part of the biomarker determinations in the population cohorts was funded by the Medical Research Council London (G0601463, identification no. 80983: Biomarkers in the MORGAM Populations). The PREVEND study was supported by the Dutch Kidney Foundation (Grant E0.13) and the Netherlands Heart Foundation (Grant NHS2010B280). The FINRISK surveys were mainly funded by budgetary funds of the National Institute for Health and Welfare, Finland (THL). Additional funding has been obtained from numerous non-profit foundations. V.S. has been supported by the Finnish Foundation for Cardiovascular Research. T.N. was supported by the Finnish Medical Foundation, the Emil Aaltonen Foundation, the Paavo Nurmi Foundation, and the Academy of Finland (Project No. 321351). T.Z. and S.B. are PIs of the German Centre for Cardiovascular Research (DZHK). RBS has received funding from the European Research Council

(ERC) under the European Union's Horizon 2020 research and innovation programme under the grant agreement No 648131, from the European Union's Horizon 2020 research and innovation programme under the grant agreement No 847770 (AFFECT-EU) and German Center for Cardiovascular Research (DZHK e.V.) (81Z1710103); German Ministry of Research and Education (BMBF 01ZX1408A) and ERACoSysMed3 (031L0239).

Conflict of interest: R.B.S. received consulting fees and speaker honoraria from BMS/Pfizer. V.S. has participated in a conference trip sponsored by Novo Nordisk and received an honorarium from the same source for participating in an advisory board meeting. He also has ongoing research collaboration with Bayer Ltd.

Data availability

The data underlying this article cannot be shared publicly to protect the privacy of the individuals that participated in the study. The data are available from the THL Biobank upon submission of a research plan and signing a data transfer Agreement (https://thl.fi/en/web/thl-biobank/for-researchers/application-process).

References

1. Christophersen IE, Rienstra M, Roselli C, Yin X, Geelhoed B, Barnard J *et al.*; METASTROKE Consortium of the ISGC. Large-scale analyses of common and rare variants identify 12 new loci associated with atrial fibrillation. *Nat Genet* 2017;**49**:946–52.
2. Lubitz SA, Brody JA, Bihlmeyer NA, Roselli C, Weng LC, Christophersen IE *et al.*; NHLBI GO Exome Sequencing Project. Whole exome sequencing in atrial fibrillation. *PLoS Genet* 2016;**12**:e1006284.
3. Hucker WJ, Saini H, Lubitz SA, Ellinor PT. Atrial fibrillation genetics: is there a practical clinical value now or in the future? *Can J Cardiol* 2016;**32**:1300–5.
4. Weng LC, Preis SR, Hulme OL, Larson MG, Choi SH, Wang B *et al.* Genetic predisposition, clinical risk factor burden, and lifetime risk of atrial fibrillation. *Circulation* 2018;**137**:1027–38.
5. Lubitz SA, Yin X, Lin HJ, Kolek M, Smith JG, Trompet S *et al.* Genetic risk prediction of atrial fibrillation. *Circulation* 2017;**135**:1311–20.
6. Everett BM, Cook NR, Conen D, Chasman DI, Ridker PM, Albert CM. Novel genetic markers improve measures of atrial fibrillation risk prediction. *Eur Heart J* 2013;**34**:2243–51.
7. 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. Nat Genet. 2018;50:1234-1239.
8. Roselli C, Chaffin MD, Weng L-C, Aeschbacher S, Ahlberg G, Albert CM *et al.* Multi-ethnic genome-wide association study for atrial fibrillation. *Nat Genet* 2018;**50**:1225–33.
9. Magnussen C, Niiranen TJ, Ojeda FM, Gianfagna F, Blankenberg S, Njølstad I *et al.* Sex differences and similarities in atrial fibrillation epidemiology, risk factors, and mortality in community cohorts. *Circulation* 2017;**136**:1–97.
10. Brouwers FP, De Boer RA, Van Der Harst P, Voors AA, Gansevoort RT, Bakker SJ *et al.* Incidence and epidemiology of new onset heart failure with preserved vs. reduced ejection fraction in a community-based cohort: 11-year follow-up of PREVEND. *Eur Heart J* 2013;**34**:1424–31.
11. Kolek MJ, Muehlschlegel JD, Bush WS, Parvez B, Murray KT, Stein CM *et al.* Genetic and clinical risk prediction model for postoperative atrial fibrillation. *Circ Arrhythm Electrophysiol* 2015;**8**:25–31.
12. Muse ED, Wineinger NE, Spencer EG, Peters M, Henderson R, Zhang Y *et al.* Validation of a genetic risk score for atrial fibrillation: a prospective multicenter cohort study. *PLoS Med* 2018;**15**:e1002525.
13. Schnabel RB, Yin X, Gona P, Larson MG, Beiser AS, McManus DD *et al.* 50 year trends in atrial fibrillation prevalence, incidence, risk factors, and mortality in the Framingham Heart Study: a cohort study. *Lancet* 2015;**386**:154–62.
14. Hijazi Z, Wallentin L, Siegbahn A, Andersson U, Christersson C, Ezekowitz J *et al.* N-terminal pro-B-type natriuretic peptide for risk assessment in patients with atrial fibrillation: insights from the ARISTOTLE trial (apixaban for the prevention of stroke in subjects with atrial fibrillation). *J Am Coll Cardiol* 2013;**61**:2274–84.
15. Sinner MF, Stepas KA, Moser CB, Krijthe BP, Aspelund T, Sotoodehnia N *et al.* B-type natriuretic peptide and C-reactive protein in the prediction of atrial fibrillation risk: the CHARGE-AF Consortium of community-based cohort studies. *Europace* 2014;**16**:1426–33.
16. Svennberg E, Henriksson P, Engdahl J, Hijazi Z, Al-Khalili F, Friberg L *et al.* N-terminal pro B-type natriuretic peptide in systematic screening for atrial fibrillation. *Heart* 2017;**103**:1271–7.
17. Takase H, Dohi Y. Kidney function crucially affects B-type natriuretic peptide (BNP), N-terminal proBNP and their relationship. *Eur J Clin Invest* 2014;**44**:303–8.
18. Hughes MF, Saarela O, Stritzke J, Kee F, Silander K, Klopp N *et al.* Genetic markers enhance coronary risk prediction in men: the MORGAM prospective cohorts. *PLoS One* 2012;**7**:e40922.
19. Glynn RJ, MacFadyen JG, Ridker PM. Tracking of high-sensitivity C-reactive protein after an initially elevated concentration: the JUPITER Study. *Clin Chem* 2009;**55**:305–12.
20. Frankenstein L, Remppis A, Frankenstein J, Hess G, Zdunek D, Slottje K *et al.* Variability of N-terminal probrain natriuretic peptide in stable chronic heart failure and its relation to changes in clinical variables. *Clin Chem* 2009;**55**:923–9.