ORIGINAL INVESTIGATION

Genetic risk score analysis indicates migraine with and without comorbid depression are genetically different disorders

Lannie Ligthart · Jouke-Jan Hottenga · Cathryn M. Lewis · Anne E. Farmer · Ian W. Craig · Gerome Breen · Gonneke Willemsen · Jacqueline M. Vink · Christel M. Middeldorp · Enda M. Byrne · Andrew C. Heath · Pamela A. F. Madden · Michele L. Pergadia · Grant W. Montgomery · Nicholas G. Martin · Brenda W. J. H. Penninx · Peter McGuffin · Dorret I. Boomsma · Dale R. Nyholt

Abstract Migraine and major depressive disorder (MDD) are comorbid, moderately heritable and to some extent influenced by the same genes. In a previous paper, we suggested the possibility of causality (one trait causing the other) underlying this comorbidity. We present a new application of polygenic (genetic risk) score analysis to investigate the mechanisms underlying the genetic overlap of migraine and MDD. Genetic risk scores were constructed based on data from two discovery samples in which genome-wide association analyses (GWA) were performed for migraine and MDD, respectively. The Australian Twin Migraine GWA study ($N = 6{,}350$) included 2,825 migraine cases and 3,525 controls, 805 of whom met the diagnostic criteria for MDD. The RADIANT GWA study ($N = 3{,}230$) included 1,636 MDD cases and 1,594 controls. Genetic risk scores for migraine and for MDD were used to predict pure and comorbid forms of migraine and MDD in an independent Dutch target sample (NTR–NESDA, $N = 2{,}966$), which included 1,476 MDD cases and 1,058 migraine cases (723 of these individuals had both disorders concurrently). The observed patterns of prediction suggest that the 'pure' forms of migraine and MDD are genetically distinct disorders. The subgroup of individuals with comorbid MDD and migraine were genetically most similar to MDD patients. These results indicate that in at least a subset of migraine patients with MDD, migraine may be a symptom or consequence of MDD.

Introduction

There is a well-established comorbidity between migraine and major depressive disorder (MDD), with a 2–3 times

D. I. Boomsma and D. R. Nyholt contributed equally to this work.

L. Ligthart (✉) · J.-J. Hottenga · G. Willemsen · J. M. Vink · C. M. Middeldorp · D. I. Boomsma
Department of Biological Psychology, VU University, van der Boechorststraat 1, 1081 BT Amsterdam, The Netherlands
e-mail: rsl.ligthart@vu.nl

L. Ligthart
EMGO+ Institute for Health and Care Research, VU University Medical Center, van der Boechorststraat 7, 1081 BT Amsterdam, The Netherlands

C. M. Lewis · A. E. Farmer · I. W. Craig · G. Breen · P. McGuffin
MRC SGDP Centre, Institute of Psychiatry, PO80, DeCrespigny Park, Denmark Hill, London SE5 8AF, UK

C. M. Middeldorp
Department of Child and Adolescent Psychiatry, GGZ inGeest/VU University Medical Center, A.J. Ernststraat 1187, 1081 HL Amsterdam, The Netherlands

E. M. Byrne
Queensland Brain Institute, The University of Queensland, QBI Building (#79), St Lucia, QLD 4072, Australia

A. C. Heath · P. A. F. Madden · M. L. Pergadia
Department of Psychiatry, Washington University School of Medicine, 660 S. Euclid Ave, St. Louis, MO 63110, USA

G. W. Montgomery · N. G. Martin · D. R. Nyholt
QIMR Berghofer Medical Research Institute, Royal Brisbane Hospital, Locked Bag 2000, Herston, QLD 4029, Australia

B. W. J. H. Penninx
Department of Psychiatry, VU University Medical Center, A.J. Ernststraat 1187, 1081 HL Amsterdam, The Netherlands

higher risk of migraine in MDD patients than in non-MDD controls (Ligthart et al. 2010b). Both conditions are around 40–50 % heritable (Levinson 2006; Mulder et al. 2003), and thought to be influenced by a large number of genetic variants with small effects. Longitudinal studies show mixed results with respect to the order of onset of migraine and depression (Antonaci et al. 2011; Swartz et al. 2000). There is some evidence suggesting that anxiety disorders tend to precede migraine, whereas depression tends to follow migraine (Breslau et al. 1991; Merikangas et al. 1990), but a bidirectional relationship has also been reported, with migraine predicting depression onset and vice versa (Breslau et al. 2000, 2003). We previously showed, in a bivariate twin study, that migraine and depression are partly influenced by the same genes, with a correlation of around 0.3 between the genetic factors affecting the two disorders (Ligthart et al. 2010a). Several other studies reported similar results (Schur et al. 2009; Stam et al. 2010).

While genetic correlations are often interpreted as the same genes directly affecting multiple conditions, alternative explanations are possible, including cross-assortative mating (Hazel 1943) or the causal effect of one trait on the other (De Moor et al. 2008). In our twin study, we found that in monozygotic twin pairs discordant for anxious depression, the twins with a high anxious depression score had an increased risk of migraine; their non-anxious-depressed co-twins did not, even though they had the same genetic make-up as the anxious-depressed twin. Similarly, in pairs discordant for migraine, only twins with migraine had an increased risk of anxious depression (Ligthart et al. 2010a). This suggests that a causal relationship (one trait causing the other), rather than genes directly involved in multiple disorders (pleiotropy), may explain the comorbidity of migraine and depression (De Moor et al. 2008; Ligthart and Boomsma 2012).

A possible causal explanation is that a severe pain disorder such as migraine might reduce a patient's quality of life, resulting in depressive symptoms. However, a study on back pain and depression found no relationship between pain duration and depression severity, which would be expected if it is the long-lasting pain and disability that causes the depressive symptoms (Von Korff and Simon 1996). Alternatively, migraine might be a symptom or consequence of depression. Pain symptoms are remarkably common in depressed patients, and some have argued that pain should be viewed as an integral part of the depression phenotype (Lépine and Briley 2004; Stahl 2002). Currently, solid evidence for this hypothesis is lacking.

Here, we investigate the comorbidity of migraine and depression at the measured genotype level, by using genetic risk scores (based on SNP risk alleles) obtained in one (discovery) sample to predict migraine and depression status in another, independent (target) sample. With this method, the combined effects of a large number of SNPs on a trait or disorder, estimated in a discovery sample, are used to construct a 'genetic risk score' for individuals in an independent target sample. The genetic risk score is a single quantitative index of genetic load for the discovery trait, and can be used to 'predict' phenotype status in individuals in the target sample (Purcell et al. 2009). This method allows the comparison of specific groups of patients in terms of their genetic similarity, and has previously been applied to demonstrate overlap in genetic influences across disorders, such as schizophrenia and bipolar disorder (Purcell et al. 2009), and anxiety and depression (Demirkan et al. 2010).

Here, we apply this method to investigate possible causal relationships between two phenotypes. Our first hypothesis is that some migraines are a consequence of depression (or are a symptom of depression; since these two hypotheses are closely related and yield the same predictions we will not distinguish between them). According to this hypothesis, migraine due to depression will be genetically similar to depression, but not to migraine without depression (further referred to as pure migraine). This is based on the principle that if disease A causes disease B, the genes that affect disease A must indirectly also affect disease B, through the causal chain (De Moor et al. 2008; Ligthart and Boomsma 2012).

Two alternative hypotheses were considered. Hypothesis 2 was the reverse of hypothesis 1: depression can be a consequence (or symptom) of migraine, and depression due to migraine is genetically similar to migraine, but not to depression without migraine (further referred to as pure depression). Hypothesis 3 was that there is only one type of migraine and one type of depression, which are in part influenced by genes involved in both disorders. However, the latter is not expected given our previous findings (Ligthart et al. 2010a).

