



OPEN Self-rated health and the risk of incident atrial fibrillation in general population

Yonggu Lee¹, Jae Han Kim² & Jin-Kyu Park³✉

The association between self-rated health (SRH) and the development of atrial fibrillation (AF) in the general population remains underexplored. We reviewed the data of 9,895 participants in the Ansong-Ansan cohort study, a community-based Korean study. SRH was categorised as 'poor', 'fair', or 'good'. A newly developed AF was identified through biennial electrocardiography examinations and/or a self-reported history of physician-determined diagnoses. Over a median follow-up of 11 years, 149 patients (1.5%) developed AF. Compared with the 'good' group, the 'poor' group exhibited a higher risk of incident AF (adjusted hazard ratio, 1.85; 95% confidence interval, 1.19–2.87). Old age, female sex, lower educational level, smoking, cardiovascular diseases (hypertension, diabetes mellitus, and coronary artery disease), and inflammation were associated with 'poor' SRH. Along with SRH, age, male sex, urban residence, hypertension, and myocardial infarction were also associated with a higher risk of incidental AF. The combined model, which included conventional risk factors and SRH, demonstrated a marginally improved performance in predicting incident AF (concordance index: 0.704 vs. 0.714, $P = 0.058$). Poor SRH is independently associated with the development of AF in Korean adults. However, it plays a limited role in AF surveillance when combined with conventional AF risk factors.

Keywords Self-rated health, Atrial fibrillation, Risk factor, Prospective study

Atrial fibrillation (AF) is a common type of cardiac arrhythmia associated with significant morbidity and mortality¹. It poses a considerable burden on healthcare systems worldwide owing to its high incidence; potential complications, such as stroke and heart failure; and associated healthcare costs². Identifying its risk factors is crucial for the development of effective prevention and management strategies.

Self-rated health (SRH) encompasses the subjective assessment of an individual of their overall health and well-being. Several studies have demonstrated the significance of SRH as a predictor of cardiovascular diseases and mortality^{3,4}. Poor SRH is associated with the development of chronic kidney disease⁵, a major risk factor for AF⁶. However, studies on the role of SRH in predicting the incidence of AF in the general population are limited.

Investigating the potential association between SRH and the development of AF could help in developing a straightforward and cost-effective method for stratifying the risk of AF in the general population, potentially facilitating more individualised screening for AF. This prospective longitudinal community-based cohort study aimed to bridge this knowledge gap by investigating the association between SRH and incident AF in the general population.

Results

Baseline characteristics

The study population had a mean age of 52.2 ± 8.9 years, with 4,687 (47.4%) participants being men. Hypertension (HTN) and diabetes mellitus (DM) were present in 1,526 (15.4%) and 659 (6.7%) patients, respectively. The average body mass index (BMI) was 24.6 ± 3.1 kg/m².

Table 1 presents the baseline characteristics of the SRH group. Among the 9,895 participants surveyed, 3,380 (34.2%) rated their health status as 'very poor' or 'poor'. Most baseline characteristics significantly differed among the SRH groups, except for physical activity and the presence of heart failure. Compared with the 'fair' and 'good' SRH groups, the participants in the 'poor' SRH group were more likely to be older, women, and rural residents and have lower education levels. The 'poor' SRH group had a lower proportion of current smokers and

¹Division of Cardiology, Hanyang University Guri Hospital, Guri, Republic of Korea. ²Division of Cardiology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea. ³Division of Cardiology, Department of Internal Medicine, Hanyang University Medical Center, Seoul, Republic of Korea. ⁴+Yonggu Lee and Jae Han Kim have contributed equally to this work. ✉email: cardiohy@gmail.com

	Self-rated health (<i>n</i> = 9,895)			P value
	Poor	Fair	Good	
	(<i>n</i> = 3,380)	(<i>n</i> = 3,521)	(<i>n</i> = 2,994)	
Age (years)	54.8 ± 9.1	50.7 ± 8.4	51.3 ± 8.8	<0.001
Males	1,215 (35.9)	1,773 (50.4)	1,699 (56.7)	<0.001
Urban residential area	1,517 (44.9)	1,734 (49.2)	1,714 (57.2)	<0.001
High education level	243 (7.2)	519 (14.8)	531 (17.8)	<0.001
Height (cm)	157.7 ± 8.5	160.7 ± 8.3	161.8 ± 8.6	<0.001
Weight (kg)	61.1 ± 10.0	63.3 ± 9.9	64.7 ± 10.1	<0.001
BMI (kg/m ²)	24.6 ± 3.4	24.5 ± 3.0	24.7 ± 3.0	0.043
WC (cm)	83.1 ± 9.3	82.3 ± 8.6	82.6 ± 8.4	<0.001
Smoking				<0.001
Never-smoker	2,207 (66.2)	1,968 (56.5)	1,583 (53.3)	
Ex-smoker	387 (11.6)	547 (15.7)	578 (19.5)	
Current smoker	742 (22.2)	971 (27.9)	809 (27.2)	
Alcohol consumption				<0.001
Never-drinker	1,862 (55.5)	1,472 (42.1)	1,218 (40.8)	
Ex-drinker	261 (7.8)	213 (6.1)	168 (5.6)	
Current drinker	1,231 (36.7)	1,814 (51.8)	1,597 (53.6)	
Physical activity (MET-h/day)	23.5 [11.3–37.2]	23.6 [11.6–35.0]	24.0 [12.8–34.2]	0.410
Comorbidity				
Hypertension	741 (21.9)	480 (13.6)	305 (10.2)	<0.001
Diabetes	407 (12.0)	156 (4.4)	96 (3.2)	<0.001
Dyslipidaemia	89 (2.6)	98 (2.8)	56 (1.9)	0.042
Myocardial infarction	49 (1.5)	28 (0.8)	14 (0.5)	<0.001
Heart failure	10 (0.3)	8 (0.2)	3 (0.1)	0.231
Coronary artery disease	48 (1.4)	20 (0.6)	8 (0.3)	<0.001
Asthma	121 (3.6)	65 (1.8)	30 (1.0)	<0.001
Chronic lung disease	32 (0.9)	19 (0.5)	14 (0.5)	0.034
Stroke	74 (2.2)	31 (0.9)	11 (0.4)	<0.001
Laboratory tests				
HbA1c (%)	6.0 ± 1.2	5.7 ± 0.8	5.7 ± 0.8	<0.001
Total-C (mg/dL)	191.7 ± 37.8	190.5 ± 35.0	192.6 ± 34.7	0.052
HDL-C (mg/dL)	44.3 ± 10.3	44.8 ± 10.1	44.8 ± 9.7	0.090
Triglyceride (mg/dL)	166.7 [104–193]	159.7 [97–191]	161.0 [98–187]	0.022
WBC (×10 ³ /uL)	6.6 ± 1.9	6.6 ± 1.8	6.5 ± 1.8	0.003
CRP (mg/dL)	0.29 [0.08–0.29]	0.21 [0.06–0.23]	0.20 [0.06–0.23]	<0.001

