Title: Polygenic risk for schizophrenia and season of birth within the UK Biobank cohort.

Valentina Escott-Price[1*], Daniel J. Smith[2], Kimberley Kendall[1], Joey Ward[2], George Kirov[1], Michael J. Owen[1], James Walters[1], Michael C. O'Donovan[1]

[1] MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University, Cardiff, UK

[2] Institute of Health and Wellbeing, University of Glasgow, Glasgow, UK

Abstract word count: 155

Total word count: 3166

Number of figures: 3

Number of tables: 2

Number of supplements: 3

Funding: The work at Cardiff University was supported by Medical Research Council (MRC) Centre (MR/L010305/1) and Program Grants (G0800509). The CLOZUK sample was genotyped with funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement n° 279227 (CRESTAR Consortium; http://www.crestar-project.eu/). JW is supported by a JMAS Sim Fellowship from the Royal College of Physicians of Edinburgh and DJS is supported by a Lister Institute Prize Fellowship. KK is supported by a Wellcome Trust Clinical Research Fellowship.

* Corresponding Author
Address: Hadyn Ellis Building, Maindy Road, Cardiff, CF24 4HQ, UK
Phone: +44 (0)29 20688429
E-mail: EscottPriceV@cardiff.ac.uk

Abstract

Background. There is strong evidence that people born in winter and in spring have a small increased risk of schizophrenia. As this 'season of birth' effect underpins some of the most influential hypotheses concerning potentially modifiable risk exposures, it is important to exclude other possible explanations for the phenomenon.

Methods. Here, we sought to determine whether the season of birth effect reflects gene-environment confounding rather than a pathogenic process indexing environmental exposure. We directly measured, in 136,538 participants from the UK Biobank, the burdens of common schizophrenia risk alleles and of CNVs known to increase risk for the disorder, and tested whether these were correlated with season of birth.

Results. Neither genetic measure was associated with season or month of birth within the UK Biobank sample.

Conclusions. As our study was highly powered to detect small effects, we conclude that the season of birth effect in schizophrenia reflects a true pathogenic effect of environmental exposure.

Introduction

People born in winter and early spring are at an elevated risk for schizophrenia (Baron and Gruen 1988, Boyd et al. 1986, Bradbury and Miller 1985, Mortensen et al. 1999). Although the effect is small, with an increase in risk of about 10%, this 'season of birth' effect is one of the most robust findings in schizophrenia epidemiology (Davies et al. 2003). It has also been influential in developing hypotheses of schizophrenia pathogenesis, forming one of the central tenets of the viral infection hypothesis of the disorder (Torrey et al. 1977), as well as other less intensively investigated putative mechanisms such as vitamin D deficiency during foetal development (McGrath 1999). However, at present, the mechanisms underpinning the season of birth effect are not known.

Schizophrenia is also known to be highly heritable (Cardno and Gottesman 2000, Sullivan et al. 2003) and polygenic (International Schizophrenia et al. 2009, Purcell et al. 2014). Genomic studies have identified large numbers of risk alleles that contribute to risk of the disorder, with risk being conferred by large numbers of variants spanning the full spectrum of population frequencies (Rees et al. 2014, Schizophrenia Working Group of the Psychiatric Genomics 2014, Singh et al. 2016). The relative contributions of alleles of various frequencies is not fully resolved, but recent studies estimate that common alleles, captured by genome-wide association study (GWAS) arrays, capture between a third and one half of the genetic variance in liability (Schizophrenia Working Group of the Psychiatric Genomics 2014).