Materials and methods

Discovery sample migraine: the Australian twin migraine (ATM) GWA study

The ATM study includes data from Australian twins and their families. Detailed migraine data were collected by means of semi-structured telephone interviews, with a mean age at interview of 37.5 years (SD = 11.3) (Nyholt et al. 2004, 2005). The questionnaire was based on the symptoms from the ICHD-II diagnostic criteria (Headache Classification Committee of the International Headache Society 2004). All subjects gave informed consent and approval to conduct the research was obtained from the QIMR Human Research Ethics Committee. Participants were classified as affected or unaffected with latent class analysis (LCA), as described

Table 1 Sample sizes and number of cases with migraine and/or MDD

Australian twin migraine (ATM) study (discovery)	Total sample	MDD cases	Excluding MDD
Total sample	6,350	805	5,545
Migraine cases	2,825	670	2,155
Controls	3,525	135	3,390
RADIANT MDD study (discovery)	Total sample		
Total sample	3,230	–	–
MDD cases	1,636	–	–
Controls	1,594	–	–
NTR–NESDA study (target)	Total sample	NTR	NESDA
Total sample	2,966	1,442	1,524
Controls	1,155	1,050	105
Pure migraine	335	296	39
Pure MDD	753	48	705
Comorbid MDD + migraine	723	48	675
Total MDD cases	1,476	96	1,380
Total migraine cases	1,058	344	714

in detail in previous work (Nyholt et al. 2004, 2005). All genome-wide association (GWA) cases screened positive for migraine according to ICHD-II migraine with aura (MA) and/or migraine without aura (MO) diagnostic criteria, or the ID Migraine™ Screener (Lipton et al. 2003), which has been shown to accurately identify 93 % of people with migraines. Additional migraine cases were identified from other questionnaire data as those self-reporting ("yes" or "no") to suffering migraine. All cases and controls were unrelated; one individual was selected from each family. The cases [$N = 2,825$; 752 (27 %) male, 2,073 (73 %) female] were preferentially selected from each family based on migraine severity. The population controls [$N = 3,525$; 1,814 (51 %) male, 1,711 (49 %) female] were randomly selected from families containing no known migraine cases. There were 805 individuals known to have MDD, based on DSM-IV criteria (Wray et al. 2010); after exclusion of these MDD cases there were 2,155 cases and 3,390 controls (for details see Table 1).

The subjects were drawn from the QIMR GWA cohort of over 19,000 individuals genotyped using a variety of Illumina GWA arrays. After strict QC, allelic association analysis was performed using the PLINK program. For a detailed description of the QIMR 19K GWA cohort, including QC, see Medland et al. (2009).

Discovery sample MDD: the RADIANT study

The UK RADIANT study consisted of 1,636 depression cases and 1,594 non-MDD controls from the UK (Table 1). The participants came from two studies on recurrent depression [the Depression Case Control (DeCC) study and the Depression Network (DeNT) study] and a pharmacogenetic study [the Genome-Based Therapeutic Drugs from Depression (GENDEP) study], which have been described in more detail previously (Cohen-Woods et al. 2009; Farmer et al. 2004; Uher et al. 2009). The mean age of onset was 23.1 years (SD = 11.4) and 66.4 % of cases were female.

MDD was diagnosed with the Schedules for Clinical Assessment in Neuropsychiatry (SCAN) interview. A case was defined as a patient with recurrent depression (DeCC and DeNT) or experiencing a current episode of depression (GENDEP) of at least moderate severity, according to either DSM-IV or ICD-10 criteria. The controls were volunteers recruited via the Medical Research Council general practice research framework and staff and students of King's College London, who were screened for lifetime depressive symptoms with a telephone interview.

Genotyping was performed using the Illumina Human Hap610-Quad BeadChips. Exclusion criteria were a missing genotype rate >1 %, abnormal heterozygosity, inconsistencies with respect to sex assignment, relatedness with other participants or non-Caucasian ancestry. SNPs were excluded based on MAF <1 %, or significant departure from HWE ($P < 1 \times 10^{-5}$). Genome-wide association analysis was performed with logistic regression analysis, assuming a log-additive model and including ancestry principal components as covariates. Details can be found elsewhere (Lewis et al. 2010).

Target sample: the NTR–NESDA study

The target sample included participants from the Netherlands Twin Registry (NTR) (Boomsma et al. 2006) and the

Netherlands Study of Depression and Anxiety (NESDA) (Penninx et al. 2008). Genotyping was performed in the context of the GAIN study on MDD (Boomsma et al. 2008; Sullivan et al. 2008). The NTR data were part of an ongoing longitudinal study on health, personality and lifestyle in Dutch twin families (Boomsma et al. 2006). NESDA is a longitudinal study on anxiety and depression, with participants from the general population, primary care, and specialised mental health care (Penninx et al. 2008). The mean age of the participants was 42.3 years (SD = 16.5; 65 % female).

Cases with lifetime MDD ($N = 1{,}480$) were diagnosed based on DSM-IV criteria with the Composite International Diagnostic Interview (Wittchen 1994). The control group ($N = 1{,}491$) consisted of individuals selected for low risk of MDD based on measures of depression, anxiety and neuroticism (Boomsma et al. 2008). Migraine was measured with questionnaire items based on the symptoms included in the ICHD-II criteria for migraine (2004). As in the ATM study, affection status was determined by means of LCA, as described in previous work (Ligthart et al. 2006). Table 1 provides an overview of the numbers of patients with "pure" and comorbid forms of migraine and MDD.

The analyses in this paper are based on 2,966 unrelated individuals from the GAIN-MDD study (Boomsma et al. 2008; Sullivan et al. 2008) with data on both migraine and MDD, who were genotyped on the Affymetrix Perlegen 5.0 platform. Genotype imputation was performed on a larger dataset, including samples genotyped on several different platforms. Inclusion criteria for SNPs were MAF >1 %, HWE $P > 0.00001$ and call rate >95 %. Individuals were excluded when their expected sex and IBD status did not match, when the mean heterozygosity was <0.30 or >0.35, or when the missing genotype rate exceeded 10 %. All SNPs were aligned to the positive strand of the Hapmap 2 Build 36 release 24 CEU reference set. The alignment was checked using individuals and family members tested on multiple platforms. SNPs were excluded per platform if allele frequencies differed more than 15 % with the reference set and/or the other platforms. The data were subsequently merged into a single dataset ($N = 5856$), and imputed against the reference set using IMPUTE v2 (Marchini and Howie 2010). After imputation, genotype dosage was calculated if the highest genotype probability was above 90 %. Badly imputed SNPs were removed based on HWE $P < 0.00001$, proper info <0.40, MAF <1 %, allele frequency difference >0.15 against reference.

Genetic risk score analyses

To explore whether common variants with small effects on MDD and migraine status together predicted affection status in an independent sample, a genetic risk score test was performed following the method introduced in Purcell et al. (2009), and implemented in the PLINK program (Purcell et al. 2007). Twelve different SNP sets were selected for the prediction analyses (with P value thresholds at 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0). Only SNPs directly genotyped in the discovery samples were included. The number of SNPs overlapping between discovery and target samples was 249,917 for ATM and NTR–NESDA, 416,030 for RADIANT and NTR–NESDA, and 229,778 SNPs were available in all three samples. The numbers of SNPs included in each of the analyses are listed in Table 2. For each individual in the target sample, genetic risk scores were calculated based on the estimated number of 'risk alleles' an individual possessed (accounting for uncertainty of imputation in the target sample), weighted by the log odds ratio (i.e., effect size) for that allele calculated in the discovery sample. With this procedure, three genetic risk scores were calculated for each individual, one based on RADIANT MDD, one based on ATM migraine and one based on ATM migraine with MDD cases excluded.

