

1 **Polygenic risk scores in atrial fibrillation: associations and clinical utility in disease**
2 **prediction**

3
4 Joel T Gibson^a BSc (Hons), MPhil

5 James H F Rudd^b MB BCh, PhD, FRCP, FESC

6
7 ^aDepartment of Medical Genetics, University of Cambridge, Box 238, Lv 6 Addenbrooke's
8 Treatment Centre, Addenbrooke's Hospital, Cambridge, CB2 0QQ, UK

9 ^bHeart & Lung Research Institute, University of Cambridge, Papworth Road, Cambridge
10 Biomedical Campus, Cambridge, CB2 0BB, UK

11
12 **Short title:** Polygenic risk scores in atrial fibrillation

13 **Word count:** 5,812 words

14 **Disclosures:** The authors have no conflicts of interest to disclose

15
16 **Corresponding author:**

17 Professor James H F Rudd

18 Heart & Lung Research Institute, University of Cambridge

19 Papworth Road, Cambridge Biomedical Campus

20 Cambridge, CB2 0BB, UK

21 Email: jhfr2@cam.ac.uk

25 **Abbreviations**

26 AUC (Area under the receiver operating characteristic curve)

27 AF (Atrial fibrillation)

28 NT-proBNP (B Type Natriuretic peptide)

29 PRS (Polygenic risk scores)

30 GWAS (Genome-wide association studies)

31 SNP (Single-nucleotide polymorphisms)

32 **Abstract**

33 Atrial fibrillation (AF) is a common heart arrhythmia and a major cause of cardioembolic
34 stroke. Therefore, accurate prediction is desirable to allow high-risk individuals to be
35 identified early and their risk lowered before complications arise. Polygenic risk scores (PRS)
36 have become a popular method of quantifying aggregated genetic risk from common
37 variants, but their clinical value in atrial fibrillation remains uncertain. This literature review
38 discusses the associations between PRS and AF risk, and the evidence for their clinical utility
39 in AF prediction. Stroke risk in AF patients is also considered. Despite consistent associations
40 between PRS and AF risk, the performance of PRS as a standalone tool for AF prediction was
41 poor. However, addition of PRS to existing AF prediction models commonly increased
42 predictive performance above that of the clinical models alone, including in cohorts with
43 comorbid cardiovascular disease. Associations between PRS and cardioembolic stroke risk in
44 AF patients have also been reported, but improvements to stroke prediction models from
45 PRS have been minimal. PRS are likely to add value to existing clinical AF prediction models,
46 however, standardisation of PRS across studies and populations will likely be required before
47 they can be meaningfully adopted into routine clinical practice.

48

49 **Keywords:** Atrial fibrillation, stroke, risk factor, polygenic risk score, genetic risk score,
50 CHARGE-AF

51

52 **Introduction**

53 Atrial fibrillation (AF) is the most common cardiac arrhythmia and is characterised by
54 irregular beating of the atria.¹ The prevalence rises sharply with age and is approximately 2%
55 for people in their 60s, almost 10% for people in their 70s, and 15-20% for people over 80.

56

57 Patients typically experience episodes of heart palpitations, shortness of breath, light-
58 headedness, tiredness or chest pain, but many patients are asymptomatic.² While AF is not
59 usually life threatening by itself, it does increase the risk of significant events such as heart
60 failure or stroke.³ Therefore, early and accurate identification of those at high-risk of AF is
61 desirable in order that risk-lowering interventions can be introduced before its onset.

62

63 AF shares many risk factors with other cardiovascular diseases. These include older age,
64 hypertension, diabetes, obesity, smoking and excessive alcohol use.³ Family history and
65 genetic predisposition have also been implicated. Additionally, other cardiovascular diseases
66 (e.g. valvular heart disease) and diseases affecting other organ systems (e.g. chronic kidney
67 disease) also increase risk.^{4,5}

68

69 The CHARGE-AF model is one of the most widely used prediction tools for AF, and is used
70 clinically to predict 5-year AF risk based on a range of clinical risk factors.⁶ However, despite
71 having relatively good discrimination for 5-year AF risk in the derivation cohort (C-
72 index=0.765), in the validation cohorts the C-index ranged from 0.664 to 0.705, indicating
73 modest discrimination overall. The model has also shown poor calibration in some external
74 cohorts.⁷ Additionally, its usefulness comes from its ability to predict relatively short-term AF
75 risk in an older population, after other risk factors have already manifested. Thus, it does not

76 predict long-term risk in younger individuals currently without risk factors who may
77 eventually go on to develop AF. The CHA₂DS₂-VASc score similarly uses clinical risk factors to
78 predict stroke risk in patients who already have AF to inform anticoagulant use.⁸

79

80 In contrast with clinical risk factors which typically present later in life, genetic risk is
81 established at conception and remains constant throughout life. Therefore, it may be
82 possible to use genetic data to predict long-term risk of disease prior to the appearance of
83 any clinical risk factors, when lifestyle changes may have greater impact. Polygenic risk
84 scores (PRS) have become a popular method of quantifying aggregated genetic risk from
85 common risk alleles identified from genome-wide association studies (GWAS).⁹ Many PRS
86 have been derived for AF, with studies demonstrating significant associations between PRS
87 and AF risk. Despite this, the clinical utility of PRS in the prediction of AF remains undecided,
88 and these scores are not yet used in routine clinical practice for AF management nor
89 endorsed by any of the major cardiovascular society AF guidelines. This literature review
90 summarises the associations between PRS and AF risk, and discusses the evidence for the
91 clinical utility of PRS for AF prediction.

