JCN Open Access ORIGINAL ARTICLE

pISSN 1738-6586 / eISSN 2005-5013 / J Clin Neurol 2022;18(5):553-561 / https://doi.org/10.3988/jcn.2022.18.5.553

Multidimensional Early Prediction Score for Drug-Resistant Epilepsy

Kyung Wook Kang[a], Yong Won Cho[b]
Sang Kun Lee[c], Ki-Young Jung[c]
Ji Hyun Kim[d], Dong Wook Kim[e]
Sang-Ahm Lee[f], Seung Bong Hong[g,h]
In-Seop Na[i], So-Hyun Lee[j]
Won-Ki Baek[k], Seok-Yong Choi[j]
Myeong-Kyu Kim[a]

[a]Department of Neurology, Chonnam National University Hospital, Chonnam National University Medical School, Gwangju, Korea
[b]Department of Neurology, Dongsan Medical Center, Keimyung University School of Medicine, Daegu, Korea
[c]Department of Neurology, Comprehensive Epilepsy Center, Laboratory for Neurotherapeutics, Biomedical Research Institute, Seoul National University Hospital, Seoul, Korea Program in Neuroscience, Seoul National University College of Medicine, Seoul, Korea
[d]Department of Neurology, Korea University Guro Hospital, Korea University College of Medicine, Seoul, Korea
[e]Department of Neurology, Konkuk University School of Medicine, Seoul, Korea
[f]Department of Neurology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea
[g]Department of Neurology, Samsung Medical Center, Samsung Advanced Institute for Health Sciences & Technology (SAIHST), Sungkyunkwan University School of Medicine, Samsung Biomedical Research Institute (SBRI), Seoul, Korea
[h]National Epilepsy Care Center, Seoul, Korea
[i]National Program of Excellence in Software Centre, Chosun University, Gwangju, Korea
[j]Department of Biomedical Science, Chonnam National University Medical School, Hwasun, Korea
[k]Department of Microbiology, Keimyung University School of Medicine, Daegu, Korea

Received November 30, 2021
Revised March 16, 2022
Accepted March 16, 2022

Correspondence
Myeong-Kyu Kim, MD, PhD
Department of Neurology, Chonnam National University Medical School,
42 Jebong-ro, Dong-gu,
Gwangju 61469, Korea
Tel +82-62-220-6161
Fax +82-62-228-3461
E-mail mkkim@jnu.ac.kr

Seok-Yong Choi, MD, PhD
Department of Biomedical Science, Chonnam National University Medical School,
264 Seoyang-ro, Hwasun-eup,
Hwasun 58128, Korea
Tel +82-61-379-2500
Fax +82-62-228-3461
E-mail zebrafish@jnu.ac.kr

Background and Purpose Achieving favorable postoperative outcomes in patients with drug-resistant epilepsy (DRE) requires early referrals for preoperative examinations. The purpose of this study was to investigate the possibility of a user-friendly early DRE prediction model that is easy for nonexperts to utilize.

Methods A two-step genotype analysis was performed, by applying 1) whole-exome sequencing (WES) to the initial test set (n=243) and 2) target sequencing to the validation set (n=311). Based on a multicenter case–control study design using the WES data set, 11 genetic and 2 clinical predictors were selected to develop the DRE risk prediction model. The early prediction scores for DRE (EPS-DRE) was calculated for each group of the selected genetic predictors (EPS-DRE$_{gen}$), clinical predictors (EPS-DRE$_{cln}$), and two types of predictor mix (EPS-DRE$_{mix}$) in both the initial test set and the validation set.

Results The multidimensional EPS-DRE$_{mix}$ of the predictor mix group provided a better match to the outcome data than did the unidimensional EPS-DRE$_{gen}$ or EPS-DRE$_{cln}$. Unlike previous studies, the EPS-DRE$_{mix}$ model was developed using only 11 genetic and 2 clinical predictors, but it exhibited good discrimination ability in distinguishing DRE from drug-responsive epilepsy. These results were verified using an unrelated validation set.

Conclusions Our results suggest that EPS-DRE$_{mix}$ has good performance in early DRE prediction and is a user-friendly tool that is easy to apply in real clinical trials, especially by nonexperts who do not have detailed knowledge or equipment for assessing DRE. Further studies are needed to improve the performance of the EPS-DRE$_{mix}$ model.

Keywords epilepsy; drug resistant epilepsy; genome-wide association study; genetic predictor.

INTRODUCTION

Epilepsy is a common neurological disorder that affects approximately 8 in 1,000 individuals worldwide,[1] of which more than one-third experience drug-resistant seizures even in industrialized countries where most antiseizure medications (ASMs) are readily available.[2] People with drug resistant epilepsy (DRE) are well known to have minimal chances of seizure freedom based on additional medication trials,[3] and they suffer from substantial disabilities including reduced quality of life (QOL), serious psychosocial consequences, cognitive problems, and increased morbidity and mortality in addition to seizure itself.[4-6] Furthermore, studies of epilepsy subpopulations have revealed a consistent pattern of markedly higher healthcare costs for those with DRE and for those with comorbidities.[7] Thus, the early identification of DRE during ASM treatment has a high priority over early introduction of late-alternative treatment modalities such as epilepsy surgery. This approach is important for preventing the associated disability outcomes and improving prognoses.

It is well known that epilepsy surgery is an evidence-based treatment option for people with DRE that is superior to drug therapy.[8-13] Also, the only known modifiable factor associated with a favorable postsurgical seizure outcome is the duration of epilepsy prior to epilepsy surgery.[14] Systematic reviews and a meta-analysis of the effects of epilepsy surgery time revealed that a short epilepsy duration has a favorable effects on QOL, cognitive and psychosocial function, cost and mortality as a determinant of postoperative seizure-free.[14-19] This indicates the importance of early referral of patients with DRE, including avoiding the delays associated with presurgical investigations to determine the suitability for epilepsy surgery.[14] However, several studies have shown that the duration of epilepsy before temporal lobe epilepsy is about 20 years[8,9,20-22] and that the delay to surgery has not decreased for more than 10 years, despite the increasing evidence for the efficacy of epilepsy surgery.[23,24]

The reason for delaying surgery includes neurologists finding it difficult to identify DRE or being reluctant to consider epilepsy surgery early in the course of the disease.[21,22,25-28] This difficulty or reluctance is at least partly attributable to general neurologists not being able to accurately identify potential surgical candidates.[22] Despite remarkable advancements in the understanding of epilepsy pathogenesis, the mechanism underlying ASM resistance is not yet sufficiently clear to allow the identification of DRE biomarkers.[29] Also, clinical predictors of DRE identified in outcome studies in epilepsy do not fully predict ASM resistance by single or combination.[30] On the other hand, Kwan and Brodie[3] proposed that DRE may be present from the start of epilepsy rather than evolving over time, since certain clinical characteristics of this type of epilepsy are obvious at the beginning of the disease, raising the possibility of early DRE diagnosis.