A key assumption underlying research into the causes of the season of birth effect is that season acts as a proxy for one or more environmental exposures (e.g. virus infection). While an environmental origin for the season of birth effect seems the most plausible explanation,

it is also possible that the effect is the result of gene-environment correlation. There are already examples where correlation of genetic and apparent environmental risks have been observed in psychiatric research; one such example, the link between maternal smoking and ADHD, is at least partly driven by shared genetic liability to both ADHD and smoking rather than by exposure to smoking *per se* (Thapar et al. 2009). The link between obstetric complications and schizophrenia may be confounded by correlation between genetic liability to both (Ellman et al. 2007). This is also likely to apply to cannabis and psychosis (Power et al. 2014, Vaucher et al. 2017) and indeed to the link between all substance use disorders and psychosis (Adan et al. 2017).

Gene-environment correlations are sometime classified as passive, active and evocative; regardless of the nature of the correlation, each predicts people born in the winter should have elevated genetic liability to the disorder, even those who do not manifest the disorder. In contrast, if the association between winter birth and schizophrenia is *not* the result of gene-environment correlation, there should be no link between winter birth and genetic liability to the disorder. Evocative correlation is said to occur when an individual's behaviour evokes an environmental response and is therefore not clearly applicable to the phenomenon under investigation. Passive gene-environment correlation could, in principle, operate if parents with elevated trait liability have seasonal patterns of sexual activity that favour winter birth. If this is the case, then their offspring are expected to have elevated liability to the disorder (on average the mean liability of the parents which will be higher than the population as a whole) and liability to winter birth. While this seems *a priori* unlikely, previous studies have shown that high schizophrenia trait liability is associated with a wide range of behaviours, amongst which of potential relevance are seasonal fluctuations in mood and activity levels

(Byrne et al. 2015). Confounding through active gene-environment correlation is also possible if enhanced genetic risk for the disorder impacts upon seasonal patterns of foetal loss. Given the potential impacts on prevention and treatment of detecting modifiable environmental exposures, it is important to rigorously exclude alternative explanations for the season of birth effect on schizophrenia risk.

Population studies using family history have generally suggested that the season of birth effect is not a manifestation of genetic liability (Hettema et al. 1996, Suvisaari et al. 2004). However, the majority of people who develop schizophrenia do not have a history of the disorder in a close relative (Svensson et al. 2012). Advances in molecular genetics now allow genetic liability to be directly estimated in individuals regardless of their affected status or family history. Liability conferred by common risk alleles can be estimated through a process known as polygenic risk scoring; in a given individual, their polygenic risk score (PRS) represents the burden of common risk alleles carried by that individual. PRS studies have repeatedly been demonstrated to provide a useful index of genetic liability to the disorder (International Schizophrenia et al. 2009, Ripke et al. 2013, Schizophrenia Working Group of the Psychiatric Genomics 2014), and are increasingly being applied to investigate whether environmental risk factors operate independently of genetic liability (Power et al. 2014).

Here, we implement PRS analysis of the UK Biobank (UKBB) (Sudlow et al. 2015) to determine whether season and month of birth are associated with genetic risk for schizophrenia. The UK Biobank is a large prospective cohort of more than half a million residents of the UK for which genetic data and seasonality of birth data are available. We generated schizophrenia PRS for every participant in the UK Biobank cohort (June 2015 release, N=136,538) and tested

whether this score was associated with month or season of birth. As a secondary test of the plausibility of season of birth being a genetically correlated confound, we also conducted a GWAS of season of birth, aiming to estimate whether this phenotype is heritable.

Rare alleles in the form of pathogenic copy number variants (CNVs) are also known make a contribution to schizophrenia. Although the contribution to liability from CNVs at a population level (Purcell et al. 2014) is much smaller than that of common alleles, for completeness we also tested whether the frequencies of 93 pathogenic CNVs that have been linked with neurodevelopmental disorders (Coe et al. 2014, Dittwald et al. 2013, Kendall et al. 2016) were associated with season of birth.

UK Biobank obtained informed consent from all participants and this study was conducted under generic approval from the NHS National Research Ethics Service (approval letter dated 13 May 2016, Ref 16/NW/0274) and under UK Biobank approvals for applications #6553 (Smith) and #14421 (Kirov).