Table 1. Baseline characteristics of the study population according to the self-rated health status. Data are expressed as the mean ± SD or number (%). For variables with a skewed distribution, the data are expressed as the median [1st quartile value and 3rd quartile value]. BMI, body mass index; WC, waist circumference; MET, metabolic equivalent task; C, cholesterol; HDL-C, high-density lipoprotein cholesterol; WBC, white blood cell; CRP, C-reactive protein; SD, standard deviation.

current drinkers. Compared with the ‘fair’ and ‘good’ SRH groups, those in the ‘poor’ SRH group was more likely to experience central obesity (waist circumference [WC] ≥ 90 cm for men and ≥ 85 cm for women), HTN, DM, myocardial infarction (MI), non-MI coronary artery disease (CAD), asthma, chronic lung disease, and stroke. Laboratory data indicated that haemoglobin A1c, white blood cell count, and C-reactive protein (CRP) levels were higher in the ‘poor’ SRH group than in the ‘fair’ and ‘good’ SRH groups.

Risk of AF according to the SRH status

During a median follow-up period of 11.0 ± 4.3 years (interquartile range, 6–12 years), new-onset AF was observed in 149 (1.5%) participants. Of these, 129 participants were identified as having new-onset AF through a 12-lead electrocardiography (ECG). The total observation period amounted to 86,587.1 person-years resulting in an AF incidence rate of 172.1 cases per 100,000 person-years. The Kaplan–Meier curves for the cumulative incidence of AF according to the SRH status are shown in Fig. 1. The incidence of AF was significantly higher in the ‘poor’ SRH group than in the ‘good’ SRH group (1.9% vs. 1.2%; log-rank *P* = 0.01). In the Cox proportional hazard model, as shown in Table 2, the ‘poor’ SRH group had a higher risk of incident AF than that in the ‘good’