There is accumulating evidence that polygenic inheritance with many common genetic variants exerting modest effects plays a greater role than rare monogenic mutations in most common diseases.[31-34] Genome-wide polygenic risk scores (PRSs) have recently been developed for five common diseases such as coronary artery disease, atrial fibrillation, and breast cancer, and have been demonstrated to be reliable in predicting the risk of disease in individuals with European ancestry.[34,35] Also, such a PRS is the only approach that provides an estimate of genetic liability to a trait at the individual level, and which can be used to stratify individuals according to their risk of a given disease to improve screening and prevention strategies.[34,36] Nevertheless, it should be noted that PRSs cannot be used for all races because the characteristics of the genetic architecture underlying PRSs vary with ancestry,[34] which makes it necessary to develop PRSs based on race-specific genome-wide association study (GWAS) data. However, there are no unique GWAS data to date for estimating the DRE risk for Koreans, let alone PRSs for DRE prediction.

The purpose of this study was to investigate the possibility of a user-friendly early DRE prediction model using a relatively small Korean GWAS data set and unambiguous clinical information that is easy for nonexperts to utilize at the time of an epilepsy diagnosis. Such a model would made it easy for nonexperts to decide when to introduce late alternative treatments such as epilepsy surgery as soon as possible in the process of managing newly diagnosed epilepsy patients at outpatient clinics in Korea.

METHODS

Study design and participants

This multicenter case-control study recruited consecutive participants from 12 domestic epilepsy referral centers and performed a 2-step genotype analysis. Whole-exome sequencing (WES) was applied to 243 participants (test set) who participated in our previous study (recruited from Jan 2016 to Jun 2017),[37] and target sequencing was applied to 311 participants (validation set) who were newly involved in verifying the results of the initial study (recruited from Jul 2020 to Dec 2021). The DRE prediction scores were calculated based on genotype data and clinical information of all participants.

The inclusion and exclusion criteria in this study were identical to those in our previous study.[37] In brief, participants were eligible if they were older than 20 years and had DRE or drug-responsive epilepsy (RSE) according to the following definitions and criteria: To enhance the phenotype contrast between the RSE and DRE groups, we defined drug resistance more stringently than the conventional definition[38] as the occurrence of at least 12 unprovoked seizures during 1 year prior to recruitment, with trials of 2 or more appropriate ASMs at the maximal tolerated doses, which were established on the basis of the occurrence of clinical side effects at supramaximal doses. RSE was defined as complete freedom from seizures for at least 1 year up to the date of the last follow-up visit in patients treated with a single ASM. However, patients who had a definite history of epilepsy in first- or second-degree relatives, frequently exhibited poor compliance with ASM therapy, had experienced nonmotor seizures only without consciousness impairment, or had progressive developmental epileptic encephalopathies were excluded. An extensive historical assessment was performed in all participants using a standardized form to obtain detailed information on the epidemiology, seizure characteristics, epilepsy syndrome, electroencephalography (EEG) and magnetic resonance imaging findings, and the family history.

This study was approved by the Institutional Review Boards

at Chonnam National University Hospital (approved numbers CNUH-2016-028 and CNUH-2020-208). All research protocols were performed in accordance with relevant guidelines and regulations, and written informed consent was obtained from all participants.

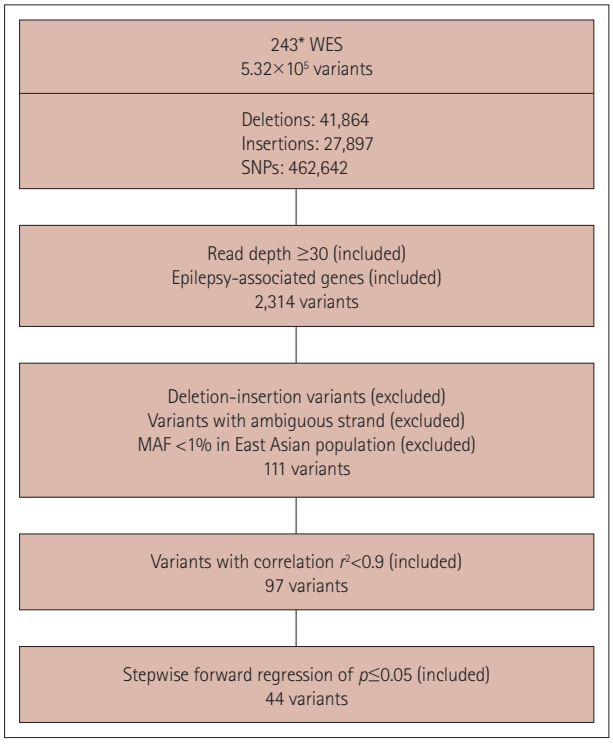

Fig. 1. Workflow of the candidate genetic variants filtering process. The customized stringent filtering process was used to identify candidate genetic variants with higher predictability of drug resistant epilepsy with a high confidence. *1 case was excluded from the final analysis due to data ambiguity. MAF, minor-allele frequency; SNPs, single nucleotide polymorphisms; WES, whole-exome sequencing.

Selection of candidate genetic variants and clinical predictors

WES data for the test set were produced following the manufacturer's protocol as described for our previous study.[37] Quality control of WES data was performed mostly according to the following standard inclusion guideline recommended by Choi et al.[36]: individual sample or variant fulfilling a genotyping call rate >0.99, sample missingness <0.02, Hardy-Weinberg equilibrium $p>1\times10^{-6}$, and heterozygosity within three standard deviations of the mean.