Material and Methods

To define risk alleles, we used the largest available schizophrenia GWAS comprising a meta-analysis of two large studies (Pardinas 2017, Schizophrenia Working Group of the Psychiatric Genomics 2014) which included 40,675 schizophrenia cases and 64,643 controls (Pardinas 2017). As PRS requires genome wide data, we did not include the extension dataset provided to the PGC for selected SNPs by deCODE genetics. Both schizophrenia datasets were imputed using the SHAPEIT/IMPUTE2 software (Delaneau et al. 2013, Howie et al. 2012) with a

combination of the 1000 Genomes phase 3 (1KGPp3) and UK10K datasets as a reference panel.

For the UK Biobank study, given that the schizophrenia GWAS we used to define risk alleles was of primarily European Ancestry, we restricted the sample to those who self-reported as being of white UK ancestry (*n*=136,538). For constructing PRS, we downloaded data that is publically available (https://www.med.unc.edu/pgc/results-and-downloads). We included autosomal SNPs that passed stringent quality control criteria (minor allele frequencies (MAF) ≥0.01) and imputation quality score greater than or equal to 0.9. This resulted in 5,471,613 SNPs. Using the UK Biobank genotypes, we pruned the SNPs keeping those which are the most significantly associated with schizophrenia in the region while excluding SNPs at which the genotypes are correlated with the selected SNPs with $r^2 \geq 0.2$. A physical distance threshold for pruning SNPs was set to 1Mb and p-value threshold was 0.5. After pruning, 118,302 independent SNPs remained. We selected markers, based upon significance thresholds, to construct a polygenic score in the UK Biobank data. The PRS was calculated from the effect size-weighted sum of associated alleles within each subject. PRS were standardised by subtracting the population mean for PRS and dividing by the standard deviation.

As a test of robustness, we constructed PRS based on risk alleles passing a range of schizophrenia association thresholds in the PGC2+CLOZUK data (e.g. significant at p≤0.01, 0.05, 0.1, …, 0.5). The primary analysis was based on P<0.05 as this is the threshold that currently maximally captures polygenic risk (Schizophrenia Working Group of the Psychiatric Genomics 2014). The polygenic risk scores of people born in each month/season were compared to those born in January/winter (baseline) using linear regression analysis with

season/month coded as a factor (glm() function in R). All analyses in the UKBB were adjusted for the array (the UK Biobank used two different arrays) and the first 8 principal components (PCs), reflecting underlying stratification in the sample due to population and/or genotyping differences. The first 8 PCs, out of 15 available in the Biobank, were selected after visual inspection of each pair of PCs, taking forward only those that resulted in multiple clusters of individuals (see (Smith et al. 2016) for detail).

AVENGEME provides a set of R functions that allow power for PRS analyses to be calculated (Dudbridge 2013). For this to be applicable, *a priori*, this method assumes that 1) there is a non-zero SNP-heritability for the season of birth and 2) there is a genetic correlation between SZ and the season of birth. As we show in the present study that both assumptions are violated, this widely used approach is not applicable. However, to illustrate that we had high power to detect the sought after effects, we calculated the power of our PRS analysis for the reported effect size OR=1.07 (Davies et al. 2003) comparing winter/spring versus summer/autumn births with pwr.norm.test() function in R.

Genome wide association analysis (GWA) of season of birth was performed using a binary phenotype defined as "0" for individuals born in winter and spring (December-May) and "1" for those born in summer and autumn (June-November). Association analysis was conducted using logistic regression with array, and the first 8 principal components as covariates (as described above). Genotype dosage data were initially converted to the most probable genotype format (with the probability>0.9), then filtered by removal of SNPs with Hardy-Weinberg equilibrium $p<10^{-6}$, MAF<0.01, info<0.4, data on <95% of the sample after excluding genotype calls made with less than 90% posterior probability after which 8,989,945 variants

were retained. LD-Score regression analysis (Bulik-Sullivan et al. 2015) was employed to estimate SNP-based heritability in this dataset.