Next, it was tested whether higher mean genetic risk scores were observed for cases than for controls in the target sample. Several different definitions of case status were used: all migraine, all MDD, pure migraine, pure MDD, and comorbid MDD + migraine. The cases were compared with controls unaffected for both migraine and MDD. The association between mean genetic risk score and case/control status in the target sample was tested using logistic regression analysis in R v2.11.0. Nagelkerke's pseudo R^2 was used to assess the variance explained. In addition, to assess predictive ability, the area under the receiver operating characteristic (ROC) curve (AUC) was calculated using the PredictABEL R package (Kundu et al. 2011). Given the unequal sex ratio of migraine and MDD, prediction analyses were compared including or excluding sex as a covariate.

Under all hypotheses, we expect prediction of a phenotype from genetic risk scores based on the same phenotype in another sample to be significant. Prediction of migraine from depression scores or vice versa will be somewhat weaker but still significant. The hypotheses yield different expected predictions when the "pure" forms of either migraine or depression are predicted from the other phenotype (see Table 3).

When pure migraine is predicted from depression genetic risk scores, we expect no prediction under hypothesis 1 (because pure migraine is genetically distinct from MDD). Whereas we expect to see modest but significant prediction under hypothesis 2 (because some cases of MDD are a symptom or consequence of migraine, and therefore, migraine risk variants will be part of the MDD risk profile). Modest but significant prediction is also expected under hypothesis 3 (because migraine and MDD

Table 2 Number of SNPs included in each prediction analysis

P value threshold	SNPs overlapping with NTR/NESDA			SNPs overlapping between all three samples		
	RADIANT	ATM (all)	ATM (exMDD)	RADIANT	ATM (all)	ATM (exMDD)
0.01	4,657	2,564	2,585	2,615	2,345	2,354
0.05	21,747	12,673	12,596	12,069	11,629	11,575
0.1	43,046	25,444	25,253	23,819	23,385	23,218
0.2	85,330	50,453	50,185	47,374	46,324	46,112
0.3	126,854	75,441	75,397	70,273	69,334	69,219
0.4	168,637	100,402	100,348	93,272	92,303	92,223
0.5	210,168	125,291	125,199	115,982	115,221	115,103
0.6	251,509	150,347	150,034	138,720	138,303	137,932
0.7	292,601	175,040	174,868	161,419	161,029	160,820
0.8	333,739	199,809	199,747	184,041	183,776	183,676
0.9	374,973	225,035	224,827	207,025	206,942	206,739
1.0	416,030	249,917	249,917	229,778	229,778	229,778

The number of SNPs used in the prediction analyses, for each *P* value threshold. On the left the numbers for the set of SNPs in each sample that overlapped with the NTR–NESDA target sample (as used in the analyses presented in Fig. 1), on the right the numbers for the set of SNPs that overlapped between all three samples (as used in the analyses presented in Fig. 4)

Table 3 Expected patterns of prediction under the three hypotheses

Genetic risk scores based on:	Predicted phenotype (NTR–NESDA)	Significant prediction expected (yes/no)		
		Hypothesis 1: migraine due to MDD	Hypothesis 2: MDD due to migraine	Hypothesis 3: shared genes
All MDD (RADIANT)	Pure migraine	No	Yes	Yes
All migraine (ATM)	Pure MDD	Yes	No	Yes
Pure migraine (ATM, excl. MDD cases)	Pure MDD	No	No	Yes

are partly affected by the same genetic risk factors, and therefore migraine patients will possess some risk variants for MDD, regardless of MDD status).

When pure depression is predicted from migraine genetic risk scores, expectations under hypothesis 1 will depend on whether MDD is predicted from all migraine or pure migraine. Pure depression prediction from all migraine is expected to be modest but significant (because all migraine is a combination of pure migraine and migraine due to MDD). Whereas pure depression predicted from pure migraine is not expected (because pure migraine and pure MDD are genetically distinct disorders). Under hypothesis 2, we expect to see no significant prediction (because pure MDD cases do not have migraine and therefore do not share migraine risk factors), and under hypothesis 3, the prediction is expected to be modest but significant (because migraine and MDD are partly affected by the same genetic risk factors).

Because there were different numbers of cases for the different phenotypes, the sample sizes differed considerably between analyses. To assess how these differences in sample size might affect our results, the smallest sample size in the target sample was taken (335 cases for pure migraine vs. 1,155 controls, see Table 1) and the prediction

analyses for the other phenotypes were repeated 10,000 times with random samples of 335 cases drawn from the total sample available for each phenotype. Mean R^2 and P values across the 10,000 permutations were calculated in order to assess the effect of sample size on the significance and explained variance, compared with the original analyses. The number of SNPs also differed between the RADIANT and ATM samples. Therefore, the analyses were repeated using only the SNPs available in all three samples, to test to what extent the results were affected by the number of SNPs included.

SNP effect (direction) concordance analysis

To further investigate the polygenic overlap of migraine (before and after excluding MDD individuals) with pure MDD, we examined the correlation of individual SNP effects across the datasets. That is, we tested whether the direction of SNP effects were positively correlated across the (i) all migraine (ATM) vs. pure MDD (NTR–NESDA), and (ii) pure migraine (ATM excluding MDD cases) vs. pure MDD (NTR–NESDA) GWA results. Firstly, the SNP effects [i.e., odds ratio (OR)] were aligned to the same reference allele at each SNP across the two GWA results

(ATM and NTR–NESDA). After alignment, a subset of independent SNPs was extracted via linkage disequilibrium (LD) clumping (Purcell et al. 2007), i.e., removal of SNPs in LD with $R^2 > 0.1$ estimated using HapMap 3 (release 2) CEU SNP genotype data. To test whether the effect directions were concordant across the datasets, we performed Fisher's Exact statistical tests of SNP effects in dataset 1 (OR_1) and dataset 2 (OR_2) after conditioning on their GWA P values. Briefly, for each subset, Fisher's Exact tests were performed on 2×2 tables containing the number of SNPs with both $OR_1 < 1$ and $OR_2 < 1$ ($OR_1 < 1 \bigcap OR_2 < 1$), $OR_1 > 1 \bigcap OR_2 < 1$, $OR_1 < 1 \bigcap OR_2 > 1$, and $OR_1 > 1 \bigcap OR_2 > 1$. SNP subsets were generated utilising the same 12 P value thresholds used in the genetic risk score analyses in ATM and NTR–NESDA. To examine the overall concordance in SNP effects between two datasets, we calculated the proportion of the total (144) threshold combination subsets where the SNP effects were nominally correlated [Fisher's Exact test P value (P_{FT}) ≤0.05]. To examine the significance of observing a specific proportion of subsets with correlated SNP effects, we performed 10,000 permutations where for each replicate the SNP effect and P value were randomly shuffled in the ATM migraine GWA results before generating SNP subsets and performing Fisher's Exact tests. Note that only the SNP name-effect relationship in the ATM migraine GWA results is destroyed by permutation to provide a new dataset sampled under the null hypothesis: the effect direction and P value pattern will remain the same under the observed and permuted subsets.

Results

The results of the genetic risk score analyses are summarised in Fig. 1. Although sex was significant in the relationship between genetic risk score and phenotype (due to the unequal sex ratio of migraine and MDD), including sex did not change the conclusions regarding the significance of the genetic risk score. Therefore, results are only provided for analyses excluding sex as a covariate.