The workflow for WES data analysis to identify candidate genetic variants with a higher prediction potential is shown in Fig. 1. First, variants with a read depth of ≥30× and variants of known epilepsy-associated genes (n=215) (Table 1) that cause pure or relatively pure epilepsies or syndromes with epilepsy as the core symptom[39] or variants of candidate genes that are associated with the hypothesized mechanisms of ASM resistance[40,41] were included. Second, deletion-insertion variants, variants with ambiguous strand (A/T or C/G), and variants with a minor allele frequency of <1% in East Asian population (gnomAD; http://gnomad.broadinstitute.org/) were excluded. Third, uncorrelated variants (correlation coefficient r^2<0.9) with lowest p value for association with overall DRE were included. Fourth, to achieve the highest prediction accuracy, only variants with a stepwise forward regression for which $p\leq0.05$ were retained. For the finally selected target variants to be included in calculating the risk score, real-time PCR using high resolution melting (HRM) analysis with appropriate primer pairs (Supplementary Table 1 in the online-

Table 1. Epilepsy associated genes

Epilepsy genes (n=105)
AARS, ADRA2B, ADSL, ALDH7A1 ALG13, ARV1, ATP6AP2, CACNA1A, CACNA1H, CACNB4, CASR, CDKL5, CERS1, CHD2, CHRNA2, CHRNA4, CHRNB2, CLCN2, CLN3, CLN5, CLN6, CLN8, CNTN2, CPA6, CSTB, CTSD, DEPDC5, DNM1, DOCK7, EEF1A2, EFHC1, EPM2A, FGF12, FOXG1, FRRS1L, GABRA1, GABRB1, GABRB3, GABRD, GABRG2, GAL, GAMT, GATM, GNAO1, GOSR2, GPR98, GRIN2A, GRIN2B, GRIN2D, GUF1, HCN1, ITPA, KCNA2, KCNB1, KCNC1, KCNMA1, KCNQ2, KCNQ3, KCNT1, KCTD7, LGI1, LMNB2, MFSD8, NECAP1, NHLRC1, NPRL2, NPRL3, NRXN1, PCDH19, PLCB1, PNPO, POLG, PPT1, PRDM8, PRICKLE1, PRIMA1, PRRT2, SCARB2, SCN1A, SCN1B, SCN2A, SCN8A, SCN9A, SIK1, SLC12A5, SLC13A5, SLC1A2, SLC25A12, SLC25A22, SLC2A1, SLC6A1, SLC9A6, SPTAN1, ST3GAL3, ST3GAL5, STX1B, STXBP1, SZT2, TBC1D24, TCF4, TPP1, UBA5, UBE3A, WWOX, ZEB2
Neurodevelopment-associated epilepsy genes (n=73)
ANKLE2, AMPD2, ARFGEF2, ARX, ASPM, ATN1, CASK, CCDC88C, CDK5, CENPE, CENPJ, CLP1, CNTNAP2, COL4A2, DCX, DIAPH1, EMX2, EPSECS, ERMARD, EXOSC3, FIG4, FLNA, GPR56, HERC1, IER3IP1, KATNB1, KIF11, KIF2A, KIF5C, LAMB1, LAMC3, MED17, MFSD2A, MPDZ, NDE1, NSDHL, OCLN, OPHN1, PAFAH1B1, PCLO, PIK3R2, PLEKHG2, PNKP, PPP1R15B, PTCH1, QARS, RELN, RTTN, SASS6, SLC12A6, SLC20A2, SNIP1, SPATA5, SRPX2, STAMBP, STRADA, SYN1, TRMT10A, TSC1, TSC2, TSEN15, TSEN2, TSEN54, TUBA1A, TUBA8, TUBB2A, TUBB2B, TUBB3, TUBG1, VPS53, WDR62, WDR73, XPR1
Other genes* associated with the hypothesized mechanisms of ASM resistance (n=37)
ABCB1, ABCC2, ACTB, ACTG1, AKT3, AVO3, CHD7, COL18A1, CYP1A1, CYP2C19, CYP2C9, DCHS1, DYNC1H1, EOMES, EZH2, FAT4, FH, FMR1, KIAA1279, KIAA1303, KIAA1999, LIS1, LRP2, MTOR, PAX6, PIK3CA, PTENRAB18, RAB3GAP1, RAB3GAP2, RICTOR, RPTOR, SNAP29, TBR2, TUBB, UGT1A1, UGT2B7, VLDR

*The genes common to the previous two categories were omitted. ASM, antiseizure medication.

only Data Supplement) was performed according to the manufacturer's protocol[42] to determine the genotype of the new samples for validation purposes (validation set). Any ambiguous results from HRM analysis were re-examined using Sanger sequencing.

Clinical factors associated with DRE were selected through a literature review. The following clinical factors were considered as candidate clinical predictors: age at seizure onset, epilepsy duration, initial responses to treatment, pretreatment seizures frequency, symptomatic epilepsy, seizure type, abnormal neuroimaging findings, abnormal EEG findings, febrile seizure, and status epilepticus.[2,43,44] Seizure type was classified into focal and generalized seizures. An abnormal neuroimaging finding was defined as the presence of potential epileptogenic lesions in the brain, such as hippocampal sclerosis, tumor, vascular malformation, focal cortical dysplasia, stroke, or brain trauma. An asymptomatic periventricular white-matter hyperintensity was not considered as a potential epileptogenic lesion in this study. Symptomatic epilepsy was defined based on the presence of abnormal neuroimaging findings as listed above. An abnormal EEG finding was defined as one indicating either epileptiform discharges or abnormal slow waves (abnormal EEG1), or epileptiform discharges only (abnormal EEG2). Status epilepticus was defined in accordance with the classification of the International League Against Epilepsy task force.[45] Reported clinical predictors of DRE[2,43,44] that could be subject to observer-based bias and thus affect the study outcomes were excluded from the analysis for calculating the early prediction score for DRE (EPS-DRE).

