Stroke

CLINICAL AND POPULATION SCIENCES

A Family and a Genome-Wide Polygenic Risk Score Are Independently Associated With Stroke in a Population-Based Study

Michelle Hämmerle, BSc; Lukas Forer, PhD; Sebastian Schönherr, PhD; Annette Peters, PhD; Harald Grallert, PhD; Florian Kronenberg, MD; Christian Gieger, PhD*; Claudia Lamina, PhD*

BACKGROUND: Positive family history and genetic risk scores have been shown to independently capture those individuals with high risk for stroke. The aim of our study was to evaluate the amount of shared information between family history and genetic risk and to investigate their combined effect on the association with prevalent and incident stroke cases.

METHODS: We obtained a family risk score (FamRS), weighted for disease onset and family size as well as genome-wide polygenic risk score (PGS) including over 3.2 million single-nucleotide polymorphisms in the population-based prospective KORA F3 (Cooperative Health Research in the Region of Augsburg) study (n=3071) from Southern Germany. FamRS and PGS were evaluated separately and combined. The measures were once treated as continuous variables but also divided in the highest 20%, 10%, 5%, and 1% percentiles. Odds ratios via logistic regression and hazard ratios via Cox regression were estimated. A stroke event was defined as a hospitalization for stroke that was self-reported in a standardized interview by certified and supervised personnel.

RESULTS: The FamRS outperformed other simplified family measures such as affected parents or number of affected family members. FamRS and PGS were not correlated, and no individuals were observed with both very high FamRS and very high PGS (top 1% percentile). In a combined model, both FamRS and PGS were independently from each other associated with risk of stroke, also independent of other traditional risk factors (p [FamRS]=0.02, p [PGS]=0.005). Individuals in the top 1% of either FamRS or PGS were found to have >5-fold risk for stroke (odds ratios, 5.82 [95% CI, 2.08–14]; *P*=0.0002). The results for incident stroke events showed the same trend but were not significant.

CONCLUSIONS: Our study shows that a family risk score and PGS capture different information concerning individual stroke risk. Combining the risk measures FamRS and PGS increases predictive power, as demonstrated in a population-based study.

GRAPHIC ABSTRACT: A graphic abstract is available for this article.

Key Words: cause of death ■ prevalence ■ risk factors ■ sibling ■ single-nucleotide polymorphism

Stroke is the second most common cause of death worldwide. Furthermore, it is the paramount reason for neurological disability in adults.[1] It annually accounts for ≈5.5 million deaths worldwide,[2] around 440 000 in Europe. In 2017, the costs rose to 45 billion $ in Europe by including lost production and care costs.[3]

The average lifetime risk of suffering a stroke is 25%.[4] Even though the prevalence of stroke has decreased over the past decades,[1,2] the burden remains high. As low- and middle-income countries adapt to an unhealthier lifestyle[5] and populations grow older, implementing effective preventive strategies is necessary.[2] Previous research indicated that between ≈35% and 70% of all

Correspondence to: Claudia Lamina, PhD, Institute of Genetic Epidemiology, Department of Genetics and Pharmacology, Medical University of Innsbruck, Innsbruck, Austria. Email claudia.lamina@i-med.ac.at
*C. Gieger and C. Lamina contributed equally.
Supplemental Material is available at https://www.ahajournals.org/doi/suppl/10.1161/STROKEAHA.121.036551.
For Sources of Funding and Disclosures, see page 2338.

Stroke is available at www.ahajournals.org/journal/str

Nonstandard Abbreviations and Acronyms

FamRS	family risk score
GRS	genetic risk score
HR	hazard ratio
KORA	Cooperative Health Research in the Region of Augsburg
NRI	net reclassification index
OR	odds ratio
PGS	polygenic score
SNP	single-nucleotide polymorphism

stroke events are preventable, especially for those at high risk.[6–8] Additionally to traditional risk factors, positive family history and genetic risk scores (GRS) could be used to identify those individuals with high risk for stroke.

Early research[9] indicated that family history as represented by parental[10] and sibling history[11,12] might be a risk factor for stroke. They were indeed identified as risk factors independent from traditional risk factors in a study including different populations.[13] A large cohort study in China confirmed the predictive power of accumulated stroke events within families.[14]

A role of genetics in the susceptibility to stroke has been suggested by family and twin studies.[11] Until today, 35 influential genetic loci have been identified via genome-wide association studies.[15] The results of early GRS for stroke, however, were modest and clinical applicability was limited or unclear.[16,17] The predictive power was substantially improved by increasing the number of single-nucleotide polymorphisms (SNPs) within the score and thereby calculating a so-called polygenic score (PGS).[18] A subsequent study demonstrated that this score was an independent risk factor.[19] Genetic scores proved useful also for geriatric individuals with cardiovascular risk factors.[20] Furthermore, the independent association of genetic and lifestyle factors with incident stroke was shown, highlighting the potential reduction of stroke risk by adhering to a healthier lifestyle.[21] To further improve the measurement of genetic risk on an individual level, it has been suggested to combine different GRS related to the same condition.[22] This idea laid the ground for the metaGRS for stroke and was the basis for calculating the PGS in this study.[23]

With emerging availability of reasonably priced SNP-chips and, therefore, usage of GRS, one might ask whether collection of family history is still necessary. However, it was hypothesized that the predictive power might increase via the combination of family history and GRS.[24,25] Most importantly, the independent or combined effect of family history and GRS on risk of stroke has not been evaluated, yet. It was the aim of this study to fill this gap. Therefore, data for both familial and genetic risk were collected and calculated as family risk score (FamRS) and PGS. The risk measures were then applied for evaluating the association with stroke risk in a population-based study.

METHODS

The data used for this analysis are subject to national data protection laws and restrictions were imposed by the ethics committee of the Bavarian Medical Association to ensure data privacy of the study participants. However, they can be applied for through an individual project agreement with KORA (Cooperative Health Research in the Region of Augsburg) using the digital tool KORA.PASST (https://epi.helmholtz-muenchen.de/). Our study is reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology statement (http://www.strobe-statement.org).