Genetic risk scores based on the RADIANT MDD results (Fig. 1a) were significantly predictive of all MDD status in NTR–NESDA (maximum Nagelkerke's pseudo-$R^2 = 0.0080$, $P = 7.5 \times 10^{-5}$, AUC = 0.545 at $P_T = 0.5$). RADIANT MDD-scores also predicted pure MDD ($R^2 = 0.0069$, $P = 0.0019$, AUC = 0.543 at $P_T = 0.3$). The proportion of explained variance was highest for the comorbid MDD + migraine phenotype ($R^2 = 0.0091$, $P = 4 \times 10^{-4}$, AUC = 0.546 at $P_T = 0.6$). Furthermore, RADIANT-MDD was significantly predictive of all migraine in NTR–NESDA ($R^2 = 0.0061$, $P = 0.0015$, AUC = 0.537 at $P_T = 0.3$). However, consistent with hypothesis 1, and inconsistent with hypotheses 2 and 3, the RADIANT-MDD genetic risk score did not significantly predict pure migraine ($R^2 = 0.0037$, $P = 0.058$, AUC = 0.530 at $P_T < 0.05$ and considerably weaker at all other cut-offs).

Similarly, genetic risk scores for migraine based on the ATM data (Fig. 1b) significantly predicted all migraine ($R^2 = 0.0058$, $P = 0.002$, AUC = 0.533 at $P_T = 0.6$), pure migraine ($R^2 = 0.0053$, $P = 0.023$, AUC = 0.532 at $P_T = 0.3$) and comorbid MDD + migraine ($R^2 = 0.0054$, $P = 0.007$, AUC = 0.535 at $P_T = 0.2$) in NTR–NESDA. ATM migraine also predicted all MDD, albeit more weakly ($R^2 = 0.0040$, $P = 0.0052$, AUC = 0.527 at $P_T = 0.3$). ATM migraine was very modestly predictive of pure MDD, marginally significant only at $P_T < 0.3$ ($R^2 = 0.0028$, $P = 0.046$, AUC = 0.520). According to hypothesis 1, this prediction should disappear once MDD cases are removed from the analysis in the discovery sample, and this was indeed observed. After 805 MDD cases were excluded from the ATM sample (Fig. 1c), the genetic risk scores for migraine no longer predicted pure MDD ($R^2 = 0.0016$, $P = 0.14$, AUC = 0.518, at $P_T = 0.05$). To test whether this lack of prediction was due to the reduced discovery GWA sample size, we compared these results with those of 1,000 analyses in which 805 individuals were dropped at random (Fig. 2). The results indicated that although the reduced sample size had a small effect on the amount of explained variance, there was a clear relationship with the MDD status of the individuals removed. In the original analyses (Fig. 1b vs. c), the reduction in explained variance was strongly related to the presence of MDD in the case group: the drop in prediction was considerable for all phenotypes except pure migraine. In the analyses where randomly selected individuals were excluded (Fig. 2), the predictive strength dropped by roughly the same amount for all phenotypes. That is, the drop in prediction of pure MDD in NTR–NESDA using ATM migraine genetic risk scores after excluding MDD individuals was greater than expected by chance.

Next, we tested to what extent our results might be biased by differences in sample size between the different subsets. To investigate this, the prediction analyses were repeated 10,000 times with random samples of 335 cases and 1,155 controls (equal to the size of the smallest target sample, pure migraine), drawn from the total sample available for each of the other phenotypes. The results showed that, while the significance was affected by sample size (reflected by larger P values), the reduced sample size had only a very modest effect on the amount of explained variance (Fig. 3). Importantly, the observed patterns in predictive strength remained highly similar. Hence the observed differences in R^2 between phenotypes were not due to differences in sample size.
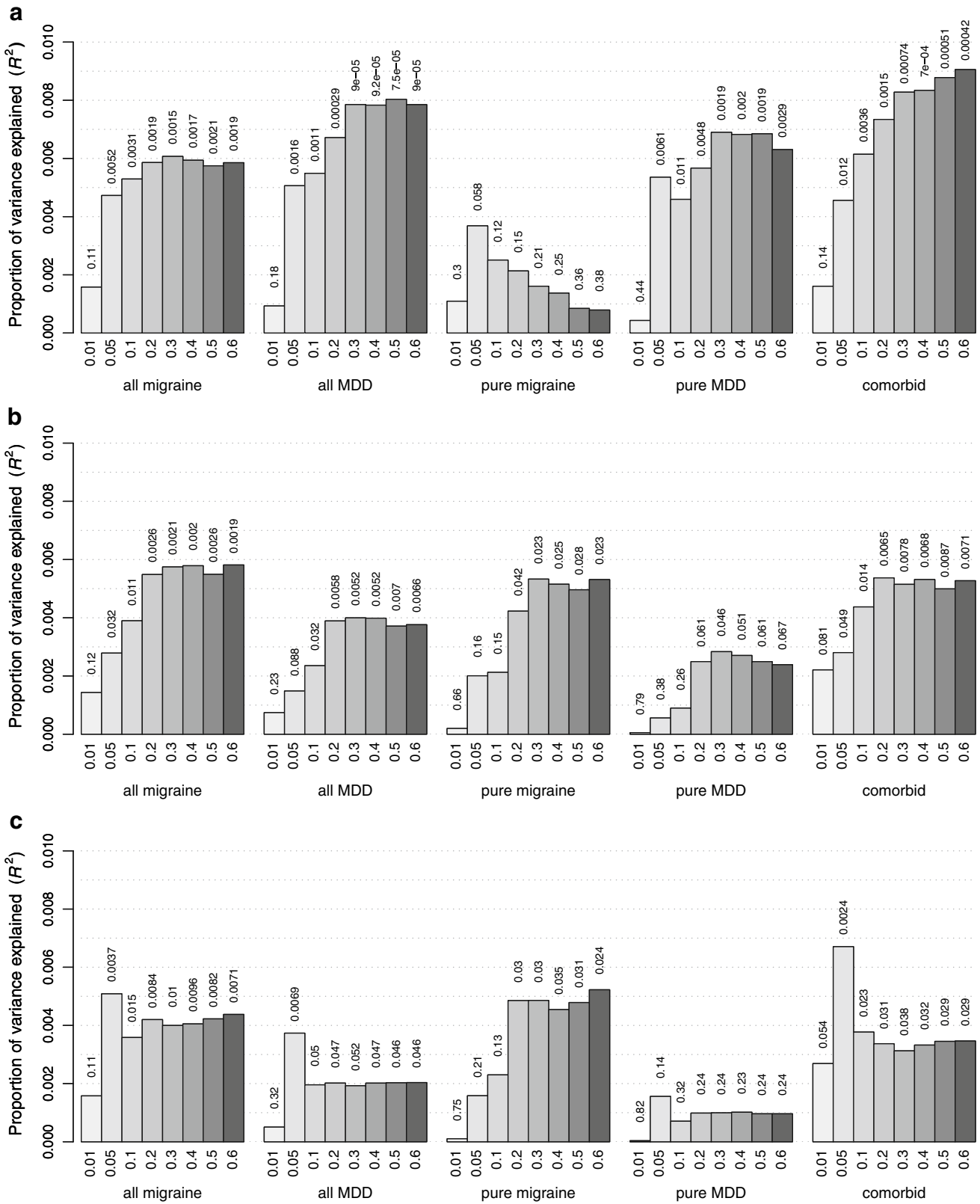
Fig. 1 Prediction of migraine/MDD phenotypes from scores based on RADIANT MDD (**a**), ATM migraine (**b**) and ATM migraine after exclusion of 805 MDD cases (**c**). Results are shown for *P* value thresholds between 0.01 and 0.6 (indicated *below the bars*). The *P* values *above the bars* indicate the significance of each test
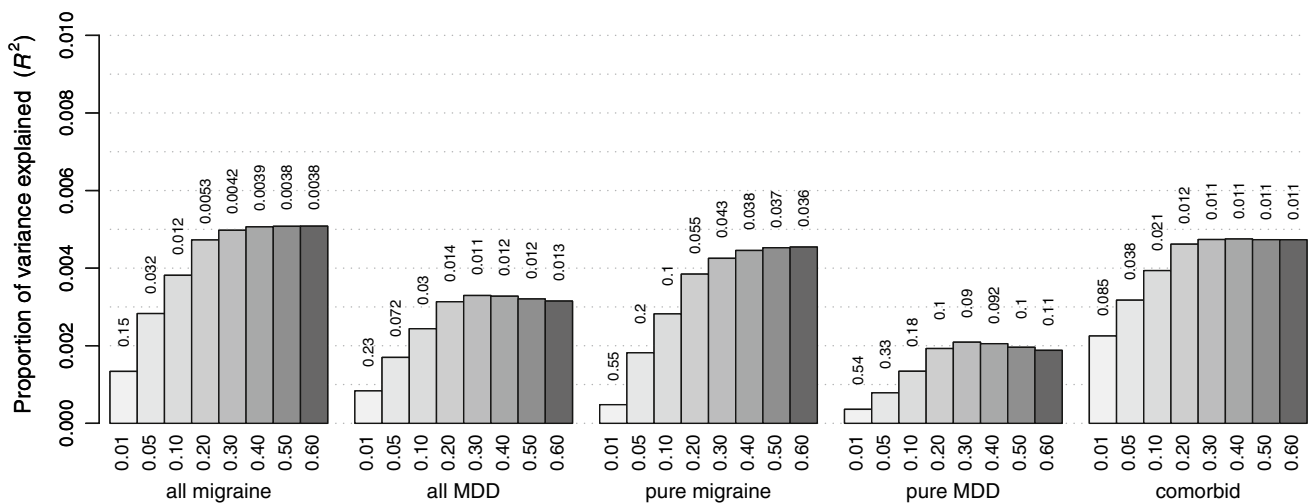
Fig. 2 Averaged results of predictions based on 1,000 permutations in the ATM data, in each of which a random sample of 805 individuals was removed from the dataset. Compared to the original analysis based on all ATM participants (Fig. 1b), this figure shows a small and consistent reduction in R^2 which is independent of the phenotype. However, in the analysis excluding the 805 MDD cases (Fig. 1c), the reduction is very limited for pure migraine, but larger for the other phenotypes (which all include some cases with MDD). This indicates that the different results in ATM before and after removal of the MDD cases are indeed explained by MDD and not merely by differences in sample size