Calculating EPS-DRE

The EPS-DRE was developed in accordance with the statistical methods used in previous studies[34,35] to calculate genome-wide polygenic scores for common diseases, with minor modifications:

$$\text{EPS-DRE} = \beta_1\chi_1 + \beta_2\chi_2 + \ldots + \beta_\kappa\chi_\kappa + \ldots \beta_n\chi_n,$$

where $\beta\kappa$ is the per-allelic or per-clinical-predictor logarithm of the odds ratio (OR) for DRE associated with genetic variant or clinical predictor κ, which is the effect size (weight) of the risk alleles or clinical factors as estimated by a GWAS on the phenotype; χ_κ is the dosage for genetic variant κ or clinical predictor κ (number of risk allele or clinical factors); and n is the total number of genetic variants and clinical predictors included in the EPS-DRE. Since the PRS provides a quantitative metric of the inherited risk of an individual based on the cumulative impact of many common polymorphisms, a weight is generally assigned to each genetic variant according to the strength of its association with disease risk.[34] In calculating PRSs for a binary (e.g., case/control) phenotype, the effect sizes used as weights are typically reported as log(OR) values.[36] Therefore, EPS-DREs were generated based on how many risk alleles they have for each genetic variant (e.g., 0, 1, 2 copies)[34,35] or whether or not they have each clinical predictor (e.g., 0 or 1) in the present study. More simply, the EPS-DRE was calculated by computing the sum of the value: [the dosage of each risk allele or each clinical predictor × respective log(OR) value].[36] Finally, the standardized EPS-DRE [(each calculated EPS-DRE value minus the mean of EPS-DREs)/(standard deviation of EPS-DREs)] was used to measure the predictive ability and validity of the EPS-DRE model. For validation testing, the dosage and effect size of each risk allele or each clinical predictor generated from the reference test set were applied as it was to the validation set.

Statistical analysis

Student's *t*-test and chi-square test were used to examine the difference of demographic variables between the DRE and RSE groups. In addition, EPS-DREs were compared between these two groups using *t*-tests. Univariate linear regression analysis was performed to identify the candidate genetic and clinical predictors for calculating EPS-DRE using the test set. Multivariate linear regression analysis was performed to adjust for covariates. If multiple candidate predictors were significantly correlated, the predictor with the lowest *p* value was selected. Furthermore, the missing values for each variable were excluded from the analysis. The correlation between the DRE and EPS-DREs was also determined using regression analysis. The significance level was evaluated using the Wald test, and the explanatory ability was evaluated through Nagelkerke's R^2 value. Receiver operating characteristic (ROC) analysis was used to determine the effect size of the continuous variable (i.e., seizure-onset age) and to evaluate the accuracy of the EPS-DRE model, which distinguishes DRE from RSE. The area under the ROC curve (AUC) was categorized into the following discrimination abilities:[46] excellent discrimination, AUC of ≥0.90; good discrimination, 0.80≤AUC<0.90; fair discrimination, 0.70≤AUC<0.80; and poor discrimination, AUC<0.70. The Youden index[47] was used to determine the optimal cutoff value for estimating the effect size of the continuous variable. Analyses were carried out using SPSS statistics (version 26; IBM Corp., Armonk, NY, USA), and all tests of statistical significance were two-sided.

RESULTS

Demographic data

The test set consisted of 120 unrelated patients with RSE and

122 patients with DRE (1 patient who was previously classified into the RSE group was excluded from the analysis due to data ambiguity). The validation set consisted of 157 unrelated patients with RSE and 154 patients with DRE. There was no significant difference in sex (p=0.797) or drug response (p=0.864) between the two sets. The mean age at recruitment and the mean age at seizure onset were significantly higher in the validation set than the test set (43.6 vs. 39.6 years [p=0.001] and 25.8 vs. 20.6 years [p<0.001], respectively). The mean age at seizure onset was significantly higher in the RES group than in the DRE group (25.5 vs. 15.8 years [p<0.001] in the test set and 30.3 vs. 21.1 years [p<0.001] in the validation set), while the mean age at recruitment did not differ significantly between the RES and DRE groups in each set (Table 2).

Selection of candidate genetic and clinical predictors

WES on 242 samples in the test set yielded about 5.3×10^5 genetic variants. Among the 2,314 variants of 215 known epilepsy-associated genes with a read depth of ≥30×, 97 variants remained after applying a customized stringent filtering process to enhance data accuracy (Fig. 1). Since the genotyping call rate was complete for the filtered 97 variants, imputation was not required to test the association between each variant and DRE. To avoid the risk of ignoring information from large numbers of variants that are likely to be associated with DRE, we adopted a general approach to test the association based on a stepwise forward regression with $p\leq0.05$ rather than a stringent significance threshold for GWAS such as $p=5\times10^{-8}$.[35] The association study resulted in 44 variants being moved to estimate the effect size (OR) of each variant using simple regression analysis (Fig. 1). Finally, 11 variants of the following 9 genes with an estimated effect (the regression coefficient) of $p\leq0.05$ were chosen for calculating EPS-DRE: *DOCK7* (1), *LRP2* (1), *PLCB1* (2), *SIK1* (1), *GPR98* (1), *DIAPH1* (1), *LAMB1* (1), *CNTNAP2* (2), and *TSC1* (1) (Supplementary Table 1 in the online-only Data Supplement). Notably, there was a large racial difference in the reference-allele frequency of some variants based on gnomAD (version 2.1.1) (Supplementary Table 2 in the online-only Data Supplement).

The initial responses to treatment, the pretreatment seizure frequency, and symptomatic epilepsy were excluded from the analysis of the candidate clinical predictors of DRE. A previous retrospective study[37] found that the mean epilepsy duration of the participants with DRE was approximately 24 years in the test set (Table 2). Therefore, it was difficult to obtain the clinical records of the early treatment period after epilepsy diagnosis in most DRE patients. Since symptomatic epilepsy was defined based on neuroimaging findings in this study, it could be replaced by abnormal neuroimaging. The classification of abnormal EEG1 could be replaced by abnormal EEG2 (epileptiform discharges only) because these two classifications were strongly correlated (Pearson correlation coefficient r=1, p=0.01). Univariate regression analysis (Supplementary Table 3 in the online-only Data Supplement) revealed that

Table 2. Demographic and clinical characteristics of participants by disease status

