Full title: Polygenic Risk Scores Predict Hypertension Onset and Cardiovascular Risk

Short title: Polygenic risk scores predict hypertension risk

Authors: Felix Vaura[a], Anni Kauko[a], Karri Suvila[a,b], Aki S. Havulinna[c,d], Nina Mars[d], Veikko Salomaa[c], FinnGen, Susan Cheng[e,f], Teemu Niiranen[a,b,c]

Affiliations:

a. Department of Internal Medicine, University of Turku, Turku, Finland
b. Division of Medicine, Turku University Hospital, Turku, Finland
c. Department of Public Health Solutions, Finnish Institute for Health and Welfare, Helsinki, Finland
d. Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland
e. Smidt Heart Institute, Cedars-Sinai Medical Center, Los Angeles, CA
f. Division of Cardiology, Brigham and Women's Hospital, Boston, MA

Correspondence: Felix Vaura, Department of Internal Medicine; address: Kiinamyllynkatu 10, 20014 University of Turku, Finland; email: fechva@utu.fi; phone: +358 29 450 5000; fax: +358 29 450 5040.

Abstract

Although genetic risk scores have been used to predict hypertension, their utility in the clinical setting remains uncertain. Our study comprised N=218 792 FinnGen participants (mean age 58 years, 56% women) and N=22 624 well phenotyped FINRISK participants (mean age 50 years, 53% women). We used public genome-wide association data to compute polygenic risk scores (PRSs) for systolic and diastolic blood pressure (BP). Using time-to-event analysis, we then assessed (1) the association of BP PRSs with hypertension and cardiovascular disease (CVD) in FinnGen and (2) the improvement in model discrimination when combining BP PRSs with the validated 4- and 10-year clinical risk scores for hypertension and CVD in FINRISK. In FinnGen, compared with having a 20 to 80 percentile range PRS, a PRS in the highest 2.5% conferred 2.3-fold (95% CI, 2.2–2.4) risk of hypertension and 10.6 years (95% CI, 9.9–11.4) earlier hypertension onset. In subgroup analyses, this risk was only 1.6-fold (95% CI, 1.5–1.7) for late-onset hypertension (age ≥55 years) but 2.8-fold (95% CI, 2.6–2.9) for early-onset hypertension (age <55 years). Elevated systolic BP PRS also conferred 1.3-fold (95% CI, 1.2–1.4) risk of CVD and 2.3 years (95% CI, 1.6–3.1) earlier onset. In FINRISK, systolic and diastolic BP PRSs improved clinical risk prediction of hypertension (but not CVD), increasing the C statistics by 0.7% (95% CI, 0.3–1.1). We demonstrate that genetic information improves hypertension risk prediction. BP PRSs together with traditional risk factors could improve prediction of hypertension and particularly early-onset hypertension, which confers substantial CVD risk.

Keywords: association, blood pressure, genetics, cardiovascular diseases, genetics, hypertension, risk factors.

Introduction

Hypertension affects over a billion[1] people worldwide and is a central risk factor for cardiovascular disease (CVD), the leading cause of death globally.[2] To prevent CVD and its complications, a clinician must be able to accurately assess a patient's risk for developing hypertension. Although lifestyle factors predict hypertension, such behavioral factors are variably subject to measurement and control. By contrast, germline genetic factors are fixed and are known to contribute an identifiable as well as important additive risk. We and others have shown that earlier hypertension onset in parents is strongly associated with hypertension in offspring,[3] and twin studies estimate the heritability of hypertension at 50% to 60%.[4] Although parental hypertension is included in the Framingham hypertension risk score,[5] the sensitivity of self-reported parental hypertension is only 68%,[6] making it an unreliable proxy for genetic information in the clinical setting.

Several studies have successfully used genetic information in isolation to predict hypertension onset by constructing genetic risk scores (GRSs) from genetic variants associated with blood pressure (BP).[7–13] A GRS for hypertension aggregates the statistically significant single-nucleotide polymorphisms (SNPs) from genome-wide association studies (GWAS) into a single predictor. This GRS then has a stronger association with hypertension than any single SNP. However, attempts to improve the predictions of existing clinical risk scores for hypertension with BP GRSs have been unsuccessful.[9–11] One likely reason is low statistical power: in previous studies the number of SNPs used for the GRS ranged from 4 to 32, which represent <0.01% of the average number of >1 million SNP associations reported per GWAS. However, novel

polygenic risk scores (PRS) combine the effect sizes of millions of SNPs regardless of statistical significance and have recently been used to improve existing clinical risk estimates of common diseases such as type 2 diabetes and breast cancer.[14]

Successful integration of genetic information into clinical risk prediction of hypertension is long overdue. We aimed to quantify the predictive ability of novel BP PRSs in hypertension, including early-onset hypertension, and demonstrate that they improve a clinical risk score of hypertension in the general population. Similarly, we quantified the predictive ability of BP PRSs in CVD.

Methods

Study Sample

Because of the sensitive nature of the data collected for this study, requests to access the data set from qualified researchers trained in human subject confidentiality protocols may be submitted through the Finnish Biobanks' FinnBB portal (https://finbb.fi/) for FinnGen and at https://www.thl.fi/biobank/researchers for FINRISK.

The latest data freeze of the Finnish FinnGen study (Data Freeze 5, spring 2020) includes 218 792 genotyped participants (56% women, mean age 58 years) with samples collected from biobanks and prospective epidemiological surveys.[14] Because every Finnish permanent resident is linked to national hospital discharge (from 1968), cause of death (from 1969), and medical expense reimbursement registries (from

1987), follow-up is possible for all major clinical end points, including hypertension and CVD.