92

93 **Early work and proof-of-concept**

94 Three early studies examined associations between PRS and AF risk, and laid the foundation
95 for future work (**Supplementary Table 1**).¹⁰⁻¹² In each study, small PRS were constructed
96 using 12 single-nucleotide polymorphisms (SNPs) previously associated with AF at the
97 genome-wide significance level. These included one SNP at each of the *KCNN3*, *PRRX1*,
98 *CAV1*, *C9orf3*, *SYNPO2L*, *SYNE2*, *HCN4* and *ZFH3* loci, and 4 SNPs at the *PITX2* locus. At

99 most loci, choice of lead SNP was the same between each study, with slight differences at
100 the remaining loci. Both weighted and unweighted PRS were examined.

101

102 Across the three studies, an increase in both the weighted and unweighted PRS was
103 associated with an increase in AF risk. One study reported that this association remained
104 even after adjusting for age, sex, AF risk factors and cardiovascular history, with a 2-fold risk
105 gradient between those in the highest and lowest PRS quintiles.¹² This same study also
106 reported that the PRS was also associated with risk of ischaemic stroke. Another study
107 reported that those with the highest unweighted PRS (15-16 risk alleles) had up to five times
108 the risk of AF compared to those with the lowest PRS (3-4 risk alleles).¹¹

109

110 Two of these studies also examined the clinical utility of PRS for AF prediction. One study
111 reported that addition of the weighted PRS to a clinical model including age, height, weight,
112 systolic blood pressure, smoking and alcohol use resulted in better predictive ability
113 compared to the clinical model alone (C-index=0.741 (95% CI, 0.709 – 0.774), C-index=0.718
114 (95% CI, 0.684 – 0.753), respectively).¹⁰ Reclassification and discrimination were also
115 significant using the combined model (continuous net reclassification improvement
116 (NRI)=0.490 (95% CI, 0.301 – 0.670), integrated discrimination improvement (IDI)=0.0053
117 (95% CI, 0.0033 – 0.0076)). Another study reported similar results, and further
118 demonstrated that addition of the PRS to the CHADS₂ score improved reclassification of AF
119 patients at risk of stroke (continuous NRI=0.166 (95% CI, 0.018 – 0.315)).¹²

120

121 **Associations with atrial fibrillation risk**

122 Following on from these early studies, many subsequent studies have similarly
123 demonstrated a significant association between higher PRS and increased risk of AF
124 **(Supplementary Table 1)**. For example, Weng et al. investigated AF risk in a European cohort
125 of 4,606 AF-free individuals over a median 9-year follow-up period. Using a weighted PRS
126 comprising 1,168 SNPs, the authors reported that the PRS was significantly associated with
127 AF risk, even after adjusting for CHARGE-AF score (hazard ratio (HR)=1.14 (95% CI, 1.11 –
128 1.16) per unit increase in PRS).¹³ A subsequent study reported that this same PRS developed
129 by Weng et al. was consistently associated with 5-year AF risk across four independent
130 cohorts in univariate analyses.¹⁴ Comparable observations have been reported using other
131 small to medium-sized PRS in a range of different populations, and adjusting for a range of
132 different covariates.^{15–21}

133

134 In addition to the relatively small PRS discussed above, a few studies have also developed
135 much larger PRS comprising millions of SNPs across the whole genome. One influential study
136 by Khera et al. developed a genome-wide PRS comprising 6,730,541 SNPs, demonstrating
137 that those with a score in the top 1% of the cohort had almost 5 times the risk of AF
138 compared to the rest of the cohort (OR=4.63 (95% CI, 3.96 – 5.39), adjusted by age, sex,
139 genotyping array and 4 principal components of ancestry).²² Another study used this same
140 PRS with similar basic covariates and reported a 46% increase in AF risk per standard
141 deviation (SD) increase in PRS.²³ Other studies have developed similar genome-wide scores,
142 with comparable results.^{24–26}

143

144 **Associations with age at onset of atrial fibrillation**

145 A number of studies have demonstrated an association between higher PRS and younger
146 age at onset of AF (**Supplementary Table 2**). One early Japanese study used an unweighted
147 PRS of 6 SNPs and reported that those with 9-12 risk alleles were on average 5 years
148 younger at onset of AF than those with 1-4 risk alleles.²⁷ A more recent Japanese study
149 obtained similar results, showing that those in the top 1% of a larger PRS were on average 4
150 years younger at onset than the remainder of the cohort.²⁸ Similar observations have been
151 reported in European cohorts.^{29,30} A study of patients undergoing catheter ablation of AF
152 found that those with a PRS in the top quintile were younger at time of ablation than those
153 in the bottom quintile.³¹