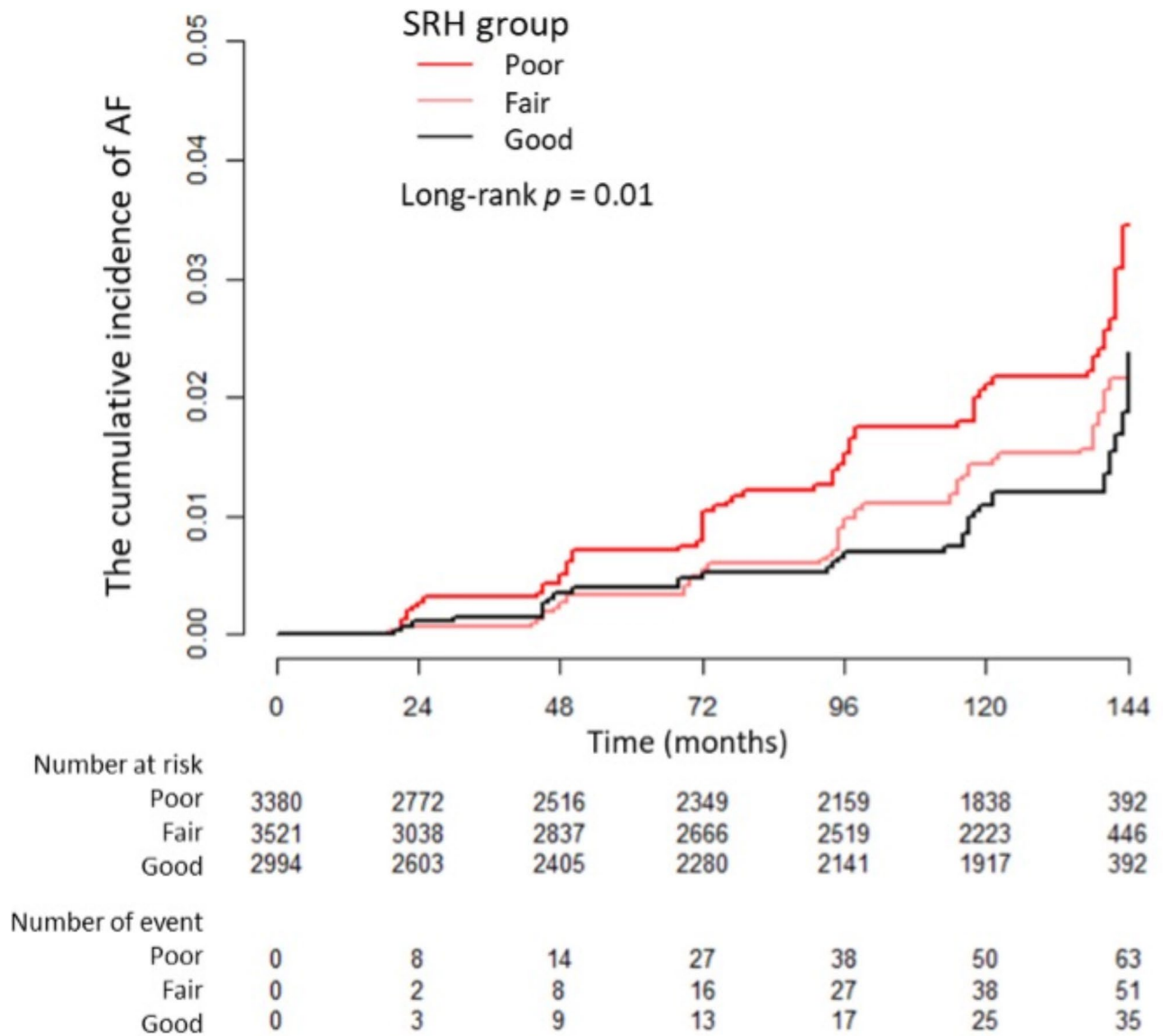


Fig. 1. The cumulative incidence of AF based on the SRH group. The incidence of AF was significantly higher in the poor SRH group than in the good SRH group. AF, atrial fibrillation; SRH, self-rated health.

SRH group after adjusting for all relevant covariates (adjusted hazard ratio, 1.79; 95% confidence interval [CI], 1.16–2.76).

Factors associated with poor SRH

Older age, HTN, DM, cardiovascular or cerebrovascular disease, asthma, chronic lung disease, and higher CRP levels were associated with poor SRH (Table 3). However, male sex, urban residence, high education level, current smoking, and drinking were negatively associated with poor SRH. In the multivariate model, various factors such as older age, sex, education level, current drinking, HTN, DM, CAD, asthma, stroke, and high CRP level remained significantly associated with poor SRH. Current smoking showed a positive association with poor SRH in the multivariate model, which was the opposite to the result in the univariate model.

SRH and AF risk discrimination

SRH alone only yielded a modest Harrell's concordance index (C-index) for predicting incident AF (Model 0; C-index, 0.607 [95% CI, 0.540–0.675]; $P < 0.001$) (Fig. 2; Table 4). The C-index for predicting incident AF using conventional risk factors was significantly higher than that achieved using SRH alone (Model 1; C-index, 0.708; 95% CI, 0.667–0.749). The predictive performance of the model comprising both SRH and conventional risk factors (Model 2) showed a slight improvement compared with that of the model comprising only conventional risk factors (Model 1), with only a marginal increase in the C-index (Model 2 vs. Model 1: 0.714 [95% CI, 0.675–0.755] vs. 0.704 [95% CI, 0.664–0.745], $P = 0.058$).

	SRH group				
	Poor		Fair		Good
The number of participants with new-onset AF	63 (1.9%)		51 (1.4%)		35 (1.2%)
	OR (95% CI)	P value	OR (95% CI)	P value	Reference
Model 0	1.78 (1.18–2.70)	0.006	1.25 (0.81–1.92)	0.315	[1]
Model 1	1.84 (1.20–2.80)	0.005	1.39 (0.90–2.14)	0.135	[1]
Model 2	1.84 (1.20–2.82)	0.005	1.41 (0.91–2.17)	0.120	[1]
Model 3	1.78 (1.16–2.75)	0.009	1.38 (0.89–2.13)	0.147	[1]
Model 4	1.79 (1.16–2.76)	0.009	1.38 (0.89–2.12)	0.148	[1]

Table 2. Hazard ratio for incident atrial fibrillation according to the self-rated health group. The HR and 95% CI were calculated using Cox proportional hazard regression analysis. Model 0: SRH Model 1: Model 0 + age, sex, residence, and education level Model 2: Model 1 + BMI, physical activity, alcohol, smoking, and eGFR (CKD-EPI) Model 3: Model 2 + HTN, DM, dyslipidaemia, heart failure, MI, non-MI CAD, asthma, chronic lung disease, and stroke Model 4: Model 3 + LDL cholesterol levels and CRP levels SRH, self-rated health; HR, hazard ratio; CI, confidence interval; eGFR, estimated glomerular filtration rate; CKD-EPI, Chronic Kidney Disease Epidemiology Collaboration; HTN, hypertension; DM, diabetes mellitus; CAD, coronary artery disease; MI, myocardial infarction; LDL, low-density lipoprotein; CRP, C-reactive protein.