The numbers of subjects with and without pathogenic CNVs born in winter/spring vs summer/autumn were compared with a chi-squared test. The list of 93 pathogenic CNVs was compiled from two widely accepted sources (Coe et al. 2014, Dittwald et al. 2013) as we have previously reported (Kendall et al. 2016). The CNV calls for UK Biobank participants were made in-house as we have previously reported in a UK Biobank CNV study (Kendall et al. 2016).

Results

The frequencies of birth by month are presented in Figure 1. The distribution of frequencies in our study is similar to the distribution in the whole UK Biobank population (N=502,632; biobank.ctsu.ox.ac.uk/crystal/field.cgi?id=52).

Correlations between timing of birth and polygenic risk.

We found no association between schizophrenia PRS and season of birth in the UK Biobank sample, with winter as the baseline. Figure 2 shows the distribution of the standardised PRS for each season. The figure clearly demonstrates no difference in mean or variance of PRS. Detailed results are presented in Table 1, with baseline winter as season of birth. Similarly, we found no differences in the PRS by month of birth with January as the baseline (Figure 3; Supplementary Table 1). Note that the specification of the baseline season is arbitrary; the results remain the same regardless which season is used.

Given that previous studies have frequently identified both winter and spring as the season of elevated risk, we also collapsed those born in winter and spring into a single group and compared their PRS with those born in summer or autumn. Our findings were not consistent with a season of birth-PRS correlation; indeed those born in winter and spring actually had a slightly lower PRS than those born in summer or autumn, and this reached nominal significance for some of the secondary tests (Table 2).

To check whether this sample maybe underpowered to detect low-effect sizes, we calculated the power of the PRS analysis for the reported effect size OR=1.07 (Davies et al. 2003). The effect size for power calculation is usually estimated as $(B_0\text{-}B)/\sigma$, where B_0 is the effect size under the null hypothesis (in our case $B_0=\log(1)=0$), B is the effect size under the alternative hypothesis (in our case $B=\log(1.07)=0.068$) and σ is the standard deviation. Since the PRS were standardised in our analyses, $\sigma=1$ under the both null and alternative hypotheses. Thus, the effect size for the power calculation was d=0.068/sqrt(2)=0.048 and the significance level was set to $\alpha=10^{-4}$, to account for multiple testing. The power to detect association between PRS and season of birth was >99.9%. We also estimated d=0.0128 as the smallest effect size which our sample is capable to detect with 80% power, which corresponds to OR=1.018 (B=0.018) excess for winter/spring births compared to summer/autumn births. As the power to detect association between PRS and season of birth is 80%, assuming association effect size as small as B=0.018 at 10^{-4} significance level (accounting for all PRS tests), we exclude seasonal variation in PRS as a being correlated with season of birth effect.

To evaluate effects that might only be observed at the extremes of liability, we also compared "winter and spring" vs "summer and autumn" births for the top and bottom deciles of the

PRS distribution; again we found no evidence for association at the primary testing threshold for risk alleles from the PGC study. We did observe a nominally significant association (p=0.02) when risk alleles were selected from the PGC GWAS at p<0.1, but again, the effect size was again inconsistent with that expected for a genotype phenotype correlation, people in the top 10% for genetic loading to schizophrenia having a slightly higher chance of being born in summer or autumn compared with the bottom 10% of people (Table 2).

Genome wide association study of timing of birth.

A Manhattan plot of the GWAS of season of birth is given as Supplementary Figure 1. No genome-wide significant associations were identified, there was no evidence for inflation in the test statistics indicative of polygenic inheritance (genomic control (Devlin and Roeder 1999) λ=0.992 (see QQ-plot in Supplementary Figure 2), and no evidence that SNP heritability contributed to this phenotype (total liability scale h^2=-0.002 SE=0.0052) as estimated by LD-Score regression.