Furthermore, to investigate the impact of the larger number of SNPs in the RADIANT than in the ATM sample, the prediction analyses were repeated using only the SNPs that overlapped between all three samples ($N = 229,778$). This produced only a minor decrease in the explained variance (Fig. 4), and very similar prediction patterns.

Finally, examination of individual SNP effects across the datasets supported the results of our genetic risk scores analyses. For the all migraine and pure MDD GWA results, SNP effects were nominally correlated [Fisher's Exact Test P value (P_{FT}) ≤0.05] in 21 of the total 144 threshold combination subsets. Importantly, for all 21 subsets the SNP effects were in the same direction producing a Fisher's Exact Test Odds Ratio (OR_{FT}) >1. Out of 10,000 permuted replicates, only 360 produced 21 or more subsets with $OR_{FT} > 1$ and $P_{FT} \leq 0.05$ ($P_{FT\text{-permuted}} = 0.036$; 95 %CI: 0.033–0.040). In contrast, for the pure migraine and pure MDD GWA results, SNP effects were nominally correlated ($P_{FT} \leq 0.05$) in only two of the total 144 threshold combination subsets. Importantly, the SNP effects were in the opposite direction in these two subsets ($OR_{FT} < 1$). Out of 10,000 replicates, 5,510 produced zero subsets with $OR_{FT} > 1$ and $P_{FT} \leq 0.05$ ($P_{FT\text{-permuted}} = 0.551$; 95 %CI: 0.541–0.561)—in line with that expected under the null hypothesis of the SNP effects being uncorrelated between datasets.

The substantial and significant reduction in correlation of SNP effects between pure migraine and pure MDD compared to all migraine and pure MDD provides clear support for the observed polygenic overlap between migraine

and depression being due to individuals comorbid for both traits and that the 'pure' forms of migraine and MDD are genetically distinct disorders.