Characteristic	Test set (n=242)		Validation set (n=311)		p
Drug response					0.864
RSE	120 (49.6)		157 (50.5)		
DRE	122 (50.4)		154 (49.5)		
Age (yr)					
At recruitment	39.6±13.3 (range, 20–84)		43.6±14.9 (range, 20–89)		0.001
RSE	39.3±15.1	(p=0.745)	42.9±16.8	(p=0.447)	
DRE	39.9±11.2		44.2±12.8		
At seizure onset	20.6±13.7 (range, 0–68)		25.8±16.8 (range, 0–80)		<0.001
RSE	25.5±15.2	(p<0.001)	30.3±17.4	(p<0.001)	
DRE	15.8±10.1		21.1±14.9		
Sex, male	128 (52.9)		168 (54.0)		0.797
RSE	64 (53.3)	(p=0.898)	88 (56.1)	(p=0.496)	
DRE	64 (52.5)		80 (51.9)		
Epilepsy duration (yr)	19.2±12.2		17.95±13.2		0.255
RSE	14.0±10.1	(p<0.001)	12.6±10.7	(p<0.001)	
DRE	24.3±11.9		23.4±13.2		
Abnormal imaging	(n=RSE:113, DRE:112)		(n=RSE:156, DRE:150)		0.593
RSE	38 (33.6)	($p\leq$0.001)	54 (34.6)	(p<0.001)	
DRE	61 (54.5)		102 (68.0)		

Values are presented as n (%) or mean±standard deviation unless otherwise indicated.
DRE, drug resistant epilepsy; RSE, drug-responsive epilepsy.

epilepsy duration (p<0.001), abnormal EEG2 (p<0.001), age at seizure onset (p<0.001), seizure type (p=0.002), and abnormal neuroimaging (p=0.002) were significantly associated with DRE. Age at seizure onset was strongly correlated with epilepsy duration (r=-0.464, p=0.01) and abnormal EEG2 (r=-0.197, p=0.01), and abnormal neuroimaging was strongly correlated with seizure type (r=0.233, p=0.01). Finally, age at seizure onset and abnormal neuroimaging were selected as candidate clinical predictors for calculating EPS-DRE. These clinical factors were preferred because they are less ambiguous, and moreover they can be obtained at earlier stages of epilepsy treatment; for example, epilepsy duration cannot be an important DRE predictor in the early stages of epilepsy treatment. Also, one of the most intractable types of seizures (a focal impaired awareness seizure without motor symptoms) might be underrecognized by clinicians who are not epilepsy specialists. This exclusion prevented overfitting of the prediction model and also simplified it, consistent with the aim of this study. ROC analysis determined the cutoff value for measuring the effect size (OR) of age at seizure onset as 13.5 years. The dosage of the effect was assigned as 1 if the age at seizure onset was ≤13.5 years, and as 0 otherwise. Simple regression analysis was performed to estimate the effect size.

Early prediction score for DRE

EPS-DRE of test set

The mean EPS-DRE$_{gen}$, EPS-DRE$_{cln}$, and EPS-DRE$_{mix}$ values in the RSE group differed significantly from the corresponding scores in the DRE group (p<0.001 for each test). Logistic regression analysis revealed that each EPS-DRE$_{gen}$, EPS-DRE$_{cln}$, and EPS-DRE$_{mix}$ model was significantly and independently associated with DRE: OR=3.725 (95% confidence interval [CI], 2.525−5.495), OR=2.915 (95% CI, 2.123−4.003), and OR=2.873 (95% CI, 2.195−3.761), respectively. The corresponding Nagelkerke R^2 values were 0.314, 0.271, and 0.446, respectively, indicating that the EPS-DRE$_{mix}$ model provided the best fit to the outcome data. In ROC analysis (Fig. 2), AUC was 0.777 (95% CI, 0.720−0.834), 0.766 (95% CI, 0.706−0.826), and 0.842 (95% CI, 0.793−0.890), respectively, indicating that the EPS-DRE$_{mix}$ model had good discrimination ability[46] in distinguishing DRE from RSE.

EPS-DRE of validation set

As in the test set, the mean EPS-DRE$_{gen}$, EPS-DRE$_{cln}$, and EPS-DRE$_{mix}$ values differed significantly between the RSE and DRE (p=0.004, p<0.001, and p< 0.001 in t-tests, respectively) and each EPS-DRE model was also independently associated with DRE: OR=1.416 (95% CI, 1.112−1.804), OR=2.514 (95% CI, 1.923−3.287), and OR=2.093 (95% CI, 1.692−2.590),

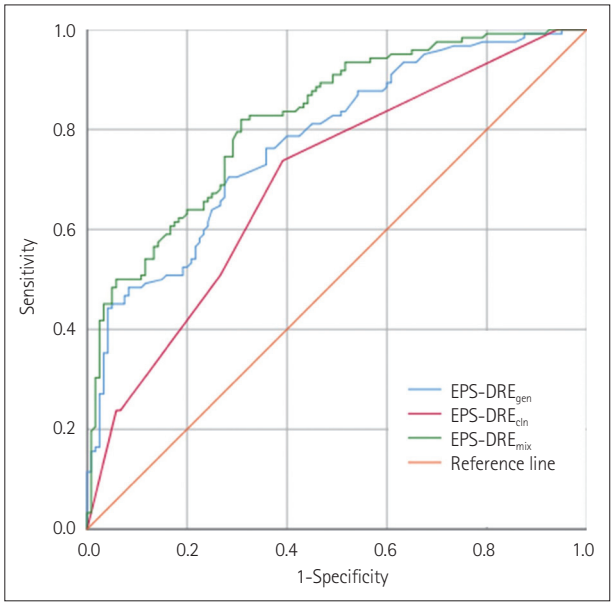

Fig. 2. Receiver operating characteristics (ROC) curves of three early prediction score for drug resistant epilepsy (EPS-DRE) models. The area under the ROC curve of EPS-DRE$_{mix}$ (green) was higher than that of EPS-DRE$_{cln}$ (red) or EPS-DRE$_{gen}$ (blue), indicating that EPS-DRE$_{mix}$ model was the best for distinguishing DRE from drug-responsive epilepsy. The reference line is the ROC curve corresponding to random chance.

respectively. Also like in the test set, Nagelkerke R^2 (0.232) was highest in the EPS-DRE$_{mix}$ model, whose discrimination ability was fair (AUC=0.739).[46]

DISCUSSION

It could have been assumed that the initial efforts to create PRS for common diseases were hampered by the smallness of the initial GWAS, which affected the precision of the estimated impact of individual variants on disease risk, and by the lack of large data sets needed to test and verify PRS.[34] In fact, summary statistics of a large GWAS data involving 184,305 participants were used to create the candidate PRSs of coronary artery disease (PRS$_{CAD}$) and a total of 6,630,150 variants were involved in calculation of the best PRS defined as the maximum AUC (0.81).[34] More recently, the best PRS for breast cancer (PRS$_{BREAST\ CANCER}$) with an AUC of 0.630 was developed using 313 variants and verified with a data set comprising about 30,000 participants.[35] These results showed that PRS is a powerful and reliable predictor of common polygenic diseases but the burden of developing a prediction model of common disease is pronounced. To make matter worse, PRS is essentially race-specific, and so individual PRS system for each race may be needed for multiracial societies. This is expensive and requires large amounts of time and knowledge of basic or clinical science that cannot be accessed at the individual level. This suggests that meeting the needs of individ-

ual races requires the size of the data set for calculating the PRS to be much smaller than in previous studies, while also maintaining a high prediction accuracy.