Study Population

KORA F3 is a population-based study in Augsburg, Germany, and surrounding counties, enrolling individuals of German nationality aged 35 to 84.[26] Study methods included ascertaining basic sociodemographic information, family and medical history, standardized interviews, and medical examinations.[27]

Between February 2004 and May 2005, 3184 individuals participated in the KORA F3 study[28]; for 3071 individuals, genetic data and family history information was available to obtain the PGS and FamRS. In the follow-up study from 2016, information about morbidity and mortality for 2904 individuals was available (Figure S1).

Definition of Stroke Events

A stroke event was defined as hospitalization due to stroke which was self-reported in a structured and standardized interview performed by certified and supervised personnel. In the interview, there was no distinction between different stroke types. In the 2016 follow-up, prevalent stroke was assessed alike, but for incident stroke the self-reported information was validated. In the validation step, different stroke types and timepoint of stroke events were assessed. More details can be found in the Supplemental Material.

Definition and Calculation of the FamRS

The FamRS is a quantitative measure to ascertain an individual's risk of developing a specific disease.[29] The basis for the calculation is the difference between expected (E) and observed (O) affected family members. The information needed to calculate the FamRS was obtained via a standardized interview. If first-grade relatives (father, mother, brothers, and sisters) suffered a stroke, it was further ascertained whether the disease occurred before or after age 60. The age was taken into account by assigning scores to the different age groups: w=2 for those <60, w=1 for those ≥60, and w=1.5 if the age was unknown. Summing these weights of the affected family members yields the observed values. There was no distinction between different types of stroke. As we described earlier,[28] the expected values for each 10-year age group could be calculated within KORA F3 as this is a population-based study with no specific inclusion criteria other than age range. Taken together, the FamRS basically takes into account the number

 Stroke. 2022;53:2331–2339. DOI: 10.1161/STROKEAHA.121.036551

of affected family members, but additionally weights the age of disease onset and number of relatives. For example, no event in a big family would lead to negative values of the FamRS. One early or 2 events at any age within a small or averaged size family would lead to a FamRS of higher than 1. Thus, the difference between these 2 scenarios corresponds to about 1 unit difference in the FamRS. Further details, the formula for calculation and an example can be found in the Supplemental Material and in Williams et al.[29]

Definition and Calculation of the PGS

The basic idea behind (poly)genic risk scores is to calculate a quantitative measure for predicting an individual's genetic risk of developing a specific condition. The contribution of genetic variants is quantified with effect estimates according to their association with diseases.[30] Then, GRS were developed including several million SNPs scattered over the whole genome without limitation by *P* value.[30] It was reasoned that it would be sensible to include more than one risk score and calculate an average of the standardized risk scores to make risk prediction more accurate.[22] This approach was applied to ischemic stroke and was used for calculating the PGS in this study.[23] Nineteen stroke-related genetic scores were included in the so-called metaGRS: any stroke, ischemic stroke, cardioembolic stroke, small and large artery stroke, 3 cardiovascular disease scores, total cholesterol, LDL (low-density lipoprotein), HDL (high-density lipoprotein), triglycerides, systolic blood pressure, diastolic blood pressure, body mass index, and T2D. The metaGRS, as denoted in Abraham et al,[23] was calculated individually for 3.2 million SNPs for all 19 GRS and then summed up as follows. The overall PGS (aka metaGRS) is the weighted ($\beta_{i,}$ i=1…19) sum of the single PGS:

$$PGS = \beta_1 \cdot PGS_{any\ stroke} + \beta_2 \cdot PGS_{ischaemic\ stroke} + \ldots + \beta_{19} \cdot PGS_{T2D}$$

More detailed information can be obtained from Abraham et al.[23]

Details on genotyping in KORA F3 can be found in the Supplemental Material. The metaGRS was calculated using PGS-Calc (https://github.com/lukfor/pgs-calc).

Statistical Analysis

For the baseline characteristics, the mean with SD and the median with 25% and 75% quantile were calculated, and percentages were denoted. For logistic and Cox regression, FamRS and PGS were treated as continuous variables, but also divided into the quantiles (top 20%, 10%, 5%, 1%, and 0.5%) to evaluate the effect of the extremes of the distribution. To ensure the validity of the estimates and sufficient power, a minimum of 5 stroke cases was set as a threshold. The analyses were conducted with the reference group lower 80%, 90%, 95%, and 99%, respectively. Different adjustments were used: (1) PGS (or FamRS), unadjusted, (2) adjustment for age and sex, (3) as (2) plus traditional risk factors (body mass index, cholesterol, diabetes, hypertension, and smoking), and (4) as 3) plus additional adjustment for FamRS (or likewise PGS, if FamRS was the initial variable of interest). Nonlinear P-splines[31] were conducted to check, if a linear relationship can be readily applied on the continuous variables PGS and FamRS. Further sensitivity analyses accounting for population stratification and relatedness are described in the Supplemental Material.

To ascertain the combination of FamRS and PGS, there were also 3 models for the variable FamRS or PGS, firstly unadjusted, secondly adjusted for age and sex, and third adjusted for traditional risk factors. This analysis could be conducted only for the stratifications in the extreme quantile groups of FamRS and PGS. The predictive ability of FamRS and PGS was evaluated with categorical-free net reclassification index (NRI) and Integrated Discrimination Improvement using function improveProb in package Hmisc. False discovery rates were calculated to account for the number of models and subgroup analyses. Table S1 includes all *P* values with their corresponding false discovery rates. All *P* values ≤0.0088 have a false discovery rates of <0.05.

RESULTS

Descriptive Statistics

In total, 71 individuals have had a stroke at study start and 186 first incident strokes were recorded (72 ischemic strokes, 39 embolic strokes, 42 ransient ischemic attack/prolonged reversible ischemic neurological deficit, 33 others, or unknown). Table 1 displays the descriptive statistics of the study cohort. The distribution of the PGS in the KORA F3 study compared with 1000 Genomes reference populations is given in Figure S2, indicating that KORAF3 resembles a European population, as expected.