Discussion

Several novel and important findings emerge from this study. First, the results confirm the assumed polygenic nature of both MDD and common migraine and indicate that many SNPs conferring true, albeit weak risk, remain to be identified. Second, this is the first study showing the overlap in genetic influences on migraine and MDD, previously found in twin and family studies, at the measured genotype level. Third, and most importantly, we developed a new application of genetic risk scores to investigate causality, and our results support the hypothesis that in a subset of patients, migraine is a symptom or consequence of MDD. Although MDD genetic risk scores predict migraine and vice versa, pure migraine appears to be genetically distinct from MDD. This means that the observed overlap in the genetic factors influencing migraine and MDD is most likely explained by the subgroup of individuals with comorbid MDD + migraine. If migraine is indeed the result and not the cause of depression, these comorbid cases should be genetically very similar to MDD patients. Figure 1 shows that this is indeed the case: comorbid MDD + migraine was in fact the phenotype most strongly predicted by genetic risk scores for MDD (Fig. 1a), possibly reflecting that these patients have a more severe form
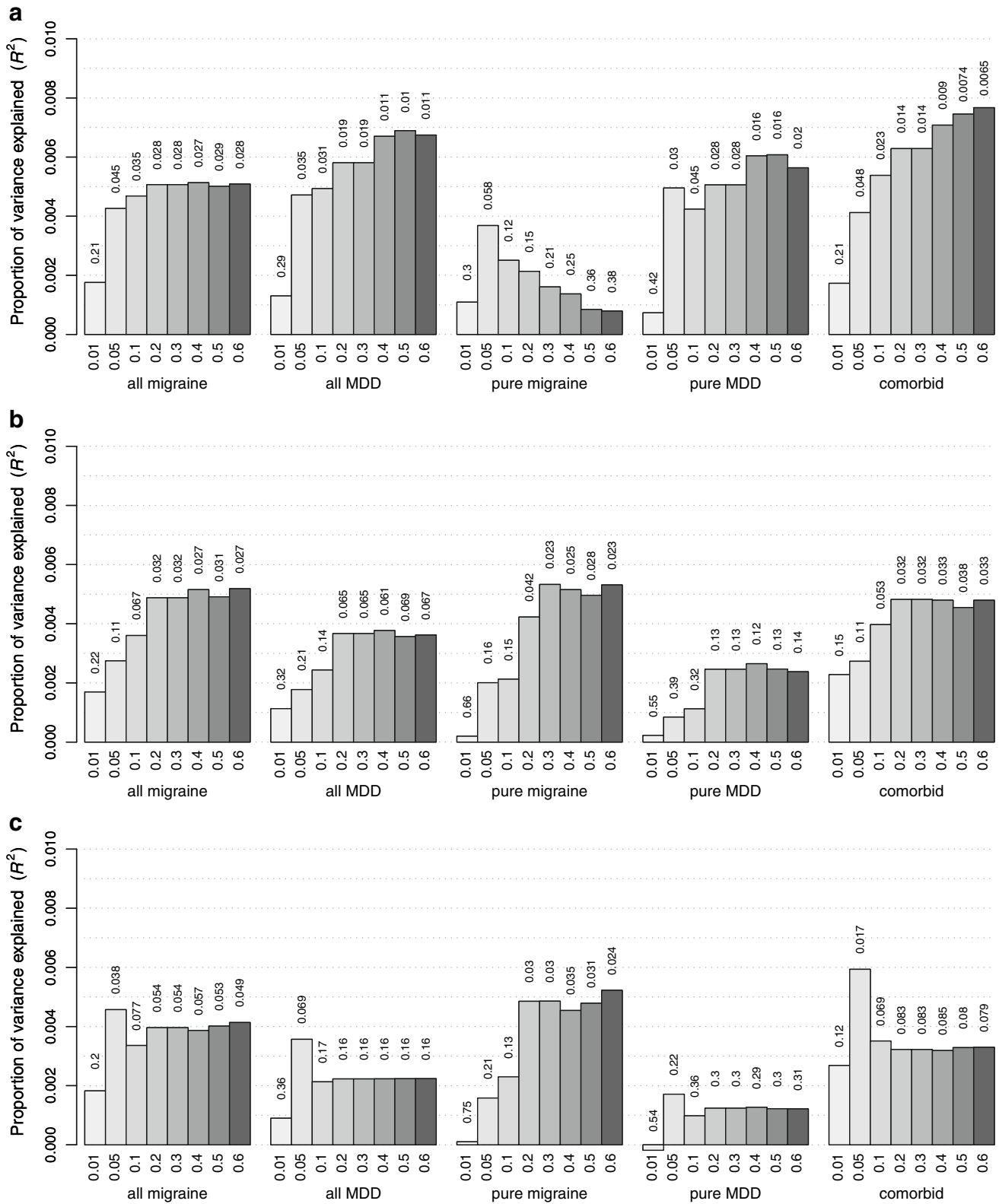
Fig. 3 Averaged results based on 10,000 permutations in which random samples were drawn with sample size equal to the smallest dataset (pure migraine, $N = 335$ cases and 1,155 controls). Results are shown for predictions based on RADIANT (**a**), ATM (**b**), and ATM after exclusion of 805 MDD cases (**c**). Note that for pure migraine, the original results are shown since no permutations were run for the smallest dataset
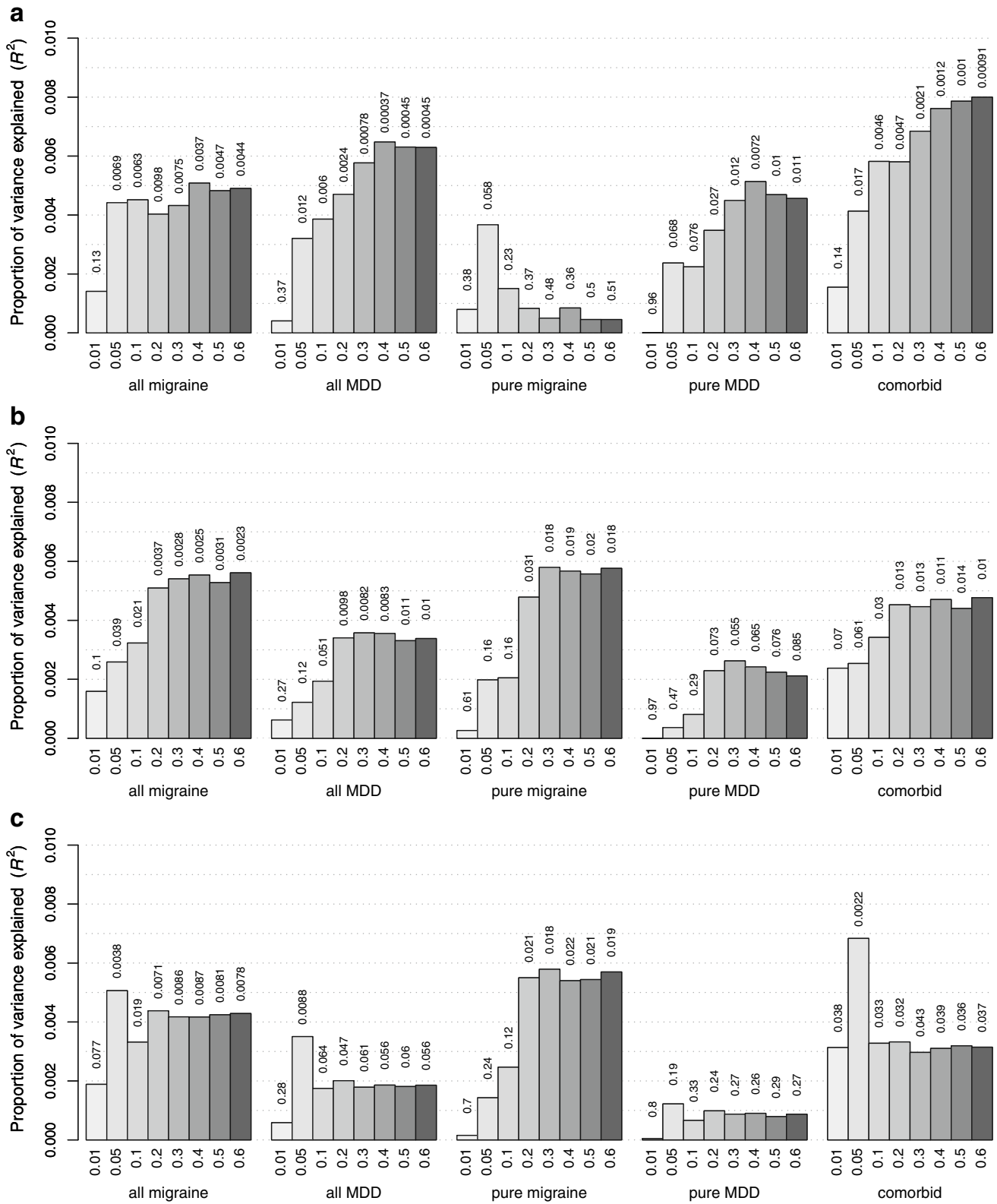
Fig. 4 Results of analyses restricted to SNPs overlapping between all three samples ($N = 229{,}778$), for predictions based on RADIANT (**a**), ATM (**b**) and ATM after exclusion of 805 MDD cases (**c**)

of depression, and are therefore more likely to experience migraine headaches, as would be expected if migraine were a symptom or consequence of MDD. On the other hand, comorbid MDD + migraine was also (more weakly) predicted by all migraine (Fig. 1b), and some marginally significant prediction remained even after exclusion of MDD cases from the migraine discovery sample (Fig. 1c). This could be explained by the presence of individuals with undiagnosed MDD within the ATM sample, or the high prevalence of migraine and MDD, where some individuals are expected to possess risk variants for both migraine and MDD by chance. Therefore, the comorbid group will be a mixture of 'true', 'genetically comorbid' cases, and cases who have both migraine and MDD by chance.

Consistency with previous findings

The finding that migraine may be a symptom or consequence of depression is consistent with our previous findings based on discordant twin analysis, which suggested that the mere existence of genes affecting multiple disorders is insufficient to explain the relationship between migraine and depression, and that a causal or syndromic relationship would be a plausible explanation for the observed comorbidity. Although it might be expected that this type of relationship is associated with a specific order of onset (first depression, then migraine), this is not necessarily the case. If migraine is the 'byproduct' of an underlying depression, the migraine may have its onset either before or after the patient is officially diagnosed with MDD (assuming the first symptoms of depression occur sometime before all diagnostic criteria for MDD are fulfilled). This is entirely consistent with the mixed findings of previous longitudinal studies (Antonaci et al. 2011; Breslau et al. 1991, 2000, 2003; Merikangas et al. 1990; Swartz et al. 2000). Detailed longitudinal studies at the level of symptoms are required to achieve more insight into the developmental trajectories associated with the combination of depressive symptoms and migraine headaches.