Generally, the larger number of predictors involved in the PRS calculation is, the better the fit of the PRS is.[36] As noted above, $PRS_{BREAST\ CANCER}$ and PRS_{CAD} needed hundreds to millions of predictors for developing PRSs with best performance.[34,35] However, in practice, the inclusion of too many predictors can be a barrier to clinical applications at the individual level due to the economic and time burdens of performing the associated genetic analyses. It is particularly interesting that recent studies have shown that the best-performing PRS does not necessarily involve the largest number of predictors; for example, the $PRS_{BREAST\ CANCER}$ calculated from 313 predictors (PRS_{313}) showed better performance than that calculated using 5,194 predictors ($PRS_{5,194}$).[34,35]

In this study, the EPS-DRE_{mix} model had a good discrimination ability in distinguishing between RSE and DRE (AUC=0.842), even though it was derived from much smaller data sets than in previous studies,[34,35] and using only 13 multidimensional predictors. This shows that assumptions about the size of the data set or the number of predictors required to develop the best PRS might be invalid, especially under certain conditions that can improve the performance of the PRS. The conditions underlying such a good performance of the EPS-DRE_{mix} model in this study were presumed as follows: First, DRE was more strictly defined than the conventional definition[38] in order to increase the phenotypic contrast between the case and control groups. Second, candidate genetic variants selected for calculating EPS-DRE_{mix} were restricted to known epilepsy-related genes or genes associated with potential mechanisms of DRE (Table 1). This consequently seems to rule out unnecessary false positive associations between possible confusing factors and DRE. Third, any ambiguity was also ruled out in selecting candidate genes and clinical predictors; for example, no imputation that can cause up to 22% of the wrong genotypes[48] was required, and no clinical predictors that can cause observer bias were used. Fourth, to improve the performance of DRE prediction model, a multidimensional approach was applied by combining genetic and clinical risk scores, as asserted by Rudolph et al.[49]

The present study has several limitations. First, the EPS-DRE_{mix} model is a preliminary result of research to explore the possibility of developing a DRE prediction model that can be easily used by nonexperts. Therefore, further work might be needed before it can be applied directly in the field. The discrimination ability of the EPS-DRE_{mix} model remained statistically significant for both the test and the validation sets, but its explanatory power (Nagelkerke R^2) in the validation set reduced compared to the test set. The difference between the two data sets may reflect overfitting associated with choosing the optimal threshold for the p value.[35] Adding new effective genetic or clinical DRE predictors not considered in the present study in future studies will improve the performance of the EPS-DRE_{mix} model. For example, a prospective study design could test whether the initial response to treatment or pretreatment seizure frequency can improve the model's performance. Second, since the allele frequencies, linkage disequilibrium patterns, and effect sizes of common polymorphisms vary with ancestry, certain PRSs for one racial group will not necessarily provide optimal predictive abilities for other racial groups.[34,50] Such differences by race were observed in this study (Supplementary Table 2 in the online-only Data Supplement), which means that the application of the EPS-DRE_{mix} model should be restricted to Koreans and East Asians. Third, candidate genetic predictors for calculating EPS-DRE_{mix} were restricted to known epilepsy-related genes and genes associated with the potential mechanisms of DRE, thus excluding the opportunity to identify new genes that cause DRE. Fourth, since no functional studies could be conducted to determine the impact of the finally selected candidate gene predictors on the DRE mechanism, further studies are warranted to determine how the 11 candidate genetic predictors contribute to the underlying mechanisms of DRE. Fifth, the main purpose of this study was to develop an early prediction model for DRE using unambiguous clinical information available at the time of diagnosis, and so the approach taken did not reflect epilepsy being a progressive disease. Sixth, since the EPS-DRE_{mix} model was developed based on the strict customized definition of DRE, care should be taken when applying the model to DRE patients identified using the conventional definition.[38]

In conclusion, the new multidimensional EPS-DRE_{mix} model developed in this study showed good discrimination ability in distinguishing between RSE and DRE, even though it was derived from a small GWAS data set and only 11 genetic and 2 clinical predictors. This suggests that the performance of the EPS-DRE is not as strictly dependent on the size of the GWAS data set or the number of predictors included when developing the score, which contrasts with previous reports.[34,35] The present results suggest that the performance can be improved by developing and adding predictive factors that have sufficiently large effect size rather than by maximizing the size of the data set. In addition, the predictors involved in the calculating the EPS-DRE_{mix} model in this study are easily accessible at the time of an epilepsy diagnosis without ambiguity, indicating that clinical application of the EPS-DRE_{mix} model in the early stage of medical care will be sufficiently possible at the individual level of nonexperts. It is hoped that these preliminary results will contribute to the early identifi-

cation of DRE, especially by general neurologists and physicians who do not have detailed knowledge or equipment for DRE.

Supplementary Materials

The online-only Data Supplement is available with this article at https://doi.org/10.3988/jcn.2022.18.5.553.

Availability of Data and Material

The datasets generated or analyzed during the study are available from the corresponding author on reasonable request.