First, we used data from all FinnGen participants to assess the independent predictive value of BP PRSs in hypertension and CVD. Then, we used a well-phenotyped subset of FinnGen, the prospective epidemiological FINRISK surveys (N=22 624; years 1997, 2002, 2007, and 2012), to assess the predictive value of BP PRSs relative to clinical risk scores of hypertension and CVD. Baseline data in FINRISK include anthropometric and BP measurements, blood samples, and self-reported questionnaires, and the methodology has been previously described in detail.[15] In every analysis, we excluded individuals with missing values in any variables or prevalent diseases. In analyses using the clinical risk score for CVD, we additionally excluded individuals under 40 years old (see *Clinical Risk Scores*). For individuals present in more than one FINRISK study, we included data entries with the longest follow-ups.

The Coordinating Ethical Committee of the Hospital District of Helsinki and Uusimaa approved both FinnGen and FINRISK study protocols. All participants gave informed written consent. BP measurement methods are described in the **Data Supplement**.

Disease End Points

We studied the incidence of 4 diseases: hypertension, coronary heart disease (CHD), stroke, and CVD (defined as CHD or stroke), whose diagnoses were based on the *International Classification of Diseases (ICD) 8*, *ICD-9*, and *ICD-10* (see **Data Supplement**).

We excluded individuals with prevalent disease events, defined as those diagnosed before the baseline examination. In FINRISK analyses, we combined baseline data and registry data to define prevalent hypertension as: systolic BP (SBP) ≥140 mm Hg OR diastolic BP (DBP) ≥90 mm Hg OR antihypertensive medication use within the last 7 days OR registry-based hypertension. In FinnGen analyses, we reasoned that the baseline examination date was at birth because all covariates (sex and genetic information) stay unchanged over the lifetime follow-up of an individual.

In all survival analyses, we censored individuals at death or at the end of follow-up on December 31, 2018. We had the following number of incident cases and total individuals (cases/total) available for each end point in FinnGen: 55 917/218 754 for hypertension, 29 350/218 792 for CVD, 21 012/218 792 for CHD, and 11 734/212 866 for stroke. The corresponding numbers in FINRISK were as follows: 725/9906 for hypertension, 1647/12 889 for CVD, 1136/13 098 for CHD, and 746/13 355 for stroke. For subgroup analyses in FinnGen, we divided hypertension into early-onset hypertension (age <55 years; 27 361 cases) and late-onset hypertension (age ≥55 years; 28 556 cases).

Genotyping and Imputation

We genotyped FinnGen samples with Illumina and Affymetrix arrays and generated genotype calls with zCall or GenCall (for Illumina) and AxiomGT1 (for Affymetrix) at the Institute for Molecular Medicine Finland (FIMM). We performed quality control exclusions sample-wise: ambiguous gender, missingness >5%, heterozygosity >4 SD,

or non-European ancestry; and variant-wise: missingness >2%, Hardy-Weinberg equilibrium $P<1\times10^{-6}$, minor allele count <3 (for zCall) or <10 (for GenCall). After quality control, we first prephased the samples with Eagle 2.3.5 and then imputed genotypes with Beagle 4.1 (version 08Jun17.d8b, protocol described elsewhere[16]) using a Finnish population-specific SISu[17] v3 reference panel. Finally, to account for population structure in downstream analyses, we performed genetic principal component analysis (PCA) using a pruned set of SNPs of unrelated individuals. Detailed documentation of genotyping, imputation, and principal component analysis is available online.[18]

Polygenic Risk Scores

We computed PRSs for SBP and DBP using the PRS-CS[19] pipeline with default parameters. PRS-CS computes SNP effect sizes by high-dimensional Bayesian regression using GWAS summary statistics and a linkage disequilibrium reference panel. We used publicly available[20] GWAS summary statistics from the UK Biobank[21] based on 340,000 individuals (independent from FinnGen) and a European linkage disequilibrium reference panel with 1.1 million variants derived from samples of the 1000 Genomes Project.[22] The SBP and DBP PRSs were based on 1 098 015 genetic variants common in the linkage disequilibrium reference panel and FinnGen.

To compare our PRSs with previous studies, we also computed GRSs for SBP and DBP based on the variants and weights provided in the article by Evangelou et al.[23] We excluded variants not found in FinnGen (37) and variants with ambiguous allele pairs C/G or A/T (126), using 723 out of the original 886 variants for the GRS. Choosing

proxies for ambiguous variants was not feasible due to inconsistent strand alignments between FinnGen and the GRS variants.

Clinical Risk Scores

For hypertension, we extracted the 4-year predicted clinical risk from Cox proportional hazards models fitted in FINRISK. As predictors, we included the clinical risk factors of the previously validated[24] Framingham 4-year risk score[5] for near-term incidence of hypertension: age, sex, SBP, DBP, body mass index, and current smoking, as well as diabetes. We made the following modifications to the original Framingham score: we included diabetes because it was not considered in the original Framingham score due to the low number of participants with diabetes,[5] and we did not include parental hypertension as a covariate because it was not available in FINRISK.

For atherosclerotic CVD, we evaluated the 10-year clinical risk in FINRISK with the pooled cohort equations according to guidelines of the American College of Cardiology/American Heart Association,[25] using age, sex, total cholesterol, high-density lipoprotein cholesterol, SBP, antihypertensive medication, diabetes, and current smoking as predictors. We then used the 10-year atherosclerotic CVD risk for CVD, CHD, and stroke. Because the American College of Cardiology/American Heart Association risk score was derived in a cohort aged 40 to 79, we excluded individuals under 40 years old for these analyses. While a Finnish risk score for CVD also exists, its training sample includes FINRISK cohorts 2002 and 2007, making it unsuitable for our study due to the risk of overfitting.