	Univariate		Multivariate	
	OR (95% CI)	P	OR (95% CI)	P
Age (years)	1.049 (1.044–1.054)	<0.001	1.035 (1.029–1.041)	<0.001
Males	0.492 (0.452–0.536)	<0.001	0.493 (0.438–0.554)	<0.001
Urban residential area	0.724 (0.666–0.787)	<0.001	-	-
High education level	0.404 (0.349–0.468)	<0.001	0.584 (0.498–0.684)	<0.001
BMI	0.999 (0.986–1.012)	0.900	-	-
Current smoking	0.751 (0.681–0.829)	<0.001	1.422 (1.252–1.607)	<0.001
Current drinking	0.522 (0.479–0.569)	<0.001	0.746 (0.673–0.826)	<0.001
Hypertension	2.049 (1.835–2.289)	<0.001	1.420 (1.256–1.607)	<0.001
Diabetes	3.402 (2.890–4.004)	<0.001	2.889 (2.423–3.445)	<0.001
Dyslipidaemia	1.117 (0.857–1.455)	0.413	-	-
Myocardial infarction	2.267 (1.498–3.430)	<0.001	-	-
Coronary artery disease	3.337 (2.090–5.328)	<0.001	2.692 (1.641–4.416)	<0.001
Asthma	2.509 (1.911–3.293)	<0.001	2.019 (1.510–2.698)	<0.001
Chronic lung disease	1.877 (1.152–3.058)	0.011	-	-
Stroke	3.449 (2.356–5.049)	<0.001	2.467 (1.635–3.721)	<0.001
CRP (mg/dL)	1.471 (1.321–1.638)	<0.001	1.453 (1.298–1.626)	<0.001

Table 3. Factors associated with poor self-rated health. OR, odds ratio; CI, confidence interval; BMI, body mass index; CRP, C-reactive protein

Original SRH response and new-onset AF

Although the log-rank test was not significant, the Kaplan–Meier survival curve showed a significant trend toward increased incidence of AF as the original SRH responses worsened (Supplementary Fig S2). The multivariate Cox proportional hazards (CPH) model showed that poor SRH was significantly associated with a higher risk of new-onset AF after adjusting for covariates. Additionally, a marginally significant variation of incident AF was found among the original SRH responses in the analysis of variance ($P=0.0585$, Supplementary Table S1).

Discussion

In this large population-based cohort study, the incidence of newly developed AF was significantly higher in patients with poor SRH than in those with good SRH. The findings indicate that the likelihood of developing AF varied based on the responses to the SRH questionnaire. New-onset AF only occurred in 62.3% of the participants who rated their health as 'good' compared with those who rated it as 'poor'. This result suggests a significant link between subjective health perceptions and the pathophysiology of AF development in the general population. Poor SRH was significantly associated with conventional risk factors including old age, low education level, HTN, DM, cardiovascular diseases, and high inflammatory marker levels. Although 'poor' SRH was significantly associated with incident AF, its inclusion did not improve the predictive power of the model that incorporated conventional AF risk factors.

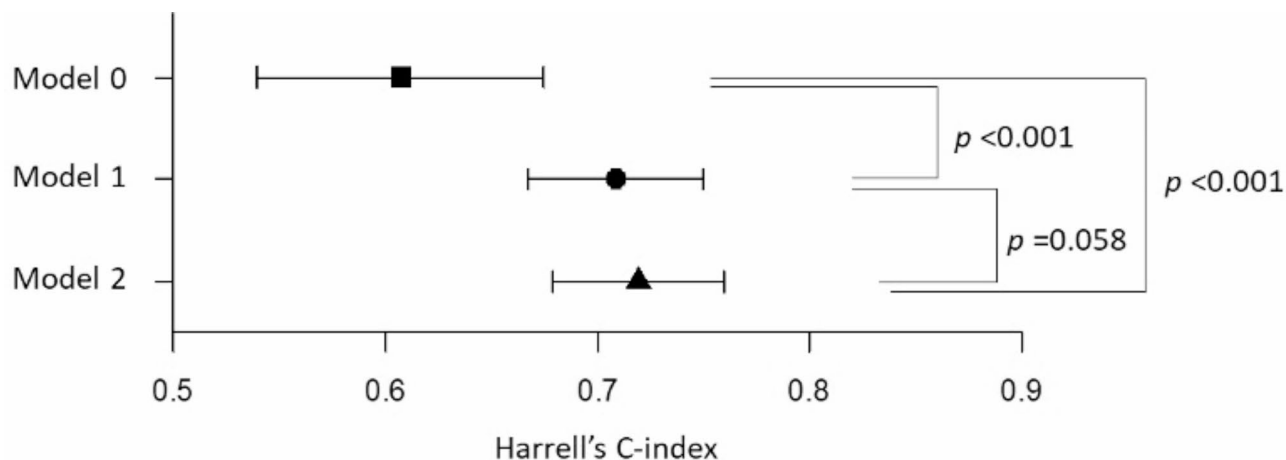


Fig. 2. The Harrell's concordance index of the prediction models for the incident AF. Model 0 included SRH; Model 1 included age, sex, urban residence, hypertension, and MI; and Model 2 included SRH and the covariates in Model 1. AF, atrial fibrillation; SRH, self-rated health; MI, myocardial infarction.