Figure 1 shows that the distribution of PGS is shifted to higher values in those who suffered a stroke (mean value, −0.33) compared with those, who did not (mean value, −0.43). All individuals with a very high FamRS (essentially >1, which corresponds to the top 3%), had a stroke, but at the same time did not show a high PGS (ranging between the seventh and 94th percentile). Thus, high FamRS scores do not necessarily go hand in hand with high PGS scores. As ascertained via Spearman correlation, PGS and FamRS are not significantly correlated (*P*=0.4173).

The proportion of individuals suffering a stroke rose consistently with the higher quantiles of both PGS and FamRS. Notably, there was no individual with both very high FamRS and very high PGS, that is, highest 1% or 0.5% (Table S2).

For incident stroke events, the tendency of rising proportions in the upper quantiles could no longer be observed as clearly as before. For the PGS, the proportions even declined. Combining both measures demonstrated inconsistent trends (Table S3).

Comparing Measures for Family History

Family history is often defined as disease occurred in parents or number of affected relatives. The FamRS is comprised out of this information but additionally weighs for age of disease onset and number of relatives. Therefore, these 3 measures for family history were evaluated

Table 1. Descriptive Statistics of KORA F3 at Baseline (n=3071) as Well as the 2016 Follow-Up (n=2904)

Baseline characteristics (n=3071)	Mean±SD; median [25%Q;75%Q] or n (%)
Women (%)	1575 (51.3%)
Age, y	57.4±12.9; 57.0 [46;67]
BMI, kg/m²	27.7±4.6; 27.1 [24.4;30.3]
SBP, mm Hg	130.8±20.0; 129 [117;142.5]
DBP, mm Hg	81.9±10.9; 81.0 [74.5;88.5]
Hypertension (%)*	1535 (50.2%)
Cholesterol, mg/dL	218.3±39.9; 216 [191;243]
LDL-C, mg/dL	128±32.6; 127 [105;148]
HDL-C, mg/dL	58.8±17.2; 56 [46;69]
Triglycerides, mg/dL	165.2±126.1; 136 [88;201]
Smoker (%)	542 (18.7%)
≥1hphysical activity per week (%)	2049 (66.7%)
Any type of diabetes (%)†	268 (8.7%)
Prevalent myocardial infarction (%)‡	82 (2.7%)
Prevalent stroke (%)§	71 (2.3%)
Age at first stroke event, only prevalent cases, y	62.67±11.6; 65 [52.5;72]
Incident stroke (%) at follow-up	186 (6.9%)
Total mortality (%) at follow-up	413 (14.2%)
Death due to stroke (%) at follow-up	38 (1.3%)

For continuous variables, mean, SD, and median [25% quantile;75% quantile] are given and n(%) percentages for categorical variables. BMI indicates body mass index; DBP, diastolic blood pressure; HDL-C, high-density lipoprotein cholesterol; KORA F3, Cooperative Health Research in the Region of Augsburg; LDL-C, low-density lipoprotein cholesterol; and SBP, systolic blood pressure.

*Defined as>140/80 mm Hg or known medical treatment.

†Validated diabetes.

‡Hospitalization due to MI (self-reported).

§Hospitalization due to stroke (self-reported).

on risk of stroke (unadjusted models) to be also comparable to other studies. Having at least one parent with a stroke leads to an odds ratio (OR) for prevalent stroke of 1.57 (95% CI, 0.93–2.62). The OR of suffering a stroke increased with each affected relative: from 1.63 (95% CI, 0.96–2.70) for one to 2.53 (95% CI, 0.86–5.97) for 2 up to 5.11 (95% CI, 0.28–27.61) for those with 3 or more affected relatives. None of these was significantly associated with risk of prevalent stroke, in contrast to continuous FamRS (*P*=0.0048). Therefore, more extensive models evaluating family history were conducted with the FamRS.

Association With Prevalent Stroke

The PGS as a continuous variable was significantly associated with prevalent stroke in all adjustment models, but the OR was attenuated after inclusion of the traditional risk factors (Table 2). The stratification into the top 20%, 10%, and 5% quantile demonstrated that the OR increased in higher categories; for example, from 2.01 (95% CI, 1.19–3.29) to 2.32 (95% CI, 1.24–4.12) and 2.60 (95% CI, 1.11–5.37) for the age- and sex-adjusted model. This was confirmed by a nonlinear P-spline (Figure S3A), which shows an increasing slope for upper tails of the PGS, but which can be approximated by a linear slope. Appending the FamRS to adjustment model 3 (including traditional risk factors), hardly changed the OR and *P* values.

FamRS as a continuous variable was also significant in all adjustment models (Table 3), while none of the stratifications, expect for the top 1%, were significant in the adjusted models. A straight line best represents the increase in risk by increasing FamRS values (Figure S3B). For the FamRS, appending other risk factors and PGS hardly influenced the OR. There was no interaction between the continuous FamRS and PGS (*P*=0.27).

The combination of the variables, that is, eg, calculating the risk for individuals within the top 20% of the FamRS or the PGS compared with those not in those top groups, increased the OR and decreased the *P* values (Figure 2A and Table S4). While the OR were similar between the top 20% and top 5%, the OR rose steeply up to 5.82 (95% CI, 2.08–14.00) for the top 1% and did not change when traditional risk factors were added to the model. Sensitivity analyses accounting for potential population stratification did not change the ORs for the PGS markedly and left the results for FamRS or the combination of PGS and FamRS practically unchanged (Table S5). Removing relatives from the analyses did also not change the ORs for PGS, FamRS or the combination of both (Table S6).

Association With Incident Stroke

The PGS as a continuous variable was significant in model 1 and 2 but no longer in model 3 and 4 (Table S7). The hazard ratio (HR) even decreased with the higher quantiles. Nonlinear splines showed that the hazard for incident stroke rather reaches a plateau for PGS values above the median (Figure S4A). After adding the FamRS to the model, hardly any change in the HR for the PGS could be observed.