ORCID iDs

Kyung Wook Kang	https://orcid.org/0000-0001-9362-8670
Yong Won Cho	https://orcid.org/0000-0002-6127-1045
Sang Kun Lee	https://orcid.org/0000-0003-1908-0699
Ki-Young Jung	https://orcid.org/0000-0001-5528-9081
Ji Hyun Kim	https://orcid.org/0000-0003-3411-5714
Dong Wook Kim	https://orcid.org/0000-0003-4484-0602
Sang-Ahm Lee	https://orcid.org/0000-0002-6743-0545
Seung Bong Hong	https://orcid.org/0000-0002-8933-5709
In-Seop Na	https://orcid.org/0000-0001-6471-043X
So-Hyun Lee	https://orcid.org/0000-0001-9992-6038
Won-Ki Baek	https://orcid.org/0000-0001-9123-4096
Seok-Yong Choi	https://orcid.org/0000-0001-7193-292X
Myeong-Kyu Kim	https://orcid.org/0000-0001-8673-7561

Author Contributions

Conceptualization: Kyung Wook Kang, Myeong-Kyu Kim, Seok-Yong Choi. Data curation: Kyung Wook Kang, Myeong-Kyu Kim, Yong Won Cho, Sang Kun Lee, Ki-Young Jung, Ji Hyun Kim, Dong Wook Kim, Sang-Ahm Lee, Seung Bong Hong, Won-Ki Baek, In-Seop Na, So-Hyun Lee. Formal analysis: Kyung Wook Kang, Myeong-Kyu Kim, Seok-Yong Choi, So-Hyun Lee. Funding acquisition: Myeong-Kyu Kim, Seok-Yong Choi, Seung Bong Hong. Investigation: all authors. Methodology: Kyung Wook Kang, Seok-Yong Choi, Myeong-Kyu Kim, So-Hyun Lee. Project administration: Seok-Yong Choi, Myeong-Kyu Kim. Resources: all authors. Software: Kyung Wook Kang, Seok-Yong Choi, Myeong-Kyu Kim. Supervision: Seok-Yong Choi, Myeong-Kyu Kim. Validation: Kyung Wook Kang, Seok-Yong Choi, Myeong-Kyu Kim. Visualization: Kyung Wook Kang, Seok-Yong Choi, Myeong-Kyu Kim, So-Hyun Lee. Writing—original draft: Kyung Wook Kang, Myeong-Kyu Kim, Seok-Yong Choi. Writing—review & editing: Myeong-Kyu Kim, Seok-Yong Choi, Yong Won Cho, Sang Kun Lee, Ki-Young Jung, Ji Hyun Kim, Dong Wook Kim, Sang-Ahm Lee, Seung Bong Hong, Won-Ki Baek, In-Seop Na, So-Hyun Lee.

Conflicts of Interest

The authors have no potential conflicts of interest to disclose.

Funding Statement

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) grant funded by the Ministry of Education (2020R1I1A3073861), Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI) funded by the Ministry of Health & Welfare, Republic of Korea (HI15C1559), and the budget for epilepsy support system project funded by Ministry of Health and Welfare (2020).

Acknowledgements

We are grateful to the patients for their help and participation in the study. We would like to thank Editage for English language editing.

REFERENCES

1. Fiest KM, Sauro KM, Wiebe S, Patten SB, Kwon CS, Dykeman J, et al. Prevalence and incidence of epilepsy: a systematic review and meta-analysis of international studies. *Neurology* 2017;88:296-303.
2. Chen Z, Brodie MJ, Liew D, Kwan P. Treatment outcomes in patients with newly diagnosed epilepsy treated with established and new antiepileptic drugs: a 30-year longitudinal cohort study. *JAMA Neurol* 2018;75:279-286.
3. Kwan P, Brodie MJ. Early identification of refractory epilepsy. *N Engl J Med* 2000;342:314-319.
4. McCagh J, Fisk JE, Baker GA. Epilepsy, psychosocial and cognitive functioning. *Epilepsy Res* 2009;86:1-14.
5. Mohanraj R, Norrie J, Stephen LJ, Kelly K, Hitiris N, Brodie MJ. Mortality in adults with newly diagnosed and chronic epilepsy: a retrospective comparative study. *Lancet Neurol* 2006;5:481-487.
6. Lawn ND, Bamlet WR, Radhakrishnan K, O'Brien PC, So EL. Injuries due to seizures in persons with epilepsy: a population-based study. *Neurology* 2004;63:1565-1570.
7. Begley CE, Durgin TL. The direct cost of epilepsy in the United States: a systematic review of estimates. *Epilepsia* 2015;56:1376-1387.
8. Wiebe S, Blume WT, Girvin JP, Eliasziw M; Effectiveness and Efficiency of Surgery for Temporal Lobe Epilepsy Study Group. A randomized, controlled trial of surgery for temporal-lobe epilepsy. *N Engl J Med* 2001;345:311-318.
9. Spencer SS, Berg AT, Vickrey BG, Sperling MR, Bazil CW, Shinnar S, et al. Initial outcomes in the multicenter study of epilepsy surgery. *Neurology* 2003;61:1680-1685.
10. Engel J Jr, Wiebe S, French J, Sperling M, Williamson P, Spencer D, et al. Practice parameter: temporal lobe and localized neocortical resections for epilepsy: report of the Quality Standards Subcommittee of the American Academy of Neurology, in association with the American Epilepsy Society and the American Association of Neurological Surgeons. *Neurology* 2003;60:538-547.
11. Engel J Jr, McDermott MP, Wiebe S, Langfitt JT, Stern JM, Dewar S, et al. Early surgical therapy for drug-resistant temporal lobe epilepsy: a randomized trial. *JAMA* 2012;307:922-930.
12. Jeha LE, Najm IM, Bingaman WE, Khandwala F, Widdess-Walsh P, Morris HH, et al. Predictors of outcome after temporal lobectomy for the treatment of intractable epilepsy. *Neurology* 2006;66:1938-1940.
13. Dwivedi R, Ramanujam B, Chandra PS, Sapra S, Gulati S, Kalaivani M, et al. Surgery for drug-resistant epilepsy in children. *N Engl J Med* 2017;377:1639-1647.
14. Bjellvi J, Olsson I, Malmgren K, Wilbe Ramsay K. Epilepsy duration and seizure outcome in epilepsy surgery: a systematic review and meta-analysis. *Neurology* 2019;93:e159-e166.
15. Perry MS, Duchowny M. Surgical versus medical treatment for refractory epilepsy: outcomes beyond seizure control. *Epilepsia* 2013;54:2060-2070.
16. Meguins LC, Adry RA, Silva-Junior SC, Araújo Filho GM, Marques LH. Shorter epilepsy duration is associated with better seizure outcome in temporal lobe epilepsy surgery. *Arq Neuropsiquiatr* 2015;73:212-217.
17. Pelliccia V, Deleo F, Gozzo F, Sartori I, Mai R, Cossu M, et al. Early and late epilepsy surgery in focal epilepsies associated with long-term epilepsy-associated tumors. *J Neurosurg* 2017;127:1147-1152.
18. Ramantani G, Zentner J. Epilepsy surgery in children and adolescents. *Neurol Int Open* 2017;1:E86-E97.
19. Ko A, Lee JS. Factors associated with seizure and cognitive outcomes after epilepsy surgery for low-grade epilepsy-associated neuroepithelial tumors in children. *Clin Exp Pediatr* 2020;63:171-177.
20. Gilliam F, Kuzniecky R, Meador K, Martin R, Sawrie S, Viikinsalo M, et al. Patient-oriented outcome assessment after temporal lobectomy for refractory epilepsy. *Neurology* 1999;53:687-694.
21. Berg AT, Langfitt J, Shinnar S, Vickrey BG, Sperling MR, Walczak T,