	C-index (95% CI)*	AIC	Predictors	HR (95% CI)	P value
Model 0	0.607 (0.540–0.675)	2596.2	SRH		
			Fair vs. Good	1.25 (0.81–1.92)	0.314
			Poor vs. Good	1.78 (1.18–2.70)	0.006
Model 1	0.704 (0.664–0.745)	2541.1	Age (every 10 years)	1.79 (1.46–2.20)	< 0.001
	$P < 0.001$	$\Delta AIC = 55.1$	Male sex	1.72 (1.24–2.39)	0.001
	vs. Model 0	vs. Model 0	Urban residential area	1.92 (1.35–2.72)	< 0.001
			Hypertension	1.60 (1.14–2.24)	0.006
			Myocardial infarction	3.61 (1.58–8.26)	0.002
Model 2	0.714 (0.675–0.755)	2538.0	SRH		
	$P = 0.058$	$\Delta AIC = 3.1$	Fair vs. Good	1.36 (0.88–2.09)	0.166
	vs. Model 1	vs. Model 1	Poor vs. Good	1.76 (1.15–2.70)	0.009
			Age (every 10 years)	1.75 (1.42–2.15)	< 0.001
			Male sex	1.88 (1.34–2.63)	< 0.001
			Urban residential area	1.95 (1.37–2.77)	< 0.001
			Hypertension	1.56 (1.11–2.19)	0.010
			Myocardial infarction	3.46 (1.51–7.92)	0.003

Table 4. Prediction models for incidental atrial fibrillation with self-rated health. Model 1 included age, sex, residence and education, BMI, physical activity, alcohol, smoking, eGFR, HTN, DM, dyslipidaemia, heart failure, MI, non-MI CAD, asthma, chronic lung disease, stroke, LDL cholesterol level, and CRP levels. Model 2 included covariates in Model 1 plus SRH. The models were reduced to the best-fit model using a backward variable selection procedure (criterion, $P > 0.05$). C-index, concordance index; CI, confidence interval; AIC, Akaike's information criterion; HR, hazard ratio; SRH, self-rating health index; BMI, body mass index; eGFR, estimated glomerular filtration rate; HTN, hypertension; DM, diabetes mellitus; CAD, coronary artery disease; MI, myocardial infarction; LDL, low-density lipoprotein; CRP, C-reactive protein.

To the best of our knowledge, this study is the first to investigate the association between SRH and the development of AF in a general population. In a cohort of older patients with AF, Abu et al. reported that one in every six older patients with AF rated their health status in the two lowest categories out of five, and a poorer health status was more prevalent in patients with multimorbidity⁷. However, their study did not investigate whether SRH was a risk factor for AF. Instead, they speculated that AF might serve as an indicator of frailty, which could manifest as poor SRH.

Because it is obvious that the personal perception of well-being cannot facilitate future AF development, the observed association between poor SRH at baseline and incident AF during follow-up might appear to be reversed in the temporal dimension. Nevertheless, several factors may explain this baffling association. First, poor SRH and AF share common risk factors. Advanced age, HTN, and cardiovascular diseases such as MI are all established risk factors for AF. These risk factors were also associated with poor SRH, as observed in our results and previous studies^{8–10}. Additionally, increased systemic inflammation, as indicated by elevated CRP

levels, has also been associated with poor SRH and incident AF^{10,11}. Consequently, a negative perception of one's health status resulting from these underlying comorbidities and conditions may precede the onset of AF by several years.

Second, new-onset AF was identified using a standard 12-lead ECG and self-reported physician-based diagnosis. These AF identification methods are susceptible to delays between the onset of symptoms and the formal diagnosis of AF. All cases of AF identified in this study, except for two, were persistent chronic AF. Typical AF begins with an episode of tachycardia by paroxysmal AF. This episode may be followed by recurrent episodes of paroxysmal AF, eventually progressing to chronic AF with stable symptoms over a longer follow-up period. Therefore, delays between the true onset of AF and the identification of AF are likely. Although ambulatory ECG monitoring may facilitate the detection of subclinical or paroxysmal AF¹², the scheduled biennial ECG assessment method used in our study may have caused a time delay between symptom onset and AF diagnosis. This time lag may account for the apparent predictive value of SRH for new-onset AF.

Third, unmeasured confounders, especially those related to social, psychological, or behavioural predispositions affecting both SRH status and AF incidence, may influence the observed association between SRH and new-onset AF. Obstructive sleep apnoea, a significant behavioural risk factor for AF, is also known to affect self-reported general health status, assessed using the short form (SF)-36 questionnaires¹³. Previous studies have indicated that lower educational levels are associated with poor SRH, which is consistent with our results⁷. Poor SRH is also associated with various variables related to socioeconomic status. This association creates different levels of discrepancy between SRH and physical health depending on socioeconomic status, which is considered a limitation of SRH¹⁴. However, the tendency for poor SRHs in participants with low socioeconomic status may explain the link between poor SRH and incident AF, as the association between socioeconomic status and incident AF has been established¹⁵.

SRH is useful for comparing the health status across population groups as an outcome variable in clinical trials and as a risk assessment tool in routine clinical practice¹⁶. SRH can predict mortality and hospitalisation. It also has a predictive value for stroke when added to a model that includes traditional risk factors, including smoking, high BMI, high cholesterol levels, HTN, and DM¹⁷, and the Framingham risk stratification model for coronary heart disease and cardiovascular death¹⁸. Although we found that the improvement in the predictive power of the model for incident AF was only marginally significant, further investigations in diverse populations could determine whether the model has additional predictive value for current risk assessment tools.