The FamRS was not significantly associated with time to incident stroke events, in none of the adjustment models, except for the top 1% (Table S8). However, only 5 individuals experienced a stroke event within this category. However, an increase in HR could be observed for increasing FamRS values (Table S8, Figure S4B). Adding the PGS to the fully adjusted model hardly increased the HR of the FamRS. The combination of FamRS and PGS showed an increase in HR for increasing quintiles, but that was not statistically significant (Table S9, Figure 2B).

Reclassification and Discrimination Improvement

Since there was no significant association with the PGS or FamRS with incident events in any of the models

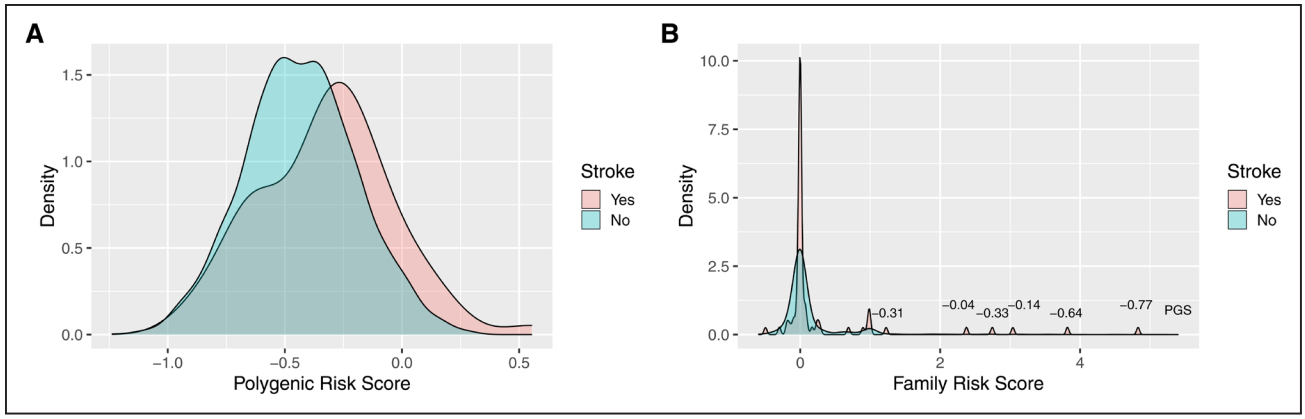

Figure 1. Density plot.
Density plot of the polygenic score (PGS; **A**) and Family Risk Score (FamRS; **B**) stratified by stroke yes (light red)/no (turquoise). The single humps for FamRS values >1 stand for single individuals, which all had a stroke and for which their respective PGS values are denoted.

adjusting for additional risk factors, predictive ability was only tested on prevalent events. The basic model included age, sex, body mass index, hypertension, total cholesterol, diabetes, and current smoking. By adding PGS into the model, the probability of events could be correctly increased in 61.3% of cases (NRI_{events}=0.226) whereas the probability was correctly decreased in 57.2% of noncases ($NRI_{non\text{-}events}$=0.145), leading to an overall NRI of 0.37 (*P*=0.003). The Integrated Discrimination Improvement was estimated to be 0.006 (*P*=0.022). Adding the FamRS alone to the basic model could only improve prediction for nonevents ($NRI_{non\text{-}events}$=0.70, NRI_{events}=−0.55, NRI=0.15, *P*=0.15; Integrated Discrimination Improvement=0.005, *P*=0.17). Taking both PGS and FamRS together in addition to the basic model leads to a better prediction for both events (NRI_{events}=0.161) and nonevents ($NRI_{non\text{-}events}$=0.30), yielding an overall NRI of 0.461 (*P*=0.0003) and Integrated Discrimination Improvement of 0.012 (*P*=0.014).

DISCUSSION

The FamRS outperformed the other family-related risk measures (affected parents and relatives), as only the FamRS was significantly associated with stroke. Each one-unit change in FamRS was leading to a 1.48-fold risk of suffering a stroke. Furthermore, a significant association between PGS and prevalent stroke was shown. A one-unit change in the PGS resulted in a 3.39-fold risk for a stroke. The adjustment of both variables for each other did not markedly influence the findings, indicating that FamRS and PGS have independent effects on the risk of stroke. As a consequence, combining both risk scores decreased the *P* values, and those being in the top 1% quantile either PGS or FamRS were found to have a 5.82-fold increased probability to have already experienced a stroke. The evaluation of different upper percentiles compared with the rest of the population and nonlinear splines showed that no specific thresholds could be derived but that the risk rather increased linearly with increasing FamRS or PGS.

The combination of both risk measures outperformed the model with solely FamRS or PGS for prevalent stroke. Adding both PGS and FamRS to a model with traditional risk factors also led to significantly improved net reclassification and discrimination. These findings might justify the additional effort in calculating the FamRS, especially since ascertaining family history, at least in a simplified form, is already an established procedure in clinical practice.[32,33] A general limitation is that the potential of family history is often underutilized, especially due to a lack of technologies to document family history and due to time limitations in the clinical routine. Therefore, a valid and intuitive tool for ascertaining and documenting family history would be useful[33–35] to make use of the FamRS.

The results on the incident stroke events were inconsistent and differed from what would be expected given previous research.[23] In the present study, however, no information on recurrent strokes was collected for those individuals who already had experienced a prevalent stroke event. Excluding prevalent cases in the incident analysis might have led to a selection bias due to selecting based on the risk factors profile. Furthermore, one might reasonably speculate that especially those with extraordinary high familial or genetic risk experience a stroke event at younger age and are consequently overrepresented in the prevalent stroke cases rather the incident ones. In line with that speculation, Marston et al[20] showed that a GRS on ischemic stroke performed better in a primary prevention cohort than in subjects with previous stroke. In our study, incident stroke events occurred on average 10 years later than prevalent events. Therefore, older age and also accumulated risk factors over years might have diminished the predictive power of FamRS and PGS. Especially as, on average, one out of 4 stroke patients experiences more than one stroke event.[36]

We demonstrated that the information included in family history and FamRS is not necessarily driven by