Statistical Analyses

In FinnGen, we used Cox proportional hazards models (R package survival[26] and survminer[27]) with age as the time scale to measure the association between BP PRSs and disease end points. We categorized every PRS into 5 bins based on percentiles (<2.5, 2.5—20, 20—80, 80—97.5, and >97.5) and used the largest 20% to 80% bin as the reference. We adjusted all Cox models by sex, collection year, genotyping batch, and the first 10 genetic principal components. Due to the large sample sizes of >200 000, we validated the proportional hazards assumptions by visually inspecting log-minus-log plots.[28] In addition to Cox models, we estimated the differences in age at disease onset between PRS categories by restricted mean survival time models (RMST; R package survRM2[29]). To assess collinearity between the SBP PRS and the DBP PRS, we computed their Pearson correlation.

In FINRISK, to quantify the added value of BP PRSs in clinical risk prediction, we first calculated the 4-year clinical risk for hypertension and the 10-year clinical risk for CVD using Cox models with age as the time scale. We then added continuous BP PRSs in the models as independent predictors and recalculated the clinical risks. For hypertension, we added BP PRSs (SBP, DBP, or both SBP and DBP), while for CVD, we added the SBP PRS. Finally, we compared the risk predictions of the models with and without PRSs for each disease end point (hypertension, CVD, CHD, and stroke) using Harrell C statistics[30] (R package compareC[31]) and the 2-category Net Reclassification Improvement[32] (R package PredictABEL[33]) with risk categories <7.5% and ≥7.5%. We checked proportional hazards assumptions by visually inspecting scaled Schoenfeld residuals[34] (R function ggcoxzph in survminer[27]). We investigated model

calibration with calibration plots and the Hosmer-Lemeshow goodness-of-fit test[35] (methodology by Cook and Ridker[36]). We adjusted each Cox model by cohort year, genotyping batch, and the first 4 genetic principal components.

To compare our BP PRSs with BP GRSs of Evangelou et al,[23] we performed the following calculations in FINRISK: difference in SBP between the upper and lower deciles of the SBP risk scores, Pearson correlations between the PRSs and the GRSs, as well as the difference in C-statistics between a clinical risk prediction model using GRSs and one using both GRSs and PRSs.

We used R 3.6.3 (R Core Team 2020) for all computations and considered a 2-sided $P<0.05$ statistically significant.

Results

BP PRSs and Hypertension Risk

In FinnGen, there were 218 754 individuals (mean age 58 years, 56% women) and 55 917 cases of hypertension. An increasing BP PRSs was associated with a higher hypertension incidence (**Figure 1**). The hazard ratios (HRs; 95% CI) per 1 SD increase in PRSs were 1.42 (1.41—1.43) for SBP and 1.41 (1.40—1.42) for DBP. For the top 2.5% SBP and DBP PRS categories, the HRs for hypertension were 2.19 and 2.26, respectively (**Table 1**). In these categories, hypertension was diagnosed 10.6 and 10.5 years earlier than in the 20% to 80% category (**Figure 2**). Similarly, for the bottom 2.5% PRS categories, the HRs for hypertension were 0.47 and 0.57, and hypertension was

diagnosed 8.6 and 6.7 years later than in the 20% to 80% category. P-values were $<1\times10^{-32}$ for all estimates. The Pearson correlation between the SBP and DBP PRSs was 0.62 (0.61—0.63).

Early-onset hypertension showed stronger associations to SBP and DBP PRSs than late-onset hypertension (**Table 1**). For early-onset hypertension, the HRs (95% CI) per 1 SD increase in PRSs were 1.54 (1.53—1.56) for SBP and 1.58 (1.56—1.60) for DBP, whereas for late-onset hypertension they were 1.31 (1.29—1.32) for SBP and 1.26 (1.25—1.28) for DBP. P-values were $<1\times10^{-292}$ for all estimates.

In FINRISK, there were 9906 individuals (mean age 46 years, 60% women) and 725 cases of hypertension. After including BP PRSs in the clinical risk prediction model for hypertension, the C statistics (%) increased from 79.7 with SBP and DBP PRSs: by 0.5 (0.01—0.9; P=0.016) with SBP PRS and by 0.6 (0.3—1.0; $P=8\times10^{-4}$) with DBP PRS (**Table 2**). Including both SBP and DBP PRSs in the clinical risk prediction model gave the largest increase in the C statistics (%): 0.7 (0.3—1.1; P=0.0017). The 2-category net reclassification improvement was not statistically significant for any BP PRS.

SBP PRS and Cardiovascular Risk

Higher SBP PRS was associated with a higher incidence of all 3 CVD end points (CVD, CHD, and stroke; **Figure 3**). The HR (95% CI) per 1 SD increase in SBP PRS was 1.13 (1.12—1.15) in CVD, 1.15 (1.13—1.17) in CHD, and 1.11 (1.09—1.13) in stroke. For the top 2.5% SBP PRS category, HRs (for CVD, CHD, and stroke) were 1.30, 1.33, and 1.29 (**Table 3**), and disease onset occurred 2.3, 2.0, and 1.4 years earlier compared

with the average 20% to 80% SBP PRS category (**Figure S1** in the **Data Supplement**). Similarly, for the bottom 2.5% SBP PRS category, HRs were 0.74, 0.69, and 0.79, and disease onset occurred 2.4, 2.4, and 1.1 years later compared with the average 20% to 80% SBP PRS category. *P* values ranged from 4×10^{-95} to 6×10^{-4}, with stroke having the weakest estimates.