Strengths and limitations

A major strength of this study is that the target sample was phenotypically well characterised for both migraine and MDD. Subjects were selected based on MDD status, such that cases met strict clinical diagnostic criteria for MDD and controls were at low risk for MDD. This allowed a strict distinction between migraineurs with and without MDD. In addition, very similar phenotyping procedures were used across target and discovery samples, which ensure that predictions were not limited by differences in phenotype definitions.

A limitation is that the genetic risk scores explained only a small proportion of variance, limiting the predictive ability of the method. This is consistent with findings in previous studies using similar designs and phenotypes (e.g., Demirkan et al. 2010; Middeldorp et al. 2011; Purcell et al. 2009), and expected given the small effects of individual genetic variants on complex traits (Dudbridge 2013). When individual genetic variants have small effects, the effect size estimates will be associated with large standard errors, which limit the strength of the correlation between predicted and observed phenotypes, and therefore the proportion of explained variance (Davies et al. 2011). Moreover, our genetic risk score was not restricted to robustly implicated risk loci; hence the majority of SNP effects in the discovery samples will not be true risk effects. However, the main aim of this study was to analyse patterns in predictive strength and association testing of genetic risk scores, which is possible even when the proportion of explained variance is low, assuming that the variance captured by the measured SNPs is representative of the remaining variance.

Using RADIANT MDD scores to predict migraine and MDD produced more accurate predictions (i.e., explained more variance) than using ATM migraine scores, even though the ATM sample was larger. This might be related to the larger number of SNPs that overlapped between NTR–NESDA and RADIANT. However, restricting the analyses to the SNP set that overlapped between all three samples ($N = 229,778$) produced only a minor decrease in the explained variance (Fig. 4). A more important factor may be that for MDD a more stringent phenotype definition was used than for migraine. The migraine cases were not ascertained based on phenotype status, but drawn from a population-based sample, which includes many genetically informative subclinical migraineurs. A stricter phenotype definition excludes many of these individuals, which effectively results in a loss of statistical power. However, similar patterns were observed under stricter definitions of migraine, for instance, ICHD-II MO/MA cases (data not shown) (Headache Classification Committee of the International Headache Society 2004).

Finally, although migraine was assessed in a large part of the RADIANT sample, for a subset of participants this information was not available. Therefore, we did not perform an analysis in which migraine cases were excluded. However, as expected, in participants with migraine data migraine status was significantly associated with MDD: the prevalence of migraine was 29 % in the MDD cases and 11 % in non-MDD controls (OR = 2.6, 95 % CI 2.0–3.4). In spite of the considerable number of migraineurs and the strong association between migraine and MDD in the RADIANT sample, the MDD genetic risk scores did not predict pure migraine, supporting the hypothesis that migraine

associated with MDD is genetically different from pure migraine. It is expected that this would not change if all migraineurs were excluded from the sample.

Implications

The results of this study imply that some cases of migraine are a result of MDD, and that 'pure' and MDD-related migraine are etiologically different disorders. This has important implications for research and treatment. Firstly, in genetic studies, a distinction should be made between pure migraine and migraine accompanied by depression. The same may apply to other, both psychiatric and non-psychiatric disorders that are comorbid with migraine. Perhaps common migraine might be better viewed as a group of disorders with similar symptomology but (genetically) different etiologies, rather than as a single disorder.

Secondly, patients with MDD-related migraine might require a different course of treatment, or respond differently to certain types of treatment (e.g., with antidepressants) than patients with migraine unrelated to MDD. An important and unanswered question is whether successful treatment of MDD leads to a reduction in migraine symptoms. Well-designed studies targeted specifically at this research question are needed to address this.

References

Antonaci F, Nappi G, Galli F, Manzoni GC, Calabresi P, Costa A (2011) Migraine and psychiatric comorbidity: a review of clinical findings. J Headache Pain 12:115–125

Boomsma DI, de Geus EJ, Vink JM, Stubbe JH, Distel MA, Hottenga JJ, Posthuma D, van Beijsterveldt TC, Hudziak JJ, Bartels M, Willemsen G (2006) Netherlands twin register: from twins to twin families. Twin Res Hum Genet 9:849–857

Boomsma DI, Willemsen G, Sullivan PF, Heutink P, Meijer P, Sondervan D, Kluft C, Smit G, Nolen WA, Zitman FG, Smit JH, Hoogendijk WJ, van Dyck R, de Geus EJ, Penninx BW (2008) Genome-wide association of major depression: description of samples for the GAIN major depressive disorder study: NTR and NESDA biobank projects. Eur J Hum Genet 16:335–342

Breslau N, Davis GC, Andreski P (1991) Migraine, psychiatric disorders, and suicide attempts: an epidemiologic study of young adults. Psychiatry Res 37:11–23

Breslau N, Schultz LR, Stewart WF, Lipton RB, Lucia VC, Welch KM (2000) Headache and major depression: is the association specific to migraine? Neurology 54:308–313

Breslau N, Schultz LR, Welch KMA, Lipton RB, Stewart WF (2003) Comorbidity of migraine and depression: investigating potential etiology and prognosis. Neurology 60:1308–1312

Cohen-Woods S, Gaysina D, Craddock N, Farmer A, Gray J, Gunasinghe C, Hoda F, Jones L, Knight J, Korszun A (2009) Depression case control (DeCC) study fails to support involvement of the muscarinic acetylcholine receptor M2 (CHRM2) gene in recurrent major depressive disorder. Hum Mol Genet 18:1504–1509

Davies G, Tenesa A, Payton A, Yang J, Harris SE, Liewald D, Ke X, Le Hellard S, Christoforou A, Luciano M, McGhee K, Lopez L, Gow AJ, Corley J, Redmond P, Fox HC, Haggarty P, Whalley LJ, McNeill G, Goddard ME, Espeseth T, Lundervold AJ, Reinvang I, Pickles A, Steen VM, Ollier W, Porteous DJ, Horan M, Starr JM, Pendleton N, Visscher PM, Deary IJ (2011) Genome-wide association studies establish that human intelligence is highly heritable and polygenic. Mol Psychiatry 16:996–1005

De Moor MH, Boomsma DI, Stubbe JH, Willemsen G, de Geus EJ (2008) Testing causality in the association between regular exercise and symptoms of anxiety and depression. Arch Gen Psychiatry 65:897–905

Demirkan A, Penninx BWJH, Hek K, Wray NR, Amin N, Aulchenko YS, Van Dyck R, de Geus EJC, Hofman A, Uitterlinden AG (2010) Genetic risk profiles for depression and anxiety in adult and elderly cohorts. Mol psychiatry 16:773–783

Dudbridge F (2013) Power and predictive accuracy of polygenic risk scores. PLoS Genet 9:e1003348

Farmer A, Breen G, Brewster S, Craddock N, Gill M, Korszun A, Maier W, Middleton L, Mors O, Owen M (2004) The depression network (DeNT) study: methodology and sociodemographic characteristics of the first 470 affected sibling pairs from a large multi-site linkage genetic study. BMC Psychiatry 4:42

Hazel LN (1943) The genetic basis for constructing selection indexes. Genetics 28:476

Headache Classification Committee of the International Headache Society (2004) The international classification of headache disorders: 2nd edition. Cephalalgia 24(suppl 1):9–160

Kundu S, Aulchenko YS, van Duijn CM, Janssens ACJ (2011) PredictABEL: an R package for the assessment of risk prediction models. Eur J Epidemiol 26:261–264

Lépine JP, Briley M (2004) The epidemiology of pain in depression. Hum Psychopharmacol 19(suppl 1):S3–S7

Levinson DF (2006) The genetics of depression: a review. Biol Psychiatry 60:84–92

Lewis CM, Ng MY, Butler AW, Cohen-Woods S, Uher R, Pirlo K, Weale ME, Schosser A, Paredes UM, Rivera M (2010) Genomewide association study of major recurrent depression in the UK population. Am J Psychiatry 167:949–957