Table 2. Results of Logistic Regression Analysis

PGS	No. stroke cases	OR	95% CI	*P* value
Model 1: Unadjusted				
Continuous	71	6.35	2.40–16.85	0.0002
Top 20%	25	2.22	1.33–3.61	0.0016
Top 10%	15	2.47	1.34–4.32	0.0023
Top 5%	8	2.48	1.08–4.98	0.0181
Model 2: Adjusted for age and sex				
Continuous	71	5.15	1.90–14.10	0.0014
Top 20%	25	2.01	1.19–3.29	0.0070
Top 10%	15	2.32	1.24–4.12	0.0056
Top 5%	8	2.60	1.11–5.37	0.0160
Model 3: As model 2 + BMI, hypertension, cholesterol, diabetes, smoker				
Continuous	71	3.39	1.12–10.36	0.0315
Top 20%	25	1.67	0.94–2.88	0.0735
Top 10%	15	2.04	1.01–3.84	0.0355
Top 5%	8	2.04	0.75–4.63	0.1177
Model 4: As model 3 + FamRS				
Continuous	71	3.71	1.21–11.47	0.0218
Top 20%	25	1.72	0.96–2.99	0.0580
Top 10%	15	2.09	1.03–3.94	0.0307
Top 5%	8	2.17	0.80–4.95	0.0901

The models 1 to 4 give OR and 95% CI for the variable PGS and PGS categories and risk of stroke with different adjustments as denoted in the table. The reference group for the categories are the lower 80%, 90%, and 95%. BMI indicates body mass index; FamRS, Family Risk Score; OR, odds ratio; and PGS, polygenic score.

Table 3. Results of Logistic Regression Analysis

FamRS	No. stroke cases	OR	95% CI	*P* value
Model 1: Unadjusted				
Continuous	71	1.48	1.09-1.90	0.0048
Top 20%	21	1.70	0.99-2.81	0.0452
Top 10%	12	1.86	0.94-3.37	0.0549
Top 5%	9	1.66	0.76-3.22	0.1610
Model 2: Adjusted for age and sex				
Continuous	71	1.51	1.11-1.95	0.0039
Top 20%	21	1.68	0.97-2.82	0.0533
Top 10%	12	1.90	0.95-3.52	0.0511
Top 5%	9	1.63	0.74-3.21	0.1880
Model 3: As model 2 + BMI, hypertension, cholesterol, diabetes, smoker				
Continuous	71	1.48	1.07-1.93	0.0080
Top 20%	21	1.66	0.92-2.87	0.0800
Top 10%	12	1.80	0.86-3.45	0.0946
Top 5%	9	1.47	0.63-3.04	0.3340
Model 4: As model 3 + PGS				
Continuous	71	1.52	1.10-2.00	0.0050
Top 20%	21	1.66	0.92-2.88	0.0797
Top 10%	12	1.82	0.87-3.50	0.0890
Top 5%	9	1.51	0.64-3.13	0.3010

The models 1 to 4 give OR and 95% CI for the variable FamRS and FamRS categories and risk of stroke with different adjustments as denoted in the table. The reference group for the categories are the lower 80%, 90%, and 95%. BMI indicates body mass index; FamRS, Family Risk Score; OR, odds ratio; and PGS, polygenic score.

genetics. The results indicated that the FamRS captures information not incorporated in the PGS and vice versa. Apart from shared genetic factors, the FamRS contains information on social determinants of health, behavior, and shared environment.[34,37] Interestingly, the OR of the FamRS was not influenced by traditional risk factors. The OR of the PGS, however, decreased after inclusion of other risk factors. This is only surprising at the first glance. The PGS also incorporates GRS for several of the risk factors, which were adjusted for (body mass index, cholesterol, blood pressure, diabetes). That means, the PGS is one cause of those risk factors, which are causes for stroke themselves. Adjusting for those risk factors removes the indirect causal paths from the PGS to stroke. Thus, the risk factor adjusted analysis likely underestimates the PGS effect. It will be relevant for future analyses, to ascertain whether the effects of the FamRS would be diminished by including recognized behavioral and environmental risk factors such as air pollution, diet, and physical activity, which are not available in sufficiently high granularity and quality in the present study.[38]

Limitations and Strengths

To the best of our knowledge, this is the first analysis combining a genetic with a familial risk score, which is the major strength of this study. Furthermore, different family-related risk factors (eg, affected parents, affected relatives, and the FamRS) were systematically compared and the FamRS outperformed other measures of family history. As KORA F3 is a population-based study, it might more accurately represent the true familial and genetic susceptibility to stroke within a population than a case-control study.[39] The population-based nature of the study is also a limitation at the same time, since there are only few stroke cases, especially in the high risk groups. The analyses in the top 5% or even top 1% quantiles are to be taken with caution. It has also to be noted that all models, which are based on different upper thresholds of FamRS, PGS, and their combinations have not been predefined and are of explorative nature. Results are pretty stable, though, and are confirmed by nonlinear splines in the extreme tails of the distribution. False discovery rates reassure the findings. There was no information on stroke subtypes for the prevalent events and also the FamRS did not take subtypes into account. This could lead to the mixing of different types of genetic mechanisms and family structures in the analyses.

In population-based studies with low stroke prevalence, a relatively large proportion of self-reported strokes may be false positives. Therefore, results should be confirmed by replications.