After including SBP PRS in the clinical risk prediction models, the C statistics did not change in any of the CVD end points. The 2-category NRI was not statistically significant, either (**Table S1**). Calibration plots for Cox models containing both clinical risk scores and PRSs indicated good calibration (**Figure S2**).

BP PRS Versus GRS

The sex-adjusted difference in SBP between the top and bottom deciles was 14.1 mm Hg (95% CI, 13.0—15.2) for the SBP PRS and 10.6 mm Hg (95% CI, 9.5—11.7) for the SBP GRS. The Pearson correlation between the PRSs and GRSs was 0.46 (0.44—0.47) for SBP and 0.41 (0.39—0.42) for DBP. After including BP PRSs in addition to BP GRSs into a clinical risk prediction model for hypertension, the C statistic increased by 0.4% (0.1—0.6; *P*=0.0020).

Discussion

We demonstrated that BP PRSs predict hypertension and improve the Framingham model for near-term incidence of hypertension. The predictive ability of BP PRSs was particularly strong for early-onset hypertension. The SBP PRS independently predicted

CVD but failed to improve the American College of Cardiology/American Heart Association clinical risk prediction model.

Previous research has shown that early-onset hypertension is associated with increased risk of cardiovascular death and target organ damage compared with late-onset hypertension.[3,37] While the heritability of hypertension is well established,[38] only early-onset hypertension in parents has been reliably shown to associate with hypertension in offspring,[3] suggesting a more robust genetic component in early-onset than in late-onset hypertension. In our study, a 1 SD increase of SBP and DBP PRSs resulted in 54% and 58% greater risks of early-onset hypertension, respectively. For late-onset hypertension, the corresponding risk increases were only 31% and 26%. Furthermore, individuals in the top 2.5% of the SBP PRS were diagnosed with hypertension on average 19 years earlier than individuals in the bottom 2.5%, which has been shown to translate to ≈2-fold greater risk of CVD and death.[39] Therefore, our study supports the prominent role of genetics in early-onset hypertension and suggests that high-risk individuals could benefit from detailed collection of parental hypertension history, early genotyping, and more intensive treatment interventions.

Previous investigations of BP PRSs have not detected improvements in clinical risk prediction of hypertension.[9–11] In this study, we used the office-based risk score[5] from the Framingham Heart Study as the clinical risk score of hypertension. We showed that including both SBP and DBP PRSs in the risk prediction model increased the C statistics by 0.7 percentage points from 79.7 % to 80.4 % (P=0.0017; **Table 2**), which demonstrates the added value of PRSs. Furthermore, individuals in the top and bottom 2.5% of the PRSs had 2-fold higher and lower risks of hypertension as compared with

those in the middle quantile. Our results underline the potential of PRSs as a complementary tool alongside traditional clinical risk factors for hypertension prediction. While the utility of PRSs for an average patient in the clinical setting is still debatable, we have taken an essential first step in providing a baseline improvement to clinical risk prediction of hypertension using germline DNA.

While our study focused on the prediction of hypertension, we also assessed the predictive value of the SBP PRS in CVD. SBP PRS was independently associated with a higher incidence of CVD, CHD, and stroke, but did not improve the discriminative value of the existing CVD risk score in any of the end points. Considering that studies on combining CHD PRSs to clinical risk factors have demonstrated varying evidence of utility,[14,40,41] it is no surprise that a PRSs for one CVD risk factor, hypertension, had limited value in CVD risk prediction. Although a BP PRS did not seem to improve CVD risk prediction, subsequent studies should examine the potential of multitrait PRSs for CVD risk prediction in large cohorts.

Several GWASs[8,23,42-46] of BP have been conducted over the past 20 years and many studies[7-13,23] have used these GWASs to combine the genome-wide significant SNPs into a GRS. Most notably, Evangelou et al[23] conducted the largest BP GWAS meta-analysis to date with 760,000 individuals and used the significant 886 SNPs to construct SBP and DBP GRSs. The SBP difference between the top and bottom SBP GRS deciles in their study sample was 12.9 mm Hg.[23] In FINRISK, using modified BP GRSs with 723 available SNPs, the difference was 10.6 mm Hg. Meanwhile, the difference in SBP between the top and bottom SBP PRS deciles in FINRISK was 14.1 mm Hg, representing 9% and 33% increases compared with the GRSs, respectively.

Furthermore, the model discrimination increased significantly even when PRSs were included in a model with GRSs and traditional clinical risk factors. Indeed, PRSs have consistently performed better than their GRS counterparts.[47] While our SBP PRS was based on a GWAS of 340 000 individuals, it used 1.1 million SNPs instead of 886 SNPs. Therefore, our study underlines the value of using the latest methodology for constructing genome-wide PRSs for BP, to better capture common variation for this highly polygenic trait.