CNVs and timing of birth.

There was no difference in the frequencies of pathogenic CNVs between groups of UK Biobank participants born in winter and spring compared with those born in summer or autumn (frequencies 0.017 [95%CI=0.016-0.018] and 0.016 [95%CI=0.015-0.017], respectively; chi-squared test p=0.137).

Conclusion

An excess of winter and spring births in people with schizophrenia is one of the most robustly supported and influential epidemiological findings in psychiatry [1-5]. Here, we test and fail to support the hypothesis that the excess of winter and spring births in people with schizophrenia is an effect of gene-environment correlation. Not only do we fail to find evidence that schizophrenia liability in the form of common alleles or rare CNVs is associated with season of birth, our GWAS suggests that season of birth is not even a heritable trait, which alone makes such a correlation an untenable explanation for the season of birth effect (with respect to common variation). Rejecting a genetic-environmental correlation, we conclude that our study strongly supports the widely held view that the excess of winter births in schizophrenia is the result of an as yet unknown environmental risk exposure with a seasonal gradient.

Although not a direct aim of our study, we also note that the absence of heritability effects on season of birth implies that other traits that exhibit relative age effects, for example personality traits, sporting ability, and general academic performance (Jeronimus et al. 2015), are similarly unlikely to represent gene environment correlation; rather, as widely interpreted, they most likely result from differential levels of maturity in school and other cohort intakes. However, some cognitive phenotypes exhibit season of birth fluctuations beyond effects attributable simply to the timing of school intake (Grootendorst-van Mil et al. 2017). It is therefore not possible to exclude the possibility that an as yet to be discovered seasonally-fluctuating risk factor that increases risk of schizophrenia also contributes to season of birth effects which have been observed for these other cognitive phenotypes.

Our study has a number of major strengths: the use of the largest schizophrenia dataset to date to identify genetic risk loci; the largest available genotyped population cohorts to generate schizophrenia PRS; and the ability to measure liability to schizophrenia directly at a molecular level. Another strength conferred by the large population sample is the power to test the sample month by month. This is important given that many studies vary in their definition of winter-spring and in the month for which risk is maximal. Together, these strengths allow us to test the genetic confounding hypothesis with extremely high power; as a result, our failure to find evidence in support of that hypothesis allows us to refute it as an explanation. In doing so, our results are consistent with, and complementary to, studies that have indirectly measured genetic liability based on family history (Hettema et al. 1996, Suvisaari et al. 2004, Svensson et al. 2012).

Our study has a number of limitations. One potential limitation of the study is that, like all environmental exposures, possible variance in the exposure rate to the pathogenic agent might mean our conclusion could be country or birth cohort specific. However, season of birth effects (and therefore exposure to the putative pathogenic environmental exposures) have been widely documented in Northern European samples [5]. They have also been shown to operate in the UK from at least 1921 up till the modern era, with the most recent study in the UK suggesting January births are associated with an OR for schizophrenia of 1.17 (Disanto et al. 2012). Another potential limitation is that our analyses did not account for the possibility that an individual's circadian biology or chronotype might interact with season of birth effects (Natale and Adan 1999, Natale et al. 2009).

Finally, our PRS analysis, and heritability estimates, were based upon common SNPs, and do not include a possible contribution from rare SNPs. However, the frequencies of rare CNVs,

linked to neurodevelopmental disorders, also did not differ by season of birth, and it seems unlikely that rare mutations that increase liability to schizophrenia would have different effects on mating behaviours than burden of common alleles. Nevertheless, when sufficient data become available, it may useful to test for seasonal burdens of rare and *de novo* mutation events.