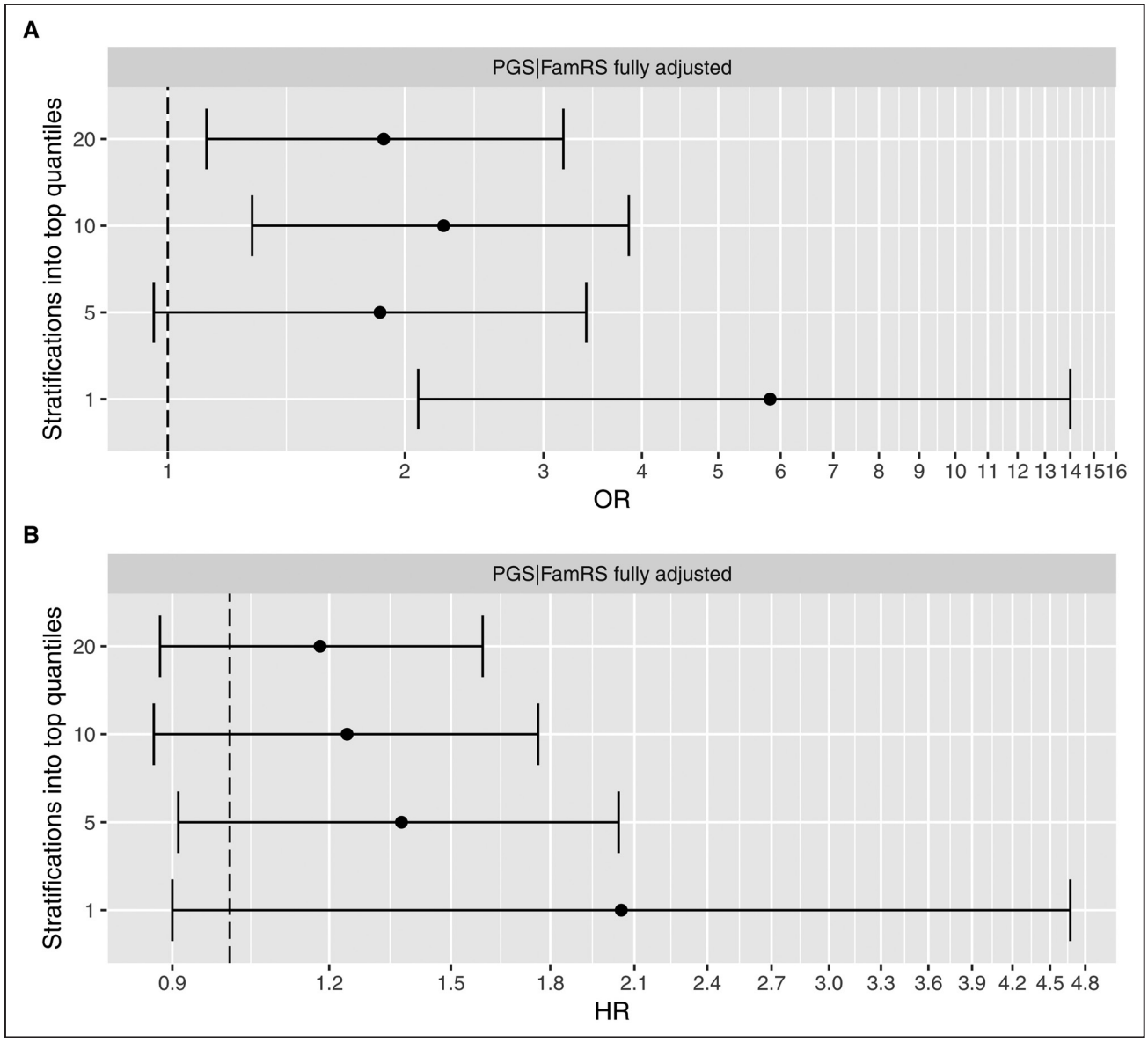

Figure 2. Results for stratifications into top quantiles of either polygenic score (PGS) or Family Risk Score (FamRS). **A**, Odds ratio (OR; 95% CI) of the logistic regression analysis for being in the top 1%, 5%, 10%, or 20% quantiles of either PGS or FamRS for the fully adjusted model. **B**, Hazard ratio (HR; 95% CI) of the Cox model for being in the top 1%, 5%, 10% or 20% quantiles of either PGS or FamRS for the fully adjusted model.

Another limiting factor is that all participants are from European-ancestry and live in close geographic proximity in Germany. Previous research demonstrated that scores derived from genome-wide association studies with only European participants are biased due to genetic drifts found in different populations.[40]

Conclusions

To conclude, we could demonstrate that the FamRS and PGS were independently and significantly associated with stroke. The FamRS included more than the genetic information as captured by the PGS, while the PGS included information different from those in the FamRS. Therefore, even in light of SNP-microarrays, which are inexpensive and make the use of polygenic scores feasible even in clinical settings, collecting family history and calculating a sophisticated and weighted family score, as the proposed FamRS, could increase accuracy. More efforts should be put into developing an easy-to-handle and reliable tools for ascertaining family history.

ARTICLE INFORMATION

Received July 6, 2021; final revision received February 1, 2022; accepted March 8, 2022.

Affiliations

Institute of Genetic Epidemiology, Department of Genetics, Medical University of Innsbruck, Innsbruck, Austria (M.H., L.F., S.H., F.K., C.L.). Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental

Health, Neuherberg, Germany (A.P., C.G., H.G.). German Center for Diabetes Research (DZD), Neuherberg, Germany (A.P., C.G., H.G.). German Research Center for Cardiovascular Research (DZHK), Partner Site Munich Heart Alliance, Munich, Germany (A.P.). Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany (C.G., H.G.).

Sources of Funding

The KORA study was initiated and financed by the Helmholtz Zentrum München – German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. This project is also funded by the Bavarian State Ministry of Health and Care through the research project DigiMed Bayern (www.digimed-bayern.de).

Disclosures

None.