Although our study has several strengths, such as a study sample of over 200 000 individuals, a validated registry-based follow-up since 1969, and a novel Bayesian PRS, it also has limitations. First, as we did not have parental history of hypertension available in FINRISK, it could not be included in the clinical risk equation when quantifying the added value of BP PRSs. However, in the clinical setting, parental history of hypertension is always self-reported by the patient. Despite improvements in hypertension awareness, 20% to 50% of individuals with hypertension are themselves unaware of it even in high-income countries.[48] Second, we used GWAS results from only 340 000 individuals (UK Biobank), instead of the large meta-analysis[23] based on 760 000 individuals (UK Biobank and International Consortium for Blood Pressure[7]). We did this to avoid overfitting because some FINRISK cohorts are part of the International Consortium for Blood Pressure consortium. Third, since both our study and the BP GWAS that we used for the PRSs comprised individuals of European ancestry, the results may not be generalizable to individuals of other ancestries. Fourth, around 36%[14] of FinnGen participants are recruited from hospital biobanks or disease-based cohorts, which may lead to overestimation of absolute, but not relative, disease risk. Fifth, while including both SBP and DBP PRSs in the Cox models can cause concern

for collinearity, correlation coefficients <0.7 between predictor variables do not distort model performance to a great extent.[49] Finally, a registry-based diagnosis of hypertension often underestimates true prevalence and could create bias in the case-control definitions.

Perspectives

We showed that novel BP PRSs predict hypertension better than previous GRSs and, for the first time, improve office-based risk estimates of hypertension. With a one-time cost that is already comparable to common laboratory tests,[50] a patient could be genotyped and have their lifetime genetic risk for hypertension determined with BP PRSs. This information could be used either together with traditional risk factors to improve clinical risk prediction or independently to estimate the lifetime risk of hypertension. While traditional risk factors vary and often worsen over time, BP PRSs could be used even when traditional risk factors are not measurable. Future research should further develop PRS methodology and quantify the effect of PRS-based genetic risk counseling on health behavior and prevention of hypertension.

Acknowledgements

We thank the participants and investigators of the FinnGen and UK Biobank studies for their invaluable contributions to this work.

Sources of Funding

This work has been funded by the Academy of Finland (321351), the Urmas Pekkala Foundation, the Paavo Nurmi Foundation, the Finnish Foundation for Cardiovascular Research, the Finnish Medical Foundation, the Emil Aaltonen Foundation, and the Hospital District of Southwest Finland. V. Salomaa was supported by the Finnish Foundation for Cardiovascular Research. K. Suvila was supported by grants from the Finnish Foundation for Cardiovascular Research, the Finnish Medical Foundation, the Turku University Foundation, and the University of Turku. S. Cheng was supported by National Institutes of Health grants R01-HL134168, R01-HL131532, R01-HL143227, and R01-HL142983.

Disclosures

VS has received honoraria for consulting from Novo Nordisk and Sanofi. He also has ongoing research collaboration with Bayer Ltd. (All unrelated to the present study).

Supplemental Materials

Supplemental Methods

Online Figures I and II

Online Table S1

References

1. World Health Organization. A global brief on hypertension : silent killer, global public health crisis: World Health Day 2013. World Health Organization Web site. https://apps.who.int/iris/handle/10665/79059. Accessed December 17, 2020.

2. Forouzanfar MH, Alexander L, Anderson HR, Bachman VF, Biryukov S, Brauer M, Burnett R, Casey D, Coates MM, Cohen A, et al. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet*. 2015; 386:2287—2323.

3. Niiranen TJ, McCabe EL, Larson MG, Henglin M, Lakdawala NK, Vasan RS, Cheng S. Heritability and Risks Associated With Early Onset Hypertension: Multigenerational, Prospective Analysis in the Framingham Heart Study. *BMJ*. 2017; 357:j1949.

4. Kupper N, Ge D, Treiber FA, Snieder H. Emergence of novel genetic effects on blood pressure and hemodynamics in adolescence: the Georgia Cardiovascular Twin Study. *Hypertension*. 2006; 47:948—954.

5. Parikh NI, Pencina MJ, Wang TJ, Benjamin EJ, Lanier KJ, Levy D, D'Agostino RB Sr, Kannel WB, Vasan RS. A risk score for predicting near-term incidence of hypertension: the Framingham Heart Study. *Ann Intern Med*. 2008; 148:102—110.

6. France CR, Page GD. Assessing parental history of hypertension: father (and mother) knows best! *Psychophysiology*. 1998;

35:341—343.

7. International Consortium for Blood Pressure Genome-Wide Association Studies, Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, Chasman DI, Smith AV, Tobin MD, Verwoert GC, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature*. 2011; 478:103—109.

8. Warren HR, Evangelou E, Cabrera CP, Gao H, Ren M, Mifsud B, Ntalla I, Surendran P, Liu C, Cook JP, et al. Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. *Nat Genet*. 2017; 49:403—415.

9. Fava C, Sjögren M, Montagnana M, Danese E, Almgren P, Engström G, Nilsson P, Hedblad B, Guidi GC, Minuz P, et al. Prediction of blood pressure changes over time and incidence of hypertension by a genetic risk score in Swedes. *Hypertension*. 2013; 61:319—326.

10. Niiranen TJ, Havulinna AS, Langén VL, Salomaa V, Jula AM. Prediction of Blood Pressure and Blood Pressure Change With a Genetic Risk Score. *J Clin Hypertens*. 2016; 18:181—186.

11. Lim N-K, Lee J-Y, Lee J-Y, Park H-Y, Cho M-C. The Role of Genetic Risk Score in Predicting the Risk of Hypertension in the Korean population: Korean Genome and Epidemiology Study. *PLoS One*. 2015; 10:e0131603.

12. Juhola J, Oikonen M, Magnussen CG, Mikkilä V, Siitonen N, Jokinen E, Laitinen T, Würtz P, Gidding SS, Taittonen L, et al. Childhood physical, environmental, and genetic predictors of adult hypertension: the cardiovascular risk in young Finns study. *Circulation*. 2012; 126:402—409.