154

155 **Clinical utility for disease prediction**

156 Despite many studies demonstrating a strong association between PRS and AF risk, the
157 clinical utility of PRS as a standalone tool for AF prediction has had mixed results
158 (**Supplementary Table 1**). In particular, PRS derived from smaller numbers of SNPs tend to
159 have a weaker performance. Using a weighted PRS comprising 6 SNPs, one Japanese study
160 found that the predictive ability of the PRS alone was modest (area under the receiver
161 operating characteristic curve (AUC)=0.641 (95% CI, 0.628 – 0.653)).²⁷ This was despite a 4-
162 fold risk gradient between those in the top and bottom PRS quartiles. Another study of
163 participants from the UK Biobank reported similar findings using a PRS calculated using 181
164 SNPs.³²

165

166 PRS derived from larger numbers of SNPs have also have varying levels of success as
167 standalone prediction tools. For example, one study showed that the genome-wide PRS
168 developed by Khera et al. had good predictive ability for lone AF, and outperformed the

169 smaller PRS developed by Weng et al. (C-index=0.76 (95% CI, 0.72 – 0.80), C-index=0.70
170 (95% CI, 0.65 – 0.75), respectively).³³ However, another study applied this same genome-
171 wide PRS to a large cohort of over 11,500 participants and reported poor predictive
172 performance for the PRS alone (AUC=0.57).³⁴

173

174 Despite these mixed results, some studies have shown that including additional covariates
175 alongside the PRS can improve the predictive performance of the model.³⁵ In the same study
176 of 11,500 participants discussed above, even the simple addition of age and sex to the PRS
177 model resulted in significantly better predictive ability compared to PRS alone (AUC=0.78,
178 AUC=0.57, respectively).³⁴ Other studies have reported even greater predictive
179 performances incorporating additional AF risk factors such as body mass index and
180 hypertension.³⁶

181

182 Several studies have examined the value of incorporating PRS into existing clinical prediction
183 models. One study reported that addition of a PRS to a clinical model including age, sex,
184 smoking status, body mass index, diabetes, hypertension and history of myocardial
185 infarction or heart failure had better predictive ability for AF than the clinical model alone (C-
186 index=0.719, C-index=0.687, respectively).³⁷ Other studies have demonstrated similar
187 improvements using additional risk models.^{15,38} Importantly, addition of PRS has been
188 demonstrated to improve the predictive performance of the widely-used CHARGE-AF risk
189 prediction model, with improvements seen across a range of different cohorts.^{14,30,39,40}
190 These findings all suggest that PRS may be a clinically useful addition to existing AF
191 prediction models.

192

193 **High-risk groups**

194 In addition to studies examining PRS and AF risk in the general population, several studies
195 have also examined cohorts of individuals already at high risk of AF (**Supplementary Table**
196 **3**). This includes individuals with chronic kidney disease, heart failure and other
197 cardiovascular diseases, as well as those who have undergone heart surgery. Establishing
198 whether genetics contributes to AF risk in these high-risk groups would allow better patient
199 stratification and management.

200

201 *Chronic kidney disease*

202 Only one study has examined PRS and AF risk in patients undergoing haemodialysis for end-
203 stage kidney disease. The authors reported that the average PRS was significantly higher in
204 those who developed AF compared to those who did not.⁴¹ Furthermore, PRS was
205 independently associated with AF risk, even after adjusting for cardiovascular history and
206 risk factors (OR=1.24 (95% CI, 1.07 – 1.43) per 0.1 unit increase in PRS). In a post-hoc
207 analysis, this model correctly classified 75% of patients with AF.

208

209 *Heart failure*

210 Using a multivariate analysis incorporating the CHARGE-AF model, one study found that a
211 higher PRS in patients with heart failure was associated with increased risk of AF (OR=2.12
212 (95% CI, 1.84 – 2.45) per 1 unit increase in PRS).⁴² Addition of the PRS to the CHARGE-AF
213 model resulted in a higher predictive ability for AF (AUC=0.721 (95% CI, 0.704 – 0.737))
214 compared to the CHARGE-AF model alone (AUC=0.699 (95% CI, 0.682 – 0.716)). Interestingly,
215 one other recent study used an unsupervised machine learning approach to identify a
216 potential subtype of heart failure characterised by subsequent development of AF.⁴³ This

217 subtype was associated with an established PRS for atrial arrhythmias, suggesting that AF
218 risk in patients with heart failure has a genetic basis.