Supplemental Material

Supplemental Materials and Methods
Figures S1–S4
Tables S1–S9

REFERENCES

1. Li L, Scott CA, Rothwell PM; Oxford Vascular Study. Trends in stroke incidence in high-income countries in the 21st century: population-based study and systematic review. *Stroke*. 2020;51:1372–1380. doi: 10.1161/STROKEAHA.119.028484
2. GBD 2016 Stroke Collaborators. Global, regional, and national burden of stroke, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet. Neurol.* 2019;18:439–458.
3. Wafa HA, Wolfe CDA, Emmett E, Roth GA, Johnson CO, Wang Y. Burden of stroke in Europe: thirty-year projections of incidence, prevalence, deaths, and disability-adjusted life years. *Stroke*. 2020;51:2418–2427. doi: 10.1161/STROKEAHA.120.029606
4. Feigin VL, Nguyen G, Cercy K, Johnson CO, Alam T, Parmar PG, Abajobir AA, Abate KH, Abd-Allah F, Abejie AN, et al. Global, regional, and country-specific lifetime risks of stroke, 1990 and 2016. *N Engl J Med*. 2018;379:2429–2437.
5. Sarikaya H, Ferro J, Arnold M. Stroke prevention–medical and lifestyle measures. *Eur Neurol*. 2015;73:150–157. doi: 10.1159/000367652
6. Chiuve SE, Rexrode KM, Spiegelman D, Logroscino G, Manson JE, Rimm EB. Primary prevention of stroke by healthy lifestyle. *Circulation*. 2008;118:947–954. doi: 10.1161/CIRCULATIONAHA.108.781062
7. Larsson SC, Åkesson A, Wolk A. Primary prevention of stroke by a healthy lifestyle in a high-risk group. *Neurology*. 2015;84:2224–2228. doi: 10.1212/WNL.0000000000001637
8. Tikk K, Sookthai D, Monni S, Gross ML, Lichy C, Kloss M, Kaaks R. Primary preventive potential for stroke by avoidance of major lifestyle risk factors: the European Prospective Investigation into Cancer and Nutrition-Heidelberg cohort. *Stroke*. 2014;45:2041–2046. doi: 10.1161/STROKEAHA.114.005025
9. Kiely DK, Wolf PA, Cupples LA, Beiser AS, Myers RH. Familial aggregation of stroke. The Framingham Study. *Stroke*. 1993;24:1366–1371. doi: 10.1161/01.str.24.9.1366
10. Seshadri S, Beiser A, Pikula A, Himali JJ, Kelly-Hayes M, Debette S, DeStefano AL, Romero JR, Kase CS, Wolf PA. Parental occurrence of stroke and risk of stroke in their children: the Framingham study. *Circulation*. 2010;121:1304–1312. doi: 10.1161/CIRCULATIONAHA.109.854240
11. Kasiman K, Lundholm C, Sandin S, Malki N, Sparén P, Ingelsson E. Familial effects on ischemic stroke: the role of sibling kinship, sex, and age of onset. *Circ Cardiovasc Genet*. 2012;5:226–233. doi: 10.1161/CIRCGENETICS.111.962241
12. Sundquist K, Li X, Hemminki K. Familial risk of ischemic and hemorrhagic stroke: a large-scale study of the Swedish population. *Stroke*. 2006;37:1668–1673. doi: 10.1161/01.STR.0000227409.59195.d1
13. Yu S, Su Z, Miao J, Yu Y, Zhang S, Wu J, Zheng H, Zhang X, Zhong S, Li H, et al. Different types of family history of stroke and stroke risk: results based on 655,552 individuals. *J Stroke Cerebrovasc Dis*. 2019;28:587–594. doi: 10.1016/j.jstrokecerebrovasdis.2018.10.038
14. Tian T, Jin G, Yu C, Lv J, Guo Y, Bian Z, Yang L, Chen Y, Shen H, Chen Z, et al; China Kadoorie Biobank Collaborative Group. Family history and stroke risk in china: evidence from a large cohort study. *J Stroke*. 2017;19:188–195. doi: 10.5853/jos.2016.01270
15. Dichgans M, Pulit SL, Rosand J. Stroke genetics: discovery, biology, and clinical applications. *Lancet Neurol*. 2019;18:587–599. doi: 10.1016/S1474-4422(19)30043-2
16. Malik R, Bevan S, Nalls MA, Holliday EG, Devan WJ, Cheng YC, Ibrahim-Verbaas CA, Verhaaren BF, Bis JC, Joon AY, et al; Wellcome Trust Case Control Consortium 2. Multilocus genetic risk score associates with ischemic stroke in case-control and prospective cohort studies. *Stroke*. 2014;45:394–402. doi: 10.1161/STROKEAHA.113.002938
17. Tada H, Shiffman D, Smith JG, Sjögren M, Lubitz SA, Ellinor PT, Louie JZ, Catanese JJ, Engström G, Devlin JJ, et al. Twelve-single nucleotide polymorphism genetic risk score identifies individuals at increased risk for future atrial fibrillation and stroke. *Stroke*. 2014;45:2856–2862. doi: 10.1161/STROKEAHA.114.006072
18. Hachiya T, Kamatani Y, Takahashi A, Hata J, Furukawa R, Shiwa Y, Yamaji T, Hara M, Tanno K, Ohmomo H, et al. Genetic predisposition to ischemic stroke: a polygenic risk score. *Stroke*. 2017;48:253–258. doi: 10.1161/STROKEAHA.116.014506
19. Hachiya T, Hata J, Hirakawa Y, Yoshida D, Furuta Y, Kitazono T, Shimizu A, Ninomiya T. Genome-wide polygenic score and the risk of ischemic stroke in a prospective cohort: the hisayama study. *Stroke*. 2020;51:759–765. doi: 10.1161/STROKEAHA.119.027520
20. Marston NA, Patel PN, Kamanu FK, Nordio F, Melloni GM, Roselli C, Gurmu Y, Weng LC, Bonaca MP, Giugliano RP, et al. Clinical application of a novel genetic risk score for ischemic stroke in patients with cardiometabolic disease. *Circulation*. 2021;143:470–478. doi: 10.1161/CIRCULATIONAHA.120.051927
21. Rutten-Jacobs LC, Larsson SC, Malik R, Rannikmäe K, Sudlow CL, Dichgans M, Markus HS, Traylor M; MEGASTROKE consortium; International Stroke Genetics Consortium. Genetic risk, incident stroke, and the benefits of adhering to a healthy lifestyle: cohort study of 306 473 UK Biobank participants. *BMJ*. 2018;363:k4168. doi: 10.1136/bmj.k4168
22. Inouye M, Abraham G, Nelson CP, Wood AM, Sweeting MJ, Dudbridge F, Lai FY, Kaptoge S, Brozynska M, Wang T, et al; UK Biobank CardioMetabolic Consortium CHD Working Group. Genomic risk prediction of coronary artery disease in 480,000 adults: implications for primary prevention. *J Am Coll Cardiol*. 2018;72:1883–1893. doi: 10.1016/j.jacc.2018.07.079
23. Abraham G, Malik R, Yonova-Doing E, Salim A, Wang T, Danesh J, Butterworth AS, Howson JMM, Inouye M, Dichgans M. Genomic risk score offers predictive performance comparable to clinical risk factors for ischaemic stroke. *Nat Commun*. 2019;10:5819. doi: 10.1038/s41467-019-13848-1
24. Do CB, Hinds DA, Francke U, Eriksson N. Comparison of family history and SNPs for predicting risk of complex disease. *PLoS Genet*. 2012;8:e1002973. doi: 10.1371/journal.pgen.1002973
25. Lambert SA, Abraham G, Inouye M. Towards clinical utility of polygenic risk scores. *Hum Mol Genet*. 2019;28(R2):R133–R142. doi: 10.1093/hmg/ddz187
26. Holle R, Happich M, Löwel H, Wichmann HE; MONICA/KORA Study Group. KORA–a research platform for population based health research. *Gesundheitswesen*. 2005;67(Suppl 1):S19–S25. doi: 10.1055/s-2005-858235
27. Löwel H, Döring A, Schneider A, Heier M, Thorand B, Meisinger C. The MONICA Augsburg surveys--basis for prospective cohort studies. *Gesundheitswesen*. 2005;67 (Suppl 1):S13–S18.
28. Lamina C, Linsenmeyer J, Weissensteiner H, Kollerits B, Meisinger C, Rantner B, Stöckl D, Stadler M, Klein-Weigel P, Peters A, et al. Correlation between a positive family risk score and peripheral artery disease in one case-control and two population-based studies. *Atherosclerosis*. 2014;237:243–250. doi: 10.1016/j.atherosclerosis.2014.08.032
29. Williams RR, Hunt SC, Heiss G, Province MA, Bensen JT, Higgins M, Chamberlain RM, Ware J, Hopkins PN. Usefulness of cardiovascular family history data for population-based preventive medicine and medical research (the Health Family Tree Study and the NHLBI Family Heart Study). *Am J Cardiol*. 2001;87:129–135. doi: 10.1016/s0002-9149(00)01303-5
30. Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, Natarajan P, Lander ES, Lubitz SA, Ellinor PT, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet*. 2018;50:1219–1224. doi: 10.1038/s41588-018-0183-z
31. Eilers PHC, Marx BD. Flexible smoothing with B-splines and penalties. *Stat Sci*. 1996;11:89–121.
32. Ginsburg GS, Wu RR, Orlando LA. Family health history: underused for actionable risk assessment. *Lancet*. 2019;394:596–603. doi: 10.1016/S0140-6736(19)31275-9