Ligthart L, Boomsma DI (2012) Causes of comorbidity: pleiotropy or causality? Shared genetic and environmental influences on migraine and neuroticism. Twin Res Hum Genet 15:158–165

Ligthart L, Boomsma DI, Martin NG, Stubbe JH, Nyholt DR (2006) Migraine with aura and migraine without aura are not distinct entities: further evidence from a large Dutch population study. Twin Res Hum Genet 9:54–63

Ligthart L, Nyholt DR, Penninx BW, Boomsma DI (2010a) The shared genetics of migraine and anxious depression. Headache 50:1549–1560

Ligthart L, Penninx BW, Nyholt DR, Distel MA, de Geus EJ, Willemsen G, Smit JH, Boomsma DI (2010b) Migraine symptomatology and major depressive disorder. Cephalalgia 30:1073–1081

Lipton R, Dodick D, Sadovsky R, Kolodner K, Endicott J, Hettiarachchi J, Harrison W (2003) A self-administered screener for migraine in primary care the ID Migraine™ validation study. Neurology 61:375–382

Marchini J, Howie B (2010) Genotype imputation for genome-wide association studies. Nat Rev Genet 11:499–511

Medland SE, Nyholt DR, Painter JN, McEvoy BP, McRae AF, Zhu G, Gordon SD, Ferreira MA, Wright MJ, Henders AK, Campbell MJ, Duffy DL, Hansell NK, Macgregor S, Slutske WS, Heath AC, Montgomery GW, Martin NG (2009) Common variants in the trichohyalin gene are associated with straight hair in Europeans. Am J Hum Genet 85:750–755

Merikangas KR, Angst J, Isler H (1990) Migraine and psychopathology. Results of the Zurich cohort study of young adults. Arch Gen Psychiatry 47:849–853

Middeldorp CM, De Moor MHM, McGrath LM, Gordon SD, Blackwood DH, Costa PT, Terracciano A, Krueger RF, De Geus EJC, Nyholt DR (2011) The genetic association between personality and major depression or bipolar disorder. A polygenic score analysis using genome-wide association data. Transl Psychiatry 1:e50

Mulder EJ, Van Baal C, Gaist D, Kallela M, Kaprio J, Svensson DA, Nyholt DR, Martin NG, MacGregor AJ, Cherkas LF, Boomsma DI, Palotie A (2003) Genetic and environmental influences on migraine: a twin study across six countries. Twin Res 6:422–431

Nyholt DR, Gillespie NG, Heath AC, Merikangas KR, Duffy DL, Martin NG (2004) Latent class and genetic analysis does not support migraine with aura and migraine without aura as separate entities. Genet Epidemiol 26:231–244

Nyholt DR, Morley KI, Ferreira MA, Medland SE, Boomsma DI, Heath AC, Merikangas KR, Montgomery GW, Martin NG (2005) Genomewide significant linkage to migrainous headache on chromosome 5q21. Am J Hum Genet 77:500–512

Penninx BW, Beekman AT, Smit JH, Zitman FG, Nolen WA, Spinhoven P, Cuijpers P, De Jong PJ, Van Marwijk HW, Assendelft WJ, Van Der Meer K, Verhaak P, Wensing M, De Graaf R, Hoogendijk WJ, Ormel J, Van Dyck R (2008) The Netherlands study of depression and anxiety (NESDA): rationale, objectives and methods. Int J Methods Psychiatr Res 17:121–140

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559–575

Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF, Sklar P, Purcell Leader SM, Stone JL, Sullivan PF, Ruderfer DM, McQuillin A, Morris DW, O'Dushlaine CT, Corvin A, Holmans PA, O'Donovan MC, Sklar P, Wray NR, Macgregor S, Sklar P, Sullivan PF, O'Donovan MC, Visscher PM, Gurling H, Blackwood DH, Corvin A, Craddock NJ, Gill M, Hultman CM, Kirov GK, Lichtenstein P, McQuillin A, Muir WJ, O'Donovan MC, Owen MJ, Pato CN, Purcell SM, Scolnick EM, St Clair D, Stone JL, Sullivan PF, Sklar Leader P, O'Donovan MC, Kirov GK, Craddock NJ, Holmans PA, Williams NM, Georgieva L, Nikolov I, Norton N, Williams H, Toncheva D, Milanova V, Owen MJ, Hultman CM, Lichtenstein P, Thelander EF, Sullivan P, Morris DW, O'Dushlaine CT, Kenny E, Quinn EM, Gill M, Corvin A, McQuillin A, Choudhury K, Datta S, Pimm J, Thirumalai S, Puri V, Krasucki R, Lawrence J, Quested D, Bass N, Gurling H, Crombie C, Fraser G, Leh Kuan S, Walker N, St Clair D, Blackwood DH, Muir WJ, McGhee KA, Pickard B, Malloy P, Maclean AW, Van Beck M, Wray NR, Macgregor S, Visscher PM, Pato MT, Medeiros H, Middleton F, Carvalho C, Morley C, Fanous A, Conti D, Knowles JA, Paz Ferreira C et al (2009) Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature 460:748–752

Schur EA, Noonan C, Buchwald D, Goldberg J, Afari N (2009) A twin study of depression and migraine: evidence for a shared genetic vulnerability. Headache 49:1493–1502

Stahl SM (2002) Does depression hurt? J Clin Psychiatry 63:273–274

Stam AH, de Vries B, Janssens AC, Vanmolkot KR, Aulchenko YS, Henneman P, Oostra BA, Frants RR, van den Maagdenberg AM, Ferrari MD, van Duijn CM, Terwindt GM (2010) Shared genetic factors in migraine and depression. Evidence from a genetic isolate. Neurology 74:288–294

Sullivan PF, de Geus EJ, Willemsen G, James MR, Smit JH, Zandbelt T, Arolt V, Baune BT, Blackwood D, Cichon S, Coventry WL, Domschke K, Farmer A, Fava M, Gordon SD, He Q, Heath AC, Heutink P, Holsboer F, Hoogendijk WJ, Hottenga JJ, Hu Y, Kohli M, Lin D, Lucae S, Macintyre DJ, Maier W, McGhee KA, McGuffin P, Montgomery GW, Muir WJ, Nolen WA, Nothen MM, Perlis RH, Pirlo K, Posthuma D, Rietschel M, Rizzu P, Schosser A, Smit AB, Smoller JW, Tzeng JY, van Dyck R, Verhage M, Zitman FG, Martin NG, Wray NR, Boomsma DI, Penninx BW (2008) Genome-wide association for major depressive disorder: a possible role for the presynaptic protein piccolo. Mol Psychiatry 14:359–375

Swartz KL, Pratt LA, Armenian HK, Lee LC, Eaton WW (2000) Mental disorders and the incidence of migraine headaches in a community sample: results from the Baltimore epidemiologic catchment area follow-up study. Arch Gen Psychiatry 57: 945

Uher R, Huezo-Diaz P, Perroud N, Smith R, Rietschel M, Mors O, Hauser J, Maier W, Kozel D, Henigsberg N (2009) Genetic

predictors of response to antidepressants in the GENDEP project. Pharmacogenomics J 9:225–233

Von Korff M, Simon G (1996) The relationship between pain and depression. Br J Psychiatry 168:101–108

Wittchen HU (1994) Reliability and validity studies of the WHO—composite international diagnostic interview (CIDI): a critical review. J Psychiatr Res 28:57–84

Wray N, Pergadia M, Blackwood D, Penninx B, Gordon S, Nyholt D, Ripke S, Macintyre D, McGhee K, Maclean A (2010) Genome-wide association study of major depressive disorder: new results, meta-analysis, and lessons learned. Mol Psychiatry 17:36–48