219

220 *Cardiac surgery*

221 Post-operative AF is a common adverse outcome following coronary revascularisation
222 procedures such as coronary artery bypass graft surgery. Using a cohort of post-cardiac
223 surgery patients, one study found that a PRS was significantly associated with risk of post-
224 operative AF, adjusting for age, sex and known predictors of AF (OR=1.92 (95% CI, 1.63 –
225 2.29) per SD increase in PRS).⁴⁴ Furthermore, addition of the PRS to standard clinical
226 predictors significantly improved the predictive ability of the model (C-index=0.782 (95% CI,
227 0.742 – 0.804), C-index=0.742 (95% CI, 0.700 – 0.764), respectively). Addition of PRS also
228 improved reclassification of intermediate-risk individuals into low and high-risk. Similar
229 findings were reported by a larger Finnish study.⁴⁵

230

231 *Cardiovascular disease*

232 In addition to heart failure and coronary revascularisation, one study examined AF risk
233 across cardiovascular disease in general using the PRS developed by Khera et al. This PRS
234 was associated with AF risk after adjusting for clinical AF risk factors (HR=1.40 (95% CI, 1.32
235 – 1.49) per 1 SD increase in PRS), even when stratifying by CHARGE-AF score.⁴⁶ Interestingly,
236 a high PRS was a stronger predictor of AF than most other clinical risk factors except age.
237 Addition of the PRS to a model incorporating the CHARGE-AF score and a clinical biomarker
238 (NT-proBNP) improved the predictive performance and showed good reclassification.

239

240 **Atrial fibrillation recurrence following treatment**

241 Given the high rate of AF recurrence following cardioversion or ablation therapy, some
242 studies have examined whether recurrence is influenced by genetic factors (**Supplementary**
243 **Table 4**). One South Korean study followed up patients for approximately 2 years following
244 radiofrequency catheter ablation and reported that a higher PRS was associated with a
245 higher risk of recurrence (HR=1.13 (95% CI, 1.03 – 1.24) per risk allele, adjusted for clinical
246 risk factors).⁴⁷ Smaller studies have reported similar findings following both catheter
247 ablation and direct current cardioversion.^{48,49} In contrast, a larger European study failed to
248 find any association between the PRS developed by Weng et al. and time to AF recurrence
249 following ablation.³¹

250

251 Only one study has examined treatment timing and AF recurrence. Using the PRS developed
252 by Khera et al. and adjusting for treatment group, the authors reported that a higher PRS
253 was associated with increased risk of AF recurrence over a mean 5-year follow-up (HR=1.08
254 (95% CI, 1.0 – 1.16)).⁵⁰ Compared to the typical standard of care, early rhythm control
255 significantly reduced risk of AF recurrence, however no interaction was found between PRS
256 and treatment strategy. This suggests that early rhythm control may be effective across all
257 levels of genetic risk.

258

259 **Stroke risk in atrial fibrillation patients**

260 Stroke is one of the most significant complications of AF. However, this risk can be lowered
261 through use of anticoagulants such as warfarin.⁵¹ Therefore, accurate identification of AF
262 patients at high risk of stroke is essential for appropriate patient management and
263 treatment selection. Several studies have examined the potential of PRS for stroke
264 prediction in AF patients with mixed results (**Supplementary Table 5**). Some studies have

265 used PRS derived from SNPs associated with AF (PRS-AF), while others have used PRS
266 derived from SNPs associated with stroke itself (PRS-stroke). Both are considered here.

267

268 Several studies have reported an association between PRS-AF and ischaemic stroke. This
269 appears to be driven by cardioembolic stroke, which is the major subtype implicated in AF.⁵²

270 One study reported that a higher PRS-AF was associated with an increased risk of
271 cardioembolic stroke after adjusting for age, sex, genotyping platform and principal
272 components of ancestry (OR=1.45 (95% CI, 1.16 – 1.83) per SD increase in PRS).³⁹ Another
273 study reported similar results, even after adjusting for known AF risk factors.⁵³ Interestingly,
274 associations with PRS-AF were not observed for other ischaemic stroke subtypes such as
275 atherothrombotic stroke or lacunar infarction, or for stroke overall, suggesting that PRS-AF
276 is highly specific to cardioembolic stroke risk.^{14,28}

277

278 In addition to PRS-AF, some studies have examined PRS-stroke for stroke prediction in AF
279 patients. In a multivariate analysis including the CHA₂DS₂-VASc model, one study
280 demonstrated that AF patients with a PRS-stroke in the highest tertile had a significantly
281 higher stroke risk than those in the lowest tertile (OR=1.29 (95% CI, 1.01 – 1.64)).⁵⁴

282 Interestingly, the predictive performance of the PRS was greatest in those with a lower
283 CHA₂DS₂-VASc score, suggesting that PRS may provide better risk stratification in AF patients
284 without high clinical risk of stroke. Despite this, another study reported that addition of a
285 PRS to the CHA₂DS₂-VASc model had minimal impact on the predictive ability compared to
286 the clinical model alone (C-index=0.61, C-index=0.60, respectively).⁵⁵

287

288 Only one study has used a multi-trait analysis to combine summary statistics for both AF and
289 cardioembolic stroke GWAS. While the predictive ability of the resulting PRS alone was
290 poorer than a basic clinical model including age, sex and hypertension (AUC=0.581,
291 AUC=0.947, respectively), addition of the PRS to the clinical model significantly improved
292 discrimination (AUC=0.950) and reclassification.⁵⁶

293

294 **Discussion**

295 Many recent studies have examined the potential for PRS to be used in AF prediction, and
296 evidence for its clinical value is growing. Following early success in the field, associations
297 between PRS and AF risk have been repeatedly found, with some studies reporting over a 3-
298 fold risk gradient between those with scores in the top and bottom quintiles of the
299 population. Importantly, these associations have persisted across a range of different cohort
300 sizes, ancestral groups and methods for PRS derivation, albeit with different effect sizes. A
301 few studies have also reported younger ages at AF onset in those with higher PRS.