The strength of the current study lies in its inclusion of a large sample size of a well-managed prospective cohort of the general population, which provides robust confidence in the association between poor SRH and the incidence of AF. The SRH employed in this study was a simple yet powerful, valid, and reproducible measure for quantifying health perception. However, this study has several limitations. First, owing to the observational nature of the current study, establishing a causal relationship was challenging. Although we adjusted for relevant covariates such as age, sex, socioeconomic status, and comorbidities, the unmeasured confounding factors may have contributed to the association between SRH and incident AF. Second, the incidence of AF was relatively low compared with that in other cohort studies^{19,20}, which may indicate an under-recognition of AF. The use of biennial 12-lead standard ECG and self-reported physician diagnosis could have resulted in under-recognition of paroxysmal AF. However, Northeast Asians have the lowest prevalence of AF worldwide²¹, and a previous study by Lee et al. reported a similarly low incidence of AF in a nationwide insurance claims database of the Korean population²². Therefore, the under-recognition of paroxysmal AF may not have significantly affected the results of this study. Third, SRH was measured once at baseline, although it would most likely have varied throughout the follow-up period. Using time-varying CPH models with SRH as a time-varying covariate may have yielded different insights into the association between SRH and the risk of AF. However, unlike variables with established causal links such as blood pressure, a poor perception of health does not have a direct causal effect on AF development. Therefore, this study focused on examining the predictive value of SRH for new-onset AF. Finally, as death prevents the observation of new-onset AF, competing risk models may have provided an unbiased and accurate estimate of the cumulative incidence of AF. However, given that the participants were relatively healthy middle-aged community residents with a low risk of death, and the primary hypothesis of the study was to examine the independent association between SRH and incident AF, competing risk models may yield results similar to those obtained in this study.

In conclusion, the incidence of AF increases in patients who perceive their health as poor compared with those who perceive it as healthy. Poor perception of health is independently associated with incident AF and shares common risk factors with AF, including advanced age, lower education level, current smoking status, HTN, DM, CAD, asthma, stroke, and inflammation. Although SRH alone does not improve the predictive power of the conventional risk factor model for the incident AF, it may serve as a simple indicator of the risk of AF in the general population.

Methods

Study population

This study included 9,895 participants, aged 40–69 years, from the Ansong–Ansan cohort of the Korean Genome Epidemiology Study, conducted by the Korea Disease Control and Prevention Agency. This study investigated the genetic and environmental factors contributing to the prevalence of metabolic and cardiovascular diseases. Individuals residing in rural (Ansung) and urban (Ansan) communities were enrolled between June 2001 and January 2003. The detailed protocols have been described in previous publications^{23,24}.

A total of 10,030 eligible individuals who had resided in Ansong ($n = 5,018$) or Ansan ($n = 5,012$) for at least 6 months were enrolled in the study. Participants diagnosed with AF at baseline ($n = 74$) and those without SRH records ($n = 61$) were excluded, resulting in a final population of 9,895 participants.

A myriad of comprehensive health examinations, detailed on-site interviews, and meticulous laboratory tests were conducted during the baseline visit to a tertiary hospital. Six serial reassessments, following the entire cohort protocol, were performed through scheduled revisits every other year until 2014. All participants voluntarily enrolled in the study and provided written informed consent at the baseline assessment and each follow-up visit. The study adhered to the principles of the Declaration of Helsinki and was approved by the Korean National Research Institute of Health and the Institutional Review Board (IRB) of Hanyang University Medical Center (IRB no. HYUH 2017-12-033).

Assessment of lifestyle and medical history, physical examination, and laboratory tests

Well-trained investigators conducted comprehensive on-site interviews, collected critical lifestyle and clinical data, and performed physical examinations at tertiary hospitals during each visit. A structured questionnaire was used to obtain data on smoking, alcohol intake, education level, and specific medical conditions, such as HTN, DM, dyslipidaemia, cerebrovascular disease, CAD, and heart failure. Higher education was defined as obtaining a college degree or higher. The type and duration of physical activity were assessed using detailed questionnaires and quantified using estimated daily metabolic equivalent task scores. Blood pressure was measured by trained examiners using a mercury sphygmomanometer positioned at the level of the heart. Measurements were taken at least twice with the participant in a sitting position, and the results were averaged. If a blood pressure difference of ≥ 5 mmHg was observed between the two measurements, a third measurement was taken, and the last two measurements were averaged. WC was measured at the midpoint between the lowest rib and iliac crest at the end of expiration in a standing position.

Blood samples were collected after overnight fasting and analysed to determine the lipid profiles, haemoglobin A1c levels, white blood cell counts, and CRP levels using an automated analyser. Laboratory evaluations were performed in a single core clinical laboratory accredited and participating annually in inspections and surveys by the Korean Association of Quality Assurance for Clinical Laboratories. Blood concentrations of glucose, total cholesterol, high density lipoprotein (HDL)-cholesterol, and triglyceride were measured using the enzyme method (ADVIA 1650 and ADVIA 1800; Siemens Healthineers). Low-density lipoprotein (LDL)-cholesterol levels were calculated using the Friedewald formula²⁵. CRP was measured using a turbidimetric assay method (ADVIA 1650 and ADVIA 1800; Siemens Healthineers). HTN was defined as a diagnosis of HTN made by a physician or the regular use of antihypertensive medications. DM was defined as a diagnosis of DM by a physician, regular use of antidiabetic medications, or a haemoglobin A1c level of $\geq 6.5\%$. Dyslipidaemia was defined as a diagnosis of dyslipidaemia by a physician, the regular use of statin without a history of cardiovascular disease or DM, or the presence of at least one of the following abnormal laboratory test results: total cholesterol level ≥ 240 mg/dL, triglyceride level ≥ 150 mg/dL, or HDL cholesterol level < 45 mg/dL.