et al. How long does it take for partial epilepsy to become intractable? *Neurology* 2003;60:186-190.
22. Benbadis SR, Heriaud L, Tatum WO, Vale FL. Epilepsy surgery, delays and referral patterns-are all your epilepsy patients controlled? *Seizure* 2003;12:167-170.
23. Choi H, Carlino R, Heiman G, Hauser WA, Gilliam FG. Evaluation of duration of epilepsy prior to temporal lobe epilepsy surgery during the past two decades. *Epilepsy Res* 2009;86:224-227.
24. Haneef Z, Stern J, Dewar S, Engel J Jr. Referral pattern for epilepsy surgery after evidence-based recommendations: a retrospective study. *Neurology* 2010;75:699-704.
25. Engel J Jr. The timing of surgical intervention for mesial temporal lobe epilepsy: a plan for a randomized clinical trial. *Arch Neurol* 1999;56:1338-1341.
26. Dlugos DJ. The early identification of candidates for epilepsy surgery. *Arch Neurol* 2001;58:1543-1546.
27. Berg AT. Understanding the delay before epilepsy surgery: who develops intractable focal epilepsy and when? *CNS Spectr* 2004;9:136-144.
28. Hakimi AS, Spanaki MV, Schuh LA, Smith BJ, Schultz L. A survey of neurologists' views on epilepsy surgery and medically refractory epilepsy. *Epilepsy Behav* 2008;13:96-101.
29. Tang F, Hartz AM, Bauer B. Drug-resistant epilepsy: multiple hypotheses, few answers. *Front Neurol* 2017;8:301.
30. Sisodiya SM. Mechanisms of antiepileptic drug resistance. *Curr Opin Neurol* 2003;16:197-201.
31. Gibson G. Rare and common variants: twenty arguments. *Nat Rev Genet* 2012;13:135-145.
32. Golan D, Lander ES, Rosset S. Measuring missing heritability: inferring the contribution of common variants. *Proc Natl Acad Sci U S A* 2014;111:E5272-E5281.
33. Fuchsberger C, Flannick J, Teslovich TM, Mahajan A, Agarwala V, Gaulton KJ, et al. The genetic architecture of type 2 diabetes. *Nature* 2016;536:41-47.
34. Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet* 2018;50:1219-1224.
35. Mavaddat N, Michailidou K, Dennis J, Lush M, Fachal L, Lee A, et al. Polygenic risk scores for prediction of breast cancer and breast cancer subtypes. *Am J Hum Genet* 2019;104:21-34.
36. Choi SW, Mak TS, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. *Nat Protoc* 2020;15:2759-2772.
37. Kang KW, Kim W, Cho YW, Lee SK, Jung KY, Shin W, et al. Genetic characteristics of non-familial epilepsy. *PeerJ* 2019;7:e8278.
38. Kwan P, Arzimanoglou A, Berg AT, Brodie MJ, Allen Hauser W, Mathern G, et al. Definition of drug resistant epilepsy: consensus proposal by the ad hoc Task Force of the ILAE Commission on therapeutic strategies. *Epilepsia* 2010;51:1069-1077.
39. Wang J, Lin ZJ, Liu L, Xu HQ, Shi YW, Yi YH, et al. Epilepsy-associated genes. *Seizure* 2017;44:11-20.
40. Lascano AM, Korff CM, Picard F. Seizures and epilepsies due to channelopathies and neurotransmitter receptor dysfunction: a parallel between genetic and immune aspects. *Mol Syndromol* 2016;7:197-209.
41. Zhang Z, Gao K, Liu Q, Zhou J, Li X, Lang N, et al. Somatic variants in new candidate genes identified in focal cortical dysplasia type II. *Epilepsia* 2020;61:667-678.
42. Kim MK, Cho YW, Shin WC, Cho JW, Shon YM, Kim JH, et al. Association of restless legs syndrome variants in Korean patients with restless legs syndrome. *Sleep* 2013;36:1787-1791.
43. Kwan P, Schachter SC, Brodie MJ. Drug-resistant epilepsy. *N Engl J Med* 2011;365:919-926.
44. Kalilani L, Sun X, Pelgrims B, Noack-Rink M, Villanueva V. The epidemiology of drug-resistant epilepsy: a systematic review and meta-analysis. *Epilepsia* 2018;59:2179-2193.
45. Trinka E, Cock H, Hesdorffer D, Rossetti AO, Scheffer IE, Shinnar S, et al. A definition and classification of status epilepticus--Report of the ILAE Task Force on classification of status epilepticus. *Epilepsia* 2015;56:1515-1523.
46. Muller MP, Tomlinson G, Marrie TJ, Tang P, McGeer A, Low DE, et al. Can routine laboratory tests discriminate between severe acute respiratory syndrome and other causes of community-acquired pneumonia? *Clin Infect Dis* 2005;40:1079-1086.
47. Fluss R, Faraggi D, Reiser B. Estimation of the Youden Index and its associated cutoff point. *Biom J* 2005;47:458-472.
48. Schurz H, Müller SJ, van Helden PD, Tromp G, Hoal EG, Kinnear CJ, et al. Evaluating the accuracy of imputation methods in a five-way admixed population. *Front Genet* 2019;10:34.
49. Rudolph A, Song M, Brook MN, Milne RL, Mavaddat N, Michailidou K, et al. Joint associations of a polygenic risk score and environmental risk factors for breast cancer in the Breast Cancer Association Consortium. *Int J Epidemiol* 2018;47:526-536.
50. Martin AR, Gignoux CR, Walters RK, Wojcik GL, Neale BM, Gravel S, et al. Human demographic history impacts genetic risk prediction across diverse populations. *Am J Hum Genet* 2017;100:635-649.