302

303 Despite this strong foundation, the clinical utility of PRS as a standalone tool for AF
304 prediction has not been consistently demonstrated. This is perhaps understandable, given
305 one study found that PRS only explained 4.7% of the variance in AF risk in the general
306 population.⁵⁷ More promising results have been reported using PRS alongside age, sex and
307 other traditional AF risk factors, with many studies demonstrating improved predictive
308 performance over clinical factors alone. Therefore, the clinical value of PRS will most likely
309 come from their inclusion in established clinical risk models such as CHARGE-AF, which at
310 present does not incorporate a genetic risk component. This may also improve AF prediction
311 in those already at high risk, particularly those with heart failure or following heart surgery.

312

313 Interestingly, the predictive ability of PRS for lone AF (typically younger AF patients without
314 known risk factors) was quite good.³³ Furthermore, another study observed that
315 reclassification following addition of PRS to the CHARGE-AF model was greatest in those
316 under 65 years.¹⁴ Even with the PRS for stroke, predictive performance in AF patients was
317 greater in those with a low CHA₂DS₂-VASc score.⁵⁴ Taken together, these suggest that PRS
318 may be more useful in younger individuals, prior to the appearance of other clinical risk
319 factors. Therefore, one potential application of PRS is in healthy individuals with a known
320 family history who may wish to better understand their own future AF risk (PRS has been
321 shown to capture AF risk information independent of that provided by family history).²⁶ It
322 may also be informative to clinically-low risk individuals with only one or two other AF risk
323 factors as an additional incentive to make healthier lifestyle changes before further
324 comorbidities develop. Such testing of premorbid populations may become increasingly
325 feasible as sequencing time and costs continue to drop, and as more people are likely to
326 have already had their genomes sequenced at some point in their lives.

327

328 Despite many studies supporting the use of PRS for the prediction of AF, further work is still
329 required prior to the adoption of PRS into routine care. While inclusion of PRS in clinical
330 models may enhance predictive performance, studies examining the extent to which patient
331 outcomes are improved are still necessary to determine whether this is worthwhile. This
332 could include randomised controlled trials assigning people to either PRS-based screening or
333 standard care to establish whether knowledge of one's PRS leads to earlier AF diagnosis,
334 better adherence to preventative treatments, or reduced AF complications. These could also
335 examine whether PRS-guided care (such as earlier/more aggressive rhythm control

336 treatment or closer monitoring for those with a high PRS) improves patient outcomes
337 compared to standard care. Furthermore, no consistent definition of 'high' and 'low' PRS
338 exists, leading to inconsistent risk stratification of individuals. Therefore, studies evaluating
339 different PRS risk thresholds are needed to optimise identification of individuals most likely
340 to benefit from early intervention. Finally, many different PRS for AF have been developed
341 using different methods and comprising different SNPs and SNP weightings. However, no
342 consensus exists as to which specific PRS will be the most suitable for clinical use. Thus,
343 benchmarking studies to identify the best-performing PRS in each ancestral group may be
344 desirable to allow for more robust comparisons of risk between individuals.

345

346 One major ethical issue in the clinical implementation of PRS in general is that the majority
347 of GWAS are carried out using primarily European cohorts.⁵⁸ However, allele frequencies can
348 differ drastically between different ancestral groups, potentially leading to differences in
349 linkage disequilibrium and tagging of causal variants. This means that PRS derived from
350 European GWAS will have the greatest clinical benefit in European individuals, with those of
351 non-European ancestry receiving a poorer service due to less accurate risk predictions. This
352 will inevitably lead to poorer patient management in those populations, further widening
353 health inequalities. This is an important point to consider here, as the large majority of AF
354 PRS have been both derived and tested using only European cohorts (**Supplementary Table**
355 **1**). Such scores have limited transferability to other populations, with one study
356 demonstrating very poor predictive performance when applied to African, East Asian and
357 South Asian cohorts (all AUCs < 0.6).⁵⁹ Two other studies also reported that PRS derived from
358 European GWAS had poorer performance in their Japanese/Korean cohorts compared to
359 PRS derived from Japanese GWAS.^{28,60} Interestingly, in both studies the highest performance

360 was obtained from PRS derived from cross-ancestral data, highlighting the need for more
361 genetically-diverse training datasets to overcome some of these biases.

362

363 **Conclusion**

364 Associations between PRS and AF risk are strong, but as a standalone tool PRS will probably
365 have little clinical value. The utility of PRS will likely come from their addition to existing
366 clinical prediction tools, with several studies reporting modest but significant improvements
367 in both predictive ability and reclassification. This has the potential to provide better AF risk
368 stratification in the general population, in particular high-risk groups, as well as in younger
369 individuals without current AF risk factors or comorbidities. Identification of AF patients at
370 high risk of stroke may also be improved using PRS to better inform treatment selection.