33. Orlando LA, Hauser ER, Christianson C, Powell KP, Buchanan AH, Chesnut B, Agbaje AB, Henrich VC, Ginsburg G. Protocol for implementation of family health history collection and decision support into primary care using a computerized family health history system. *BMC Health Serv Res.* 2011;11:264. doi: 10.1186/1472-6963-11-264
34. de Hoog CLMM, Portegijs PJM, Stoffers HEJH. Family history tools for primary care are not ready yet to be implemented. A systematic review. *Eur J Gen Pract.* 2014;20:125–133.
35. Leventer-Roberts M, Gofer I, Barak Corren Y, Reis BY, Balicer R. Constructing data-derived family histories using electronic health records from a single healthcare delivery system. *Eur J Public Health.* 2020;30:212–218. doi: 10.1093/eurpub/ckz152
36. Esenwa C, Gutierrez J. Secondary stroke prevention: challenges and solutions. *Vasc Health Risk Manag.* 2015;11:437–450. doi: 10.2147/VHRM.S63791
37. Kulshreshtha A, Vaccarino V, Goyal A, McClellan W, Nahab F, Howard VJ, Judd SE. Family history of stroke and cardiovascular health in a national cohort. *J Stroke Cerebrovasc Dis.* 2015;24:447–454. doi: 10.1016/j.jstrokecerebrovasdis.2014.09.017
38. Boehme AK, Esenwa C, Elkind MS. Stroke risk factors, genetics, and prevention. *Circ Res.* 2017;120:472–495. doi: 10.1161/CIRCRESAHA.116.308398
39. Szklo M. Population-based cohort studies. *Epidemiol Rev.* 1998;20:81–90. doi: 10.1093/oxfordjournals.epirev.a017974
40. Martin AR, Gignoux CR, Walters RK, Wojcik GL, Neale BM, Gravel S, Daly MJ, Bustamante CD, Kenny EE. Human demographic history impacts genetic risk prediction across diverse populations. *Am J Hum Genet.* 2017;100:635–649. doi: 10.1016/j.ajhg.2017.03.004
41. Zahn K, Linseisen J, Heier M, Peters A, Thorand B, Nairz F, Meisinger C. Body fat distribution and risk of incident ischemic stroke in men and women aged 50 to 74 years from the general population. The KORA Augsburg cohort study. *PLoS One.* 2018;13:e0191630. doi: 10.1371/journal.pone.0191630
42. McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, Teumer A, Kang HM, Fuchsberger C, Danecek P, Sharp K, et al; Haplotype Reference Consortium. A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet.* 2016;48:1279–1283. doi: 10.1038/ng.3643
43. Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, Vrieze SI, Chew EY, Levy S, McGue M, et al. Next-generation genotype imputation service and methods. *Nat Genet.* 2016;48:1284–1287. doi: 10.1038/ng.3656
44. Lambert SA, Gil L, Jupp S, Ritchie SC, Xu Y, Buniello A, McMahon A, Abraham G, Chapman M, Parkinson H, et al. The Polygenic Score Catalog as an open database for reproducibility and systematic evaluation. *Nat Genet.* 2021;53:420–425. doi: 10.1038/s41588-021-00783-5