Acknowledgements

This research was conducted using the UK Biobank resource. UK Biobank was established by the Wellcome Trust, Medical Research Council, Department of Health, Scottish Government and Northwest Regional Development Agency. UK Biobank has also had funding from the Welsh Assembly Government and the British Heart Foundation. Data collection was funded by UK Biobank. The work at Cardiff University was supported by Medical Research Council (MRC) Centre (MR/L010305/1) and Program Grants (G0800509). The CLOZUK sample was genotyped with funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement n° 279227 (CRESTAR Consortium; http://www.crestar-project.eu/). JW is supported by a JMAS Sim Fellowship from the Royal College of Physicians of Edinburgh and DJS is supported by a Lister Institute Prize Fellowship. KK is supported by a Wellcome Trust Clinical Research Fellowship.

References

Adan A, Arredondo AY, Capella MD, Prat G, Forero DA, Navarro JF (2017) Neurobiological underpinnings and modulating factors in schizophrenia spectrum disorders with a comorbid substance use disorder: A systematic review, *Neuroscience & Biobehavioral Reviews,* 75, 361-377.

Baron M, Gruen R (1988) Risk factors in schizophrenia. Season of birth and family history, *British Journal of Psychiatry,* 152, 460-465.

Boyd JH, Pulver AE, Stewart W (1986). Season of birth: schizophrenia and bipolar disorder, *Schizophrenia Bulletin,* 12(2), 173-186.

Bradbury TN, Miller GA (1985.) Season of birth in schizophrenia: a review of evidence, methodology, and etiology, *Psychological Bulletin,* 98(3), 569-594.

Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the PGC, Patterson N, Daly MJ, Price AL, Neale BM (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies, *Nature Genetics,* 47(3), 291-295.

Byrne EM, Psychiatric Genetics Consortium Major Depressive Disorder Working Group, Raheja UK, Stephens SH, Heath AC, Madden PA, Vaswani D, Nijjar GV, Ryan KA, Youssufi H, Gehrman, P. R., Shuldiner AR, Martin NG, Montgomery GW, Wray NR, Nelson EC, Mitchell BD, Postolache TT (2015). Seasonality shows evidence for polygenic architecture and genetic correlation with schizophrenia and bipolar disorder, *Journal of Clinical Psychiatry,* 76(2), 128-134.

Cardno AG, Gottesman II (2000). Twin studies of schizophrenia: from bow-and-arrow concordances to star wars Mx and functional genomics, *American Journal of Medical Genetics,* 97(1), 12-17.

Coe BP, Witherspoon K, Rosenfeld JA, van Bon BW, Vulto-van Silfhout AT, Bosco P, Friend KL, Baker C, Buono S, Vissers LE, Schuurs-Hoeijmakers JH, Hoischen A, Pfundt R, Krumm N, Carvill GL, Li D, Amaral D, Brown N, Lockhart PJ, Scheffer IE, Alberti A, Shaw M, Pettinato R, Tervo R, de Leeuw, N., Reijnders MR, Torchia BS, Peeters H, O'Roak BJ, Fichera M, Hehir-Kwa JY, Shendure J, Mefford HC, Haan E, Gecz J, de Vries, BB, Romano C, Eichler EE (2014). Refining analyses of copy number variation identifies specific genes associated with developmental delay, *Nature Genetics,* 46(10), 1063-1071.

Davies G, Welham J, Chant D, Torrey EF, McGrath J (2003). A systematic review and meta-analysis of Northern Hemisphere season of birth studies in schizophrenia, *Schizophrenia Bulletin,* 29(3), 587-593.

Delaneau O, Zagury JF, Marchini J (2013). Improved whole-chromosome phasing for disease and population genetic studies, *Nature Methods,* 10(1), 5-6.

Devlin B, Roeder K (1999). Genomic control for association studies, *Biometrics,* 55(4), 997-1004.

Disanto G, Morahan JM, Lacey MV, DeLuca GC, Giovannoni G, Ebers GC, Ramagopalan SV (2012). Seasonal distribution of psychiatric births in England, *PLoS One,* 7(4), e34866.