371

372 **Author Contributions**

373 **JTG:** Study design, literature search, review and critical analysis of manuscripts, writing –
374 original draft. **JHFR:** Study design, supervision, writing – review & editing. JTG undertook this
375 work while a student at the University of Cambridge.

376

377 **Funding**

378 JHFR is part-supported by the NIHR Cambridge Biomedical Research Centre, the British Heart
379 Foundation, HEFCE, the EPSRC and the Wellcome Trust.

380

382 **References**

- 383 1. Lippi G, Sanchis-Gomar F, Cervellin G: Global epidemiology of atrial fibrillation: an
384 increasing epidemic and public health challenge. *Int J Stroke SAGE Publications*, 2021;
385 16:217–221.
- 386 2. Rienstra M, Lubitz SA, Mahida S, et al.: Symptoms and functional status of patients with
387 atrial fibrillation. *Circulation American Heart Association*, 2012; 125:2933–2943.
- 388 3. Kornej J, Börschel CS, Benjamin EJ, Schnabel RB: Epidemiology of atrial fibrillation in the
389 21st century: novel methods and new insights. *Circ Res American Heart Association*,
390 2020; 127:4–20.
- 391 4. Anumonwo JMB, Kalifa J: Risk factors and genetics of atrial fibrillation. *Cardiol Clin*
392 2014; 32:485–494.
- 393 5. Rahman F, Kwan GF, Benjamin EJ: Global epidemiology of atrial fibrillation. *Nat Rev*
394 *Cardiol Nature Publishing Group*, 2014; 11:639–654.
- 395 6. Alonso A, Krijthe BP, Aspelund T, et al.: Simple risk model predicts incidence of atrial
396 fibrillation in a racially and geographically diverse population: the CHARGE-AF
397 consortium. *J Am Heart Assoc American Heart Association*, 2013; 2:e000102.
- 398 7. Kolek MJ, Graves AJ, Xu M, et al.: Evaluation of a prediction model for the development
399 of atrial fibrillation in a repository of electronic medical records. *JAMA Cardiol* 2016;
400 1:1007–1013.

- 401 8. Lip GYH, Nieuwlaat R, Pisters R, Lane DA, Crijns HJGM: Refining clinical risk stratification
402 for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-
403 based approach: the Euro heart survey on atrial fibrillation. *Chest* 2010; 137:263–272.
- 404 9. Lambert SA, Abraham G, Inouye M: Towards clinical utility of polygenic risk scores. *Hum*
405 *Mol Genet* 2019; 28:R133–R142.
- 406 10. Everett BM, Cook NR, Conen D, Chasman DI, Ridker PM, Albert CM: Novel genetic
407 markers improve measures of atrial fibrillation risk prediction. *Eur Heart J* 2013;
408 34:2243–2251.
- 409 11. Lubitz SA, Lunetta KL, Lin H, et al.: Novel genetic markers associate with atrial
410 fibrillation risk in Europeans and Japanese. *J Am Coll Cardiol* 2014; 63:1200–1210.
- 411 12. Tada H, Shiffman D, Smith JG, et al.: Twelve–single nucleotide polymorphism genetic
412 risk score identifies individuals at increased risk for future atrial fibrillation and stroke.
413 *Stroke American Heart Association*, 2014; 45:2856–2862.
- 414 13. Weng L-C, Preis SR, Hulme OL, et al.: Genetic predisposition, clinical risk factor burden,
415 and lifetime risk of atrial fibrillation. *Circulation American Heart Association*, 2018;
416 137:1027–1038.
- 417 14. Khurshid S, Mars N, Haggerty CM, et al.: Predictive accuracy of a clinical and genetic
418 risk model for atrial fibrillation. *Circ Genomic Precis Med American Heart Association*,
419 2021; 14:e003355.
- 420 15. Börschel CS, Ohlrogge AH, Geelhoed B, et al.: Risk prediction of atrial fibrillation in the
421 community combining biomarkers and genetics. *EP Eur* 2021; 23:674–681.

- 422 16. Zhang J, Chen G, Wang C, et al.: Associations of risk factor burden and genetic
423 predisposition with the 10-year risk of atrial fibrillation: observations from a large
424 prospective study of 348,904 participants. *BMC Med* 2023; 21:88.
- 425 17. Svensson T, Kitlinski M, Engström G, Melander O: Psychological stress and risk of
426 incident atrial fibrillation in men and women with known atrial fibrillation genetic risk
427 scores. *Sci Rep Nature Publishing Group*, 2017; 7:42613.
- 428 18. Tikkanen E, Gustafsson S, Ingelsson E: Associations of fitness, physical activity, strength,
429 and genetic risk with cardiovascular disease. *Circulation American Heart Association*,
430 2018; 137:2583–2591.
- 431 19. Halford JL, Weng L-C, Choi SH, et al.: Associations Between Alcohol Intake and Genetic
432 Predisposition With Atrial Fibrillation Risk in a National Biobank. *Circ Genomic Precis
433 Med American Heart Association*, 2020; 13:e003111.
- 434 20. Wang N, Sun Y, Zhang H, et al.: Long-term night shift work is associated with the risk of
435 atrial fibrillation and coronary heart disease. *Eur Heart J* 2021; 42:4180–4188.
- 436 21. Yang L, Feng H, Ai S, et al.: Association of accelerometer-derived circadian
437 abnormalities and genetic risk with incidence of atrial fibrillation. *Npj Digit Med Nature
438 Publishing Group*, 2023; 6:1–8.
- 439 22. Khera AV, Chaffin M, Aragam KG, et al.: Genome-wide polygenic scores for common
440 diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet
441 Nature Publishing Group*, 2018; 50:1219–1224.