13. Havulinna AS, Kettunen J, Ukkola O, Osmond C, Eriksson JG, Kesäniemi YA, Jula A, Peltonen L, Kontula K, Salomaa V, et al. A blood pressure genetic risk score is a significant predictor of incident cardiovascular events in 32,669 individuals. *Hypertension*. 2013; 61:987—994.

14. Mars N, Koskela JT, Ripatti P, Kiiskinen TTJ, Havulinna AS, Lindbohm JV, Ahola-Olli A, Kurki M, Karjalainen J, Palta P, et al. Polygenic and clinical risk scores and their impact on age at onset and prediction of cardiometabolic diseases and common cancers. *Nat Med*. 2020; 26:549—557.

15. Borodulin K, Tolonen H, Jousilahti P, Jula A, Juolevi A, Koskinen S, Kuulasmaa K, Laatikainen T, Männistö S, Peltonen M, et al. Cohort Profile: The National FINRISK Study. *Int J Epidemiol*. 2018; 47:696—696i.

16. Pärn K, Isokallio MA, Fontarnau JN, Palotie A, Ripatti S, Palta P. Genotype imputation workflow v3.0 v1. Protocols.io Web site. https://www.protocols.io/view/genotype-imputation-workflow-v3-0-nmndc5e. Accessed December 17, 2020.

17. Sequencing Initiative Suomi (SISu). SISu Web site. http://sisuproject.fi. Accessed July 14, 2020.

18. FinnGen. FinnGen Documentation of R4 release. Gitbook.io Web site. https://finngen.gitbook.io/documentation/. Accessed December 17, 2020.

19. Ge T, Chen C-Y, Ni Y, Feng Y-CA, Smoller JW. Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat Commun*. 2019; 10:1776.

20. UK Biobank. GWAS Round 2 Results. Neale lab Web site. http://www.nealelab.is/uk-biobank. Accessed December 17, 2020.

21. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med*. 2015; 12:e1001779.

22. 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, et al. A global reference for human genetic variation. *Nature*. 2015; 526:68—74.

23. Evangelou E, Warren HR, Mosen-Ansorena D, Mifsud B, Pazoki R, Gao H, Ntritsos G, Dimou N, Cabrera CP, Karaman I, et al. Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. *Nat Genet*. 2018; 50:1412—1425.

24. Kivimäki M, Batty GD, Singh-Manoux A, Ferrie JE, Tabak AG, Jokela M, Marmot MG, Smith GD, Shipley MJ. Validating the

Framingham Hypertension Risk Score: results from the Whitehall II study. *Hypertension*. 2009; 54:496—501.

25. Goff DC Jr, Lloyd-Jones DM, Bennett G, Coady S, D'Agostino RB, Gibbons R, Greenland P, Lackland DT, Levy D, O'Donnell CJ, et al., American College of Cardiology/American Heart Association Task Force on Practice Guidelines. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2014; 129:S49—73.

26. Therneau TM. survival: a package for survival analysis in R. R package version 3.2-7. https://CRAN.R-project.org/package=survival. Accessed December 17, 2020.

27. Kassambara A, Kosinski M, Biecek P. survminer: Drawing Survival Curves using 'ggplot2'. R package version 0.4.8. https://CRAN.R-project.org/package=survminer. Accessed December 17, 2020.

28. Schemper M. Cox Analysis of Survival Data with Non-Proportional Hazard Functions. *The Statistician*. 1992; 41:455—465.

29. Uno H, Tian L, Horiguchi M, Cronin A, Battioui C, Bell J. survRM2: Comparing Restricted Mean Survival Time. R package version 1.0-3. https://CRAN.R-project.org/package=survRM2. Accessed December 17, 2020.

30. Harrell FE. Evaluating the yield of medical tests. *JAMA*. 1982; 247:2543—2546.

31. Kang L, Chen W, Petrick NA, Gallas BD. Comparing two correlated C indices with right-censored survival outcome: a one-shot

nonparametric approach. *Stat Med*. 2015; 34:685—703.

32. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond. *Stat Med*. 2007; 27:157—172.

33. Kundu S, Aulchenko YS, van Duijn CM, Janssens ACJW. PredictABEL: an R package for the assessment of risk prediction models. *Eur J Epidemiol*. 2011; 26:261—264.

34. Schoenfeld D. Partial residuals for the proportional hazards regression model. *Biometrika*. 1982; 69:239—241.

35. Hosmer DW, Hosmer T, Le Cessie S, Lemeshow S. A comparison of goodness-of-fit tests for the logistic regression model. *Stat Med*. 1997; 16:965—980.

36. Cook NR, Ridker PM. Advances in Measuring the Effect of Individual Predictors of Cardiovascular Risk: The Role of Reclassification Measures. *Ann Intern Med*. 2009; 150:795—802.

37. Suvila K, McCabe EL, Lehtonen A, Ebinger JE, Lima JAC, Cheng S, Niiranen TJ. Early Onset Hypertension Is Associated With Hypertensive End-Organ Damage Already by MidLife. *Hypertension*. 2019; 74:305—312.

38. Miall WE, Oldham PD. The Hereditary Factor in Arterial Blood-pressure. *BMJ*. 1963; 1:75—80.

39. Wang C, Yuan Y, Zheng M, Pan A, Wang M, Zhao M, Li Y, Yao S, Chen S, Wu S, et al. Association of Age of Onset of Hypertension With Cardiovascular Diseases and Mortality. *J Am Coll Cardiol*. 2020; 75:2921—2930.