Dittwald P, Gambin T, Szafranski P, Li J, Amato S, Divon MY, Rodriguez Rojas LX, Elton LE, Scott DA, Schaaf CP, Torres-Martinez W, Stevens AK, Rosenfeld JA, Agadi S, Francis D, Kang SH, Breman A, Lalani SR, Bacino CA, Bi W, Milosavljevic A, Beaudet AL, Patel A, Shaw CA, Lupski JR, Gambin A, Cheung SW, Stankiewicz P (2013). NAHR-mediated copy-number variants in a clinical population: mechanistic insights into both genomic disorders and Mendelizing traits, *Genome Research,* 23(9), 1395-1409.

Dudbridge F (2013). Power and predictive accuracy of polygenic risk scores, *PLoS Genetics,* 9(3), e1003348

Ellman LM, Huttunen M, Lonnqvist J, Cannon TD (2007). The effects of genetic liability for schizophrenia and maternal smoking during pregnancy on obstetric complications, *Schizophrenia Research,* 93(1-3), 229-236.

Grootendorst-van Mil NH, Steegers-Theunissen RPM, Hofman A, Jaddoe VWV, Verhulst FC, Tiemeier H (2017). Brighter children? The association between seasonality of birth and child IQ in a population-based birth cohort, *BMJ Open,* 7(2)

Hettema, J M, Walsh, D and Kendler, K S (1996) 'Testing the effect of season of birth on familial risk for schizophrenia and related disorders', *British Journal of Psychiatry,* 168(2), 205-209.

Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR (2012). Fast and accurate genotype imputation in genome-wide association studies through pre-phasing, *Nature Genetics,* 44(8), 955-959.

International Schizophrenia Consortium, Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF, Sklar P (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder, *Nature,* 460(7256), 748-752.

Jeronimus BF, Stavrakakis N, Veenstra R, Oldehinkel AJ (2015). Relative Age Effects in Dutch Adolescents: Concurrent and Prospective Analyses, *PLoS One,* 10(6), e0128856

Kendall KM, Rees E, Escott-Price V, Einon M, Thomas R, Hewitt J, O'Donovan MC, Owen MJ, Walters JT, Kirov G (2017). Cognitive Performance Among Carriers of Pathogenic Copy Number Variants: Analysis of 152,000 UK Biobank Subjects, *Biological Psychiatry,* 82(2), 103-110.

McGrath J (1999). Hypothesis: is low prenatal vitamin D a risk-modifying factor for schizophrenia?', *Schizophrenia Research,* 40(3), 173-177.

Mortensen PB, Pedersen CB, Westergaard T, Wohlfahrt J, Ewald H, Mors O, Andersen PK, Melbye M (1999). Effects of family history and place and season of birth on the risk of schizophrenia, *New England Journal of Medicine,* 340(8), 603-608.

Natale V, Adan A (1999). Season of birth modulates morningness-eveningness preference in humans', *Neuroscience Letter,* 274(2), 139-141.

Natale V, Adan A, Fabbri M (2009). 'Season of birth, gender, and social-cultural effects on sleep timing preferences in humans', *Sleep,* 32(3), 423-426.

Pardinas AF, Holmans P, Pocklington AJ, Escott-Price V, et al (2018). 'Common schizophrenia alleles are enriched in mutation-intolerant genes and maintained by background selection', *Nature Genetics,* (in press)

Power RA, Verweij KJ, Zuhair M, Montgomery GW, Henders AK, Heath AC, Madden PA, Medland SE, Wray NR, Martin NG (2014). Genetic predisposition to schizophrenia associated with increased use of cannabis, *Molecular Psychiatry,* 19(11), 1201-1204.

Purcell SM, Moran JL, Fromer M, Ruderfer D, Solovieff N, Roussos P, O'Dushlaine C, Chambert K, Bergen SE, Kahler A, Duncan L, Stahl E, Genovese G, Fernandez E, Collins MO, Komiyama NH, Choudhary JS, Magnusson PK, Banks E, Shakir K, Garimella K, Fennell T, DePristo M, Grant SG, Haggarty SJ, Gabriel S, Scolnick EM, Lander ES, Hultman CM, Sullivan PF, McCarroll SA, Sklar P (2014). A polygenic burden of rare disruptive mutations in schizophrenia, *Nature,* 506(7487), 185-190.