Assessment of self-rated health

Participants were asked to evaluate their overall health by responding to the question, 'How do you generally perceive your health?' The responses were initially divided into five categories: 'very poor', 'poor', 'fair', 'good', and 'very good'. Due to the relatively small numbers of participants who responded 'very poor' ($n = 411$) and 'very good' ($n = 154$), we reclassified the responses into three categories: 'poor (very poor/poor)' SRH ($n = 3,380$), 'fair' SRH ($n = 3,521$), and 'good (good/very good)' SRH ($n = 2,994$).

Standard 12-lead ECG and identification of AF

Standard 12-lead ECG (GE Marquette MAC 5000[®], GE Marquette Inc., Milwaukee, WI, USA) was performed in all participants at baseline and during every revisit. All ECG tracings were recorded at a paper speed of 25 mm/s and an amplitude of 0.1 mV/mm. The results were interpreted by a cardiologist and coded according to the Minnesota code classification system. AF was identified either through the presence of AF on a 12-lead ECG or a self-reported history of AF using a questionnaire administered by a physician before the baseline visit or during follow-up visits. The Minnesota codes 8-3-1, 8-3-2, 8-3-3, and 8-3-4 were used to classify AF. Newly developed AF was defined as the first identification of AF between visits. The date of new AF development was defined as the date when AF was first detected on ECG or when it was diagnosed by a physician.

Statistical analysis

The baseline characteristics of the participants were compared between the groups. One-way analysis of variance was used for analysing continuous variables, such as BMI, WC, and LDL-cholesterol levels, while Pearson's chi-square test was used for analysing categorical variables, including sex, comorbidities, and smoking history. Post-hoc analyses using Bonferroni correction were used to perform multiple comparisons. For continuous variables with a skewed distribution, comparisons between groups were performed using the Kruskal–Wallis test. The Shapiro–Wilk test was used to assess the normality of the distribution of continuous variables.

All variables had approximately 1% missing values. The missing values of individual variables ranged from 0 to 4.9%. Among the covariates in the multivariate regression models, the missing values of individual variables ranged from 0 to 1.04% (Supplementary Fig. S1). Rather than excluding cases with missing data, multiple imputations were performed using a bootstrap expectation-maximization algorithm²⁶. Five possible imputed datasets were generated. The average value was used for continuous variables, while the most frequent value was adopted for categorical variables.

The associations between SRH and incident AF were evaluated using a CPH model adjusted for age, sex, residence, education, BMI, physical activity, comorbidities (including HTN, DM, heart failure, dyslipidaemia, MI, non-MI CAD, asthma, and chronic lung disease), smoking status, alcohol consumption, and laboratory data (including LDL cholesterol levels and CRP levels). The factors associated with the 'poor' SRH group were analysed using logistic regression analysis. To determine whether SRH provided an additional predictive value

for incident AF when combined with conventional risk factors, we developed prediction models for new-onset AF with and without the SRH variable using multivariate CPH models. The goodness of fit of these prediction models was estimated using Harrell's C-index and Akaike information criterion (AIC). Harrell's C-indices were compared using the method proposed by Haibe-Kains et al.²⁷. The linear predictors identified using multivariate CPH models were employed as risk predictors for the C-index estimation. A difference of > 10 between the two AIC values was considered significant.

To evaluate the impact of reclassifying SRH responses from five to three categories based on the survey results, we performed a sensitivity analysis using a Kaplan–Meier survival analysis and a multivariate CPH model with the original five-category SRH responses. Due to the small number of 'very good' responses, the 'good' response was used as a reference when calculating the coefficients in the multivariate CPH model.

All statistical analyses were performed using the statistical software R-4.3.2 (R Core Team, R Foundation for Statistical Computing, Vienna, Austria) and its packages 'tableone', 'rms', 'Amelia', 'survival', 'BicManager', and 'survcomp' in RStudio-2023.12.1, Build 402 (RStudio Team, RStudio Inc., Boston, MA, USA). A P value of < 0.05 was considered significant.

Data availability

The datasets generated and/or analysed in the current study are available from the corresponding author upon reasonable request.