40. Inouye M, Abraham G, Nelson CP, Wood AM, Sweeting MJ, Dudbridge F, Lai FY, Kaptoge S, Brozynska M, Wang T, et al. Genomic Risk Prediction of Coronary Artery Disease in 480,000 Adults: Implications for Primary Prevention. *J Am Coll Cardiol*. 2018; 72:1883—1893.

41. Abraham G, Havulinna AS, Bhalala OG, Byars SG, De Livera AM, Yetukuri L, Tikkanen E, Perola M, Schunkert H, Sijbrands EJ, et al. Genomic prediction of coronary heart disease. *Eur Heart J*. 2016; 37:3267—3278.

42. Hoffmann TJ, Ehret GB, Nandakumar P, Ranatunga D, Schaefer C, Kwok P-Y, Iribarren C, Chakravarti A, Risch N. Genome-wide association analyses using electronic health records identify new loci influencing blood pressure variation. *Nat Genet*. 2017; 49:54—64.

43. Levy D, DeStefano AL, Larson MG, O'Donnell CJ, Lifton RP, Gavras H, Cupples LA, Myers RH. Evidence for a gene influencing blood pressure on chromosome 17. Genome scan linkage results for longitudinal blood pressure phenotypes in subjects from the Framingham heart study. *Hypertension*. 2000; 36:477—483.

44. Levy D, Ehret GB, Rice K, Verwoert GC, Launer LJ, Dehghan A, Glazer NL, Morrison AC, Johnson AD, Aspelund T, et al.

Genome-wide association study of blood pressure and hypertension. *Nat Genet*. 2009; 41:677—687.

45. Newton-Cheh C, Johnson T, Gateva V, Tobin MD, Bochud M, Coin L, Najjar SS, Zhao JH, Heath SC, Eyheramendy S, et al. Genome-wide association study identifies eight loci associated with blood pressure. *Nat Genet*. 2009; 41:666—676.

46. Azam AB, Azizan EAB. Brief Overview of a Decade of Genome-Wide Association Studies on Primary Hypertension. *Int J Endocrinol*. 2018; 2018:7259704.

47. Choi SW, Mak TS-H, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. *Nat Protoc*. 2020; 15:2759—2772.

48. NCD Risk Factor Collaboration (NCD-RisC). Long-term and recent trends in hypertension awareness, treatment, and control in 12 high-income countries: an analysis of 123 nationally representative surveys. *Lancet*. 2019; 394:639—651.

49. Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G, Marquéz JRG, Gruber B, Lafourcade B, Leitão PJ et al. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*. 2013; 36:27—46.

50. Tam V, Patel N, Turcotte M, Bossé Y, Paré G, Meyre D. Benefits and limitations of genome-wide association studies. *Nat Rev Genet*. 2019; 20:467—484.

Novelty and Significance

What Is New?

- Although genetic risk scores have been used to predict hypertension, their utility in the clinical setting remains uncertain.
- We studied the association of polygenic risk scores for blood pressure with incident hypertension in >200,000 individuals.

What Is Relevant?

- Compared to having a 20 to 80 percentile range PRS, a PRS in the highest 2.5% conferred 2.3-fold risk of hypertension.
- SBP and DBP PRSs improved clinical risk prediction of hypertension, increasing the C statistics by 0.7%

Summary

Our findings demonstrate that combining genetic information with traditional risk factors improves the accuracy of hypertension risk assessment. Furthermore, individual genetic makeup is particularly strongly linked to hypertension that is diagnosed before midlife.

Figure 1. Cumulative risk of hypertension by polygenic risk score categories in FinnGen. The survival curves are from Cox proportional hazards models. There were 218 754 individuals with 55 917 cases of hypertension. We adjusted the models for sex, collection year, genotyping batch, and the first 10 genetic principal components. DBP indicates diastolic blood pressure; and SBP, systolic blood pressure.

Figure 2. Difference in age of hypertension onset estimates across blood pressure polygenic risk score categories in FinnGen. There were 218 754 individuals with 55 917 cases of hypertension. The estimates are restricted mean survival times for age at onset, and the error bars represent their 95% confidence intervals. DBP indicates diastolic blood pressure; and SBP, systolic blood pressure.

Figure 3. Cumulative risk of cardiovascular disease by systolic blood pressure PRS category in FinnGen. The survival curves are from Cox proportional hazards models. There were 218 792 individuals with 29 350 cases of CVD, 218 792 individuals with 21 012 cases of CHD, and 212 884 individuals with 11 734 cases of stroke. We adjusted the models for sex, collection year, genotyping batch, and the first 10 genetic principal components.

Table 1. HR of Hypertension Onset Between BP PRS Categories in Adjusted Cox Models (FinnGen).