- 442 23. Huang K, Trinder M, Roston TM, Laksman ZW, Brunham LR: The interplay between titin,
443 polygenic risk, and modifiable cardiovascular risk factors in atrial fibrillation. *Can J*
444 *Cardiol* 2021; 37:848–856.
- 445 24. Ritchie SC, Lambert SA, Arnold M, et al.: Integrative analysis of the plasma proteome
446 and polygenic risk of cardiometabolic diseases. *Nat Metab* Nature Publishing Group,
447 2021; 3:1476–1483.
- 448 25. Ye Y, Chen X, Han J, Jiang W, Natarajan P, Zhao H: Interactions between enhanced
449 polygenic risk scores and lifestyle for cardiovascular disease, diabetes, and lipid levels.
450 *Circ Genomic Precis Med American Heart Association*, 2021; 14:e003128.
- 451 26. Mars N, Lindbohm JV, della Briotta Parolo P, et al.: Systematic comparison of family
452 history and polygenic risk across 24 common diseases. *Am J Hum Genet* 2022;
453 109:2152–2162.
- 454 27. Liu L, Eban Y, Nitta J, et al.: Genetic variants associated with susceptibility to atrial
455 fibrillation in a Japanese population. *Can J Cardiol* 2017; 33:443–449.
- 456 28. Miyazawa K, Ito K, Ito M, et al.: Cross-ancestry genome-wide analysis of atrial
457 fibrillation unveils disease biology and enables cardioembolic risk prediction. *Nat Genet*
458 Nature Publishing Group, 2023; 55:187–197.
- 459 29. Nielsen JB, Fritsche LG, Zhou W, et al.: Genome-wide study of atrial fibrillation identifies
460 seven risk loci and highlights biological pathways and regulatory elements involved in
461 cardiac development. *Am J Hum Genet* 2018; 102:103–115.

- 462 30. Mars N, Koskela JT, Ripatti P, et al.: Polygenic and clinical risk scores and their impact on
463 age at onset and prediction of cardiometabolic diseases and common cancers. *Nat Med*
464 Nature Publishing Group, 2020; 26:549–557.
- 465 31. Shoemaker MB, Husser D, Roselli C, et al.: Genetic susceptibility for atrial fibrillation in
466 patients undergoing atrial fibrillation ablation. *Circ Arrhythm Electrophysiol American*
467 *Heart Association*, 2020; 13:e007676.
- 468 32. Lello L, Raben TG, Yong SY, Tellier LCAM, Hsu SDH: Genomic prediction of 16 complex
469 disease risks including heart attack, diabetes, breast and prostate cancer. *Sci Rep*
470 Nature Publishing Group, 2019; 9:15286.
- 471 33. Lazarte J, Dron JS, McIntyre AD, et al.: Evaluating polygenic risk scores in “lone” atrial
472 fibrillation. *CJC Open* 2021; 3:751–757.
- 473 34. Homburger JR, Neben CL, Mishne G, Zhou AY, Kathiresan S, Khera AV: Low coverage
474 whole genome sequencing enables accurate assessment of common variants and
475 calculation of genome-wide polygenic scores. *Genome Med* 2019; 11:74.
- 476 35. Wong CK, Makalic E, Dite GS, et al.: Polygenic risk scores for cardiovascular diseases and
477 type 2 diabetes. *PLOS ONE Public Library of Science*, 2022; 17:e0278764.
- 478 36. Okubo Y, Nakano Y, Ochi H, et al.: Predicting atrial fibrillation using a combination of
479 genetic risk score and clinical risk factors. *Heart Rhythm* 2020; 17:699–705.
- 480 37. Muse ED, Wineinger NE, Spencer EG, et al.: Validation of a genetic risk score for atrial
481 fibrillation: a prospective multicenter cohort study. *PLOS Med Public Library of Science*,
482 2018; 15:e1002525.