Received: 4 August 2024; Accepted: 14 October 2024

Published online: 20 October 2024

References

- Benjamin, E. J. et al. Heart Disease and Stroke Statistics-2019 update: a Report from the American Heart Association. *Circulation*. **139**, e56–e528. <https://doi.org/10.1161/cir.0000000000000659> (2019).
- Li, X., Tse, V. C., Au-Doung, L. W., Wong, I. C. K. & Chan, E. W. The impact of ischaemic stroke on atrial fibrillation-related healthcare cost: a systematic review. *Europace*. **19**, 937–947. <https://doi.org/10.1093/europace/euw093> (2017).
- Mavaddat, N., Parker, R. A., Sanderson, S., Mant, J. & Kinmonth, A. L. Relationship of self-rated health with fatal and non-fatal outcomes in cardiovascular disease: a systematic review and meta-analysis. *PLoS One*. **9**, e103509. <https://doi.org/10.1371/journal.pone.0103509> (2014).
- van der Linde, R. M. et al. Self-rated health and cardiovascular disease incidence: results from a longitudinal population-based cohort in Norfolk, UK. *PLoS One*. **8**, e65290. <https://doi.org/10.1371/journal.pone.0065290> (2013).
- Ko, H. L., Min, H. K. & Lee, S. W. Self-rated health and the risk of incident chronic kidney disease: a community-based Korean study. *J. Nephrol*. **36**, 745–753. <https://doi.org/10.1007/s40620-022-01518-3> (2023).
- Kim, S. M. et al. Association of chronic kidney Disease with Atrial Fibrillation in the General Adult Population: a Nationwide Population-based study. *J. Am. Heart Assoc.* **12**, e028496. <https://doi.org/10.1161/jaha.122.028496> (2023).
- Abu, H. O. et al. Multimorbidity, physical frailty, and self-rated health in older patients with atrial fibrillation. *BMC Geriatr*. **20**, 1–11 (2020).
- Kornej, J., Börschel, C. S., Benjamin, E. J. & Schnabel, R. B. Epidemiology of Atrial Fibrillation in the 21st Century: Novel methods and New insights. *Circ. Res.* **127**, 4–20. <https://doi.org/10.1161/circresaha.120.316340> (2020).
- Misra, S. et al. Outcomes of a remote Cardiac Rehabilitation Program for patients undergoing Atrial Fibrillation ablation: pilot study. *JMIR Cardio*. **7**, e49345. <https://doi.org/10.2196/49345> (2023).
- Lee, Y. et al. Single and persistent elevation of C-reactive protein levels and the risk of atrial fibrillation in a general population: the Ansan-Ansung Cohort of the Korean Genome and Epidemiology Study. *Int. J. Cardiol.* **277**, 240–246. <https://doi.org/10.1016/j.ijcard.2018.10.070> (2019).
- Uchino, B. N. et al. Self-rated health and inflammation: a test of Depression and Sleep Quality as mediators. *Psychosom. Med.* **81**, 328–332. <https://doi.org/10.1097/PSY.0000000000000683> (2019).
- Sanna, T. et al. Cryptogenic stroke and underlying atrial fibrillation. *N Engl. J. Med.* **370**, 2478–2486. <https://doi.org/10.1056/NEJMoa1313600> (2014).
- Finn, L., Young, T., Palta, M. & Fryback, D. G. Sleep-disordered breathing and self-reported general health status in the Wisconsin Sleep Cohort Study. *Sleep*. **21**, 701–706 (1998).
- Balaj, M. Self-reported health and the social body. *Soc. Theor. Health.* **20**, 71–89. <https://doi.org/10.1057/s41285-020-00150-0> (2022).
- Lunde, E. D. et al. Socioeconomic position and risk of atrial fibrillation: a nationwide Danish cohort study. *J. Epidemiol. Community Health*. **74**, 7–13. <https://doi.org/10.1136/jech-2019-212720> (2020).
- Jylha, M. What is self-rated health and why does it predict mortality? Towards a unified conceptual model. *Soc. Sci. Med.* **69**, 307–316. <https://doi.org/10.1016/j.socscimed.2009.05.013> (2009).
- Emmelin, M. et al. Self-rated ill-health strengthens the effect of biomedical risk factors in predicting stroke, especially for men—an incident case referent study. *J. Hypertens.* **21**, 887–896 (2003).
- May, M., Lawlor, D. A., Brindle, P., Patel, R. & Ebrahim, S. Cardiovascular disease risk assessment in older women: can we improve on Framingham? British women's heart and health prospective cohort study. *Heart*. **92**, 1396–1401 (2006).
- Rahman, F., Kwan, G. F. & Benjamin, E. J. Global epidemiology of atrial fibrillation. *Nat. Reviews Cardiol.* **11**, 639–654 (2014).
- Schnabel, R. B. et al. 50 year trends in atrial fibrillation prevalence, incidence, risk factors, and mortality in the Framingham Heart Study: a cohort study. *Lancet*. **386**, 154–162 (2015).
- Chugh, S. S. et al. Worldwide epidemiology of atrial fibrillation: a global burden of Disease 2010 study. *Circulation*. **129**, 837–847 (2014).
- Lee, S. R., Choi, E. K., Han, K. D., Cha, M. J. & Oh, S. Trends in the incidence and prevalence of atrial fibrillation and estimated thromboembolic risk using the CHA2DS2-VASc score in the entire Korean population. *Int. J. Cardiol.* **236**, 226–231 (2017).
- Lee, Y. et al. Association between insulin resistance and risk of atrial fibrillation in non-diabetics. *Eur. J. Prev. Cardiol.* **27**, 1934–1941. <https://doi.org/10.1177/2047487320908706> (2020).
- Kim, Y. & Han, B. G. Cohort Profile: the Korean Genome and Epidemiology Study (KoGES) Consortium. *Int. J. Epidemiol.* **46**, 1350. <https://doi.org/10.1093/ije/dyx105> (2017).
- Friedewald, W. T., Levy, R. I. & Fredrickson, D. S. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin. Chem.* **18**, 499–502 (1972).
- Honaker, J., King, G., Blackwell, M. & Amelia, I. I. A program for Missing Data. *J. Stat. Softw.* **45**, 1–47. <https://doi.org/10.18637/jss.v045.i07> (2011).

27. Schroder, M. S., Culhane, A. C., Quackenbush, J. & Haibe-Kains, B. Survcomp: an R/Bioconductor package for performance assessment and comparison of survival models. *Bioinformatics*. **27**, 3206–3208. <https://doi.org/10.1093/bioinformatics/btr511> (2011).

Acknowledgements

The authors would like to thank all participants and research staff of the Institute of Human Genomic Study at the Ansan Hospital of Korea University.

Author contributions

All authors have agreed to the content of this manuscript and its submission to Scientific Reports. In the case of acceptance of the manuscript, the copyright will be transferred to Scientific Reports. All data are fully available without restriction and all relevant data is included within the paper. As a corresponding author, I sign this letter on behalf of all coauthors.

Funding

None.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-76426-6>.

Correspondence and requests for materials should be addressed to J.-K.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024