	Hypertension			Early-onset hypertension			Late-onset hypertension		
PRS	**HR (95% CI)**	***P* value**	**Cases/controls**	**HR (95% CI)**	***P* value**	**Cases/controls**	**HR (95% CI)**	***P* value**	**Cases/controls**
SBP			55 917/162 837			27 361/191 393			28 556/190 198
<2.5%	0.44 (0.41—0.47)	9×10^{-107}	725/4744	0.39 (0.35—0.44)	2×10^{-52}	269/5200	0.47 (0.43—0.52)	8×10^{-57}	456/5013
2.5-20%	0.64 (0.62—0.65)	5×10^{-255}	6 911/31 367	0.55 (0.52—0.57)	3×10^{-177}	2 603/35 675	0.70 (0.68—0.73)	3×10^{-94}	4308/33 970
20-80%	1 (reference)	…	33 176/98 078	1 (reference)	…	15 658/115 596	1 (reference)	…	17 518/113 736
80-97.5%	1.54 (1.51—1.58)	$<1\times10^{-300}$	12 846/25 437	1.70 (1.66—1.75)	$<1\times10^{-300}$	7348/30 935	1.37 (1.33—1.41)	2×10^{-91}	5498/32 785
>97.5%	2.19 (2.10—2.29)	1×10^{-281}	2259/3211	2.62 (2.48—2.77)	3×10^{-271}	1483/3987	1.68 (1.57—1.81)	3×10^{-45}	776/4694
DBP			55 917/162 837			27 361/191 393			28 556/190 198
<2.5%	0.49 (0.46—0.53)	4×10^{-88}	822/4645	0.37 (0.33—0.42)	3×10^{-55}	256/5211	0.57 (0.53—0.63)	7×10^{-38}	566/4901
2.5-20%	0.67 (0.65—0.69)	2×10^{-203}	7315/30 968	0.56 (0.54—0.59)	7×10^{-162}	2675/35 608	0.75 (0.73—0.78)	2×10^{-64}	4640/33 643
20-80%	1 (reference)	…	32 873/98 375	1 (reference)	…	15 527/115 721	1 (reference)	…	17 346/113 902
80-97.5%	1.55 (1.52—1.58)	$<1\times10^{-300}$	12 662/25 624	1.73 (1.68—1.78)	$<1\times10^{-300}$	7351/30 935	1.35 (1.31—1.39)	5×10^{-80}	5311/32 975
>97.5%	2.26 (2.17—2.36)	$<1\times10^{-300}$	2245/3225	2.78 (2.64—2.93)	$<1\times10^{-300}$	1552/3918	1.60 (1.48—1.73)	3×10^{-33}	693 /4777

We defined early-onset and late-onset hypertension as age of onset <55 years and ≥55 years, respectively. We adjusted the Cox proportional hazards models for sex, collection year, genotyping batch, and the first 10 genetic principal components. DBP indicates diastolic blood pressure; HR, hazard ratio; PRS, polygenic risk score; and SBP, systolic blood pressure.

Table 2. Changes in C-index (%) and NRI (%) Assessed After Including Blood Pressure PRSs in the Clinical Risk Prediction Model (C=79.7) for Hypertension Onset (FINRISK)

	C-index (clinical risk factors + PRS)			NRI (cutoff at 7.5%)	
PRS	**C**	**Change* (95% CI)**	***P* value**	**NRI (95% CI)**	***P* value**
SBP	80.2	0.5 (0.1 to 0.9)	0.024	-0.2 (-2.0 to 1.6)	0.79
DBP	80.3	0.6 (0.3 to 1.0)	8×10^{-4}	0.2 (-1.4 to 1.9)	0.79
SBP and DBP	80.4	0.7 (0.3 to 1.1)	0.0017	-0.4 (-2.2 to 1.4)	0.66

There were 9906 individuals with 725 incident cases of hypertension. We included each blood pressure PRS as an independent covariate in a Cox model containing the following clinical risk factors for hypertension: age, sex, SBP, DBP, BMI, diabetes, and current smoking. We also included SBP PRS and DBP PRS together (SBP and DBP). BMI indicates body mass index; C-index, Harrell concordance index; DBP, diastolic blood pressure; NRI, net reclassification improvement; PRS, polygenic risk score; and SBP, systolic blood pressure. *Percentage point change in the C-index after we included the blood pressure PRSs in the clinical risk prediction model for hypertension.

Table 3. HR for Cardiovascular End Points Between Systolic Blood Pressure PRS Categories in Adjusted Cox Models (FinnGen).

End point	HR (95% CI)	*P* value	Cases/controls
CVD			29 350 / 189 442
<2.5%	0.74 (0.68—0.80)	4×10^{-13}	579 / 4891
2.5%—20%	0.84 (0.81—0.87)	2×10^{-24}	4474 / 33 815
20%—80%	1 (reference)	-	17 605 / 113 669
80%—97.5%	1.16 (1.13—1.20)	3×10^{-23}	5780 / 32 509
>97.5%	1.30 (1.22—1.39)	1×10^{-14}	912 / 4558
CHD			21 012 / 197 780
<2.5%	0.69 (0.62—0.76)	5×10^{-13}	381 / 5089
2.5%—20%	0.84 (0.80—0.87)	3×10^{-19}	3130 / 35 159
20%—80%	1 (reference)	-	12 572 / 118 702
80%—97.5%	1.20 (1.15—1.24)	1×10^{-23}	4253 / 34 036
>97.5%	1.33 (1.23—1.44)	5×10^{-13}	676 / 4794
Stroke			11 734 / 201 132
<2.5%	0.79 (0.70—0.90)	4×10^{-4}	246 / 5 095
2.5%—20%	0.86 (0.82—0.91)	3×10^{-8}	1834 / 35 480
20%—80%	1 (reference)	-	7061 / 120 602
80%—97.5%	1.11 (1.06—1.17)	2×10^{-5}	2228 / 35 010
>97.5%	1.29 (1.16—1.44)	2×10^{-6}	365 / 4945

We adjusted the Cox models for sex, collection year, genotyping batch, and the first 10 genetic principal components. CHD indicates coronary heart disease; CVD, cardiovascular disease (CHD or stroke); HR, hazard ratio; and PRS, polygenic risk score (for systolic blood pressure).