- 483 38. Lazarte J, Dron JS, McIntyre AD, et al.: Role of common genetic variation in lone atrial
484 fibrillation. *Circ Genomic Precis Med American Heart Association*, 2021; 14:e003179.
- 485 39. Lubitz SA, Yin X, Lin HJ, et al.: Genetic risk prediction of atrial fibrillation. *Circulation*
486 *American Heart Association*, 2017; 135:1311–1320.
- 487 40. Von Ende A, Casadei B, Hopewell JC: Improving prediction of atrial fibrillation: the
488 impact of polygenic risk scores over conventional risk factors amongst 270,000
489 individuals in UK Biobank. *Eur Heart J* 2020; 41:ehaa946.0491.
- 490 41. Saracyn M, Kisiel B, Bachta A, et al.: Value of multilocus genetic risk score for atrial
491 fibrillation in end-stage kidney disease patients in a Polish population. *Sci Rep Nature*
492 *Publishing Group*, 2018; 8:9284.
- 493 42. Kloosterman M, Santema BT, Roselli C, et al.: Genetic risk and atrial fibrillation in
494 patients with heart failure. *Eur J Heart Fail* 2020; 22:519–527.
- 495 43. Banerjee A, Dashtban A, Chen S, et al.: Identifying subtypes of heart failure from three
496 electronic health record sources with machine learning: an external, prognostic, and
497 genetic validation study. *Lancet Digit Health* 2023; 5:e370–e379.
- 498 44. Kertai MD, Mosley JD, He J, et al.: Predictive accuracy of a polygenic risk score for
499 postoperative atrial fibrillation after cardiac surgery. *Circ Genomic Precis Med American*
500 *Heart Association*, 2021; 14:e003269.
- 501 45. Aittokallio J, Kauko A, Vaura F, et al.: Polygenic risk scores for predicting adverse
502 outcomes after coronary revascularization. *Am J Cardiol* 2022; 167:9–14.

- 503 46. Marston NA, Garfinkel AC, Kamanu FK, et al.: A polygenic risk score predicts atrial
504 fibrillation in cardiovascular disease. *Eur Heart J* 2023; 44:221–231.
- 505 47. Choe W-S, Kang JH, Choi E-K, et al.: A genetic risk score for atrial fibrillation predicts the
506 response to catheter ablation. *Korean Circ J* 2019; 49:338–349.
- 507 48. Vogel S, Rudaka I, Rots D, et al.: A higher polygenic risk score is associated with a higher
508 recurrence rate of atrial fibrillation in direct current cardioversion-treated patients.
509 *Medicina (Mex) Multidisciplinary Digital Publishing Institute*, 2021; 57:1263.
- 510 49. Al-Kaisey A, Wong GR, Young P, et al.: Polygenic risk scores are associated with atrial
511 electrophysiologic substrate abnormalities and outcomes after atrial fibrillation
512 catheter ablation. *Heart Rhythm* 2023; 20:1188–1194.
- 513 50. Kany S, Al-Taie C, Roselli C, et al.: Association of genetic risk and outcomes in patients
514 with atrial fibrillation: interactions with early rhythm control in the EAST-AFNET4 trial.
515 *Cardiovasc Res* 2023; 119:1799–1810.
- 516 51. Freedman B, Potpara TS, Lip GYH: Stroke prevention in atrial fibrillation. *The Lancet*
517 Elsevier, 2016; 388:806–817.
- 518 52. Lubitz SA, Parsons OE, Anderson CD, et al.: Atrial fibrillation genetic risk and ischemic
519 stroke mechanisms. *Stroke American Heart Association*, 2017; 48:1451–1456.
- 520 53. Pulit SL, Weng L-C, McArdle PF, et al.: Atrial fibrillation genetic risk differentiates
521 cardioembolic stroke from other stroke subtypes. *Neurol Genet Wolters Kluwer Health*,
522 Inc. on behalf of the American Academy of Neurology, 2018; 4:e293.

- 523 54. Marston NA, Patel PN, Kamanu FK, et al.: Clinical application of a novel genetic risk
524 score for ischemic stroke in patients with cardiometabolic disease. *Circulation American*
525 *Heart Association*, 2021; 143:470–478.
- 526 55. O’Sullivan JW, Shcherbina A, Justesen JM, et al.: Combining clinical and polygenic risk
527 improves stroke prediction among individuals with atrial fibrillation. *Circ Genomic*
528 *Precis Med American Heart Association*, 2021; 14:e003168.
- 529 56. Cárcel-Márquez J, Muiño E, Gallego-Fabrega C, et al.: A polygenic risk score based on a
530 cardioembolic stroke multitrait analysis improves a clinical prediction model for this
531 stroke subtype. *Front Cardiovasc Med* 2022; 9:940696.
- 532 57. Choi SH, Jurgens SJ, Weng L-C, et al.: Monogenic and polygenic contributions to atrial
533 fibrillation risk: results from a national biobank. *Circ Res American Heart Association*,
534 2020; 126:200–209.
- 535 58. Fitipaldi H, Franks PW: Ethnic, gender and other sociodemographic biases in genome-
536 wide association studies for the most burdensome non-communicable diseases: 2005–
537 2022. *Hum Mol Genet* 2023; 32:520–532.
- 538 59. Tanigawa Y, Qian J, Venkataraman G, et al.: Significant sparse polygenic risk scores
539 across 813 traits in UK Biobank. *PLOS Genet Public Library of Science*, 2022;
540 18:e1010105.
- 541 60. Kwon O-S, Hong M, Kim T-H, et al.: Genome-wide association study-based prediction of
542 atrial fibrillation using artificial intelligence. *Open Heart* 2022; 9:e001898.

543