Rees E, Walters JT, Georgieva L, Isles AR, Chambert KD, Richards AL, Mahoney-Davies G, Legge SE, Moran JL, McCarroll SA, O'Donovan MC, Owen MJ, Kirov G (2014). Analysis of copy number variations at 15 schizophrenia-associated loci, *British Journal of Psychiatry,* 204(2), 108-114.

Ripke S, O'Dushlaine C, Chambert K, Moran JL, Kahler AK, Akterin S, Bergen SE, Collins AL, Crowley JJ, Fromer M, Kim Y, Lee SH, Magnusson PK, Sanchez N, Stahl EA, Williams S, Wray NR, Xia K, Bettella F, Borglum AD, Bulik-Sullivan BK, Cormican P, Craddock N, de Leeuw C, Durmishi N, Gill M, Golimbet V, Hamshere ML, Holmans P, Hougaard DM, Kendler KS, Lin K, Morris DW, Mors O, Mortensen PB, Neale BM, O'Neill FA, Owen MJ, Milovancevic MP, Posthuma D, Powell J, Richards AL, Riley B P, Ruderfer D, Rujescu D, Sigurdsson E, Silagadze T, Smit AB, Stefansson H, Steinberg S, Suvisaari J, Tosato, S, Verhage, M, Walters, J T, Multicenter Genetic Studies of Schizophrenia Consortium, Levinson DF, Gejman PV, Kendler KS, Laurent C, Mowry BJ, O'Donovan MC, Owen MJ, Pulver, A E, Riley BP, Schwab SG, Wildenauer DB, Dudbridge F, Shi J, Albus M, Alexander M, Campion D, Cohen D, Dikeos D, Duan J, Eichhammer P, Godard S, Hansen M, Lerer FB, Liang KY, Maier W, Mallet J, Nertney DA, Nestadt G, Norton N, O'Neill FA, Papadimitriou GN, Ribble R, Sanders AR, Silverman JM, Walsh D, Williams NM, Wormley B, Psychosis Endophenotypes International Consortium, Arranz MJ, Bakker S, Bender S, Bramon E, Collier D, Crespo-Facorro B, et al (2013). Genome-wide association analysis identifies 13 new risk loci for schizophrenia, *Nature Genetics,* 45(10), 1150-1159.

Schizophrenia Working Group of the Psychiatric Genomics Consortium (2014). Biological insights from 108 schizophrenia-associated genetic loci, *Nature,* 511(7510), 421-427.

Singh T, Kurki MI, Curtis D, Purcell SM, Crooks L, McRae J, Suvisaari J, Chheda H, Blackwood D, Breen G, Pietilainen O, Gerety SS, Ayub M, Blyth M, Cole T, Collier D, Coomber EL, Craddock N, Daly M J, Danesh J, DiForti M, Foster A, Freimer NB, Geschwind D, Johnstone M, Joss S, Kirov G, Korkko J, Kuismin O, Holmans P, Hultman CM, Iyegbe C, Lonnqvist J, Mannikko M, McCarroll SA, McGuffin P, McIntosh AM, McQuillin A, Moilanen JS, Moore C, Murray RM, Newbury-Ecob R, Ouwehand W, Paunio T, Prigmore E, Rees E, Roberts D, Sambrook J, Sklar P, StClair D, Veijola J, Walters JT, Williams H, Swedish Schizophrenia Study I, DDD, Consortium UK, Sullivan PF, Hurles ME, O'Donovan MC, Palotie A Owen, MJ, Barrett JC (2016) Rare loss-of-function variants in SETD1A are associated with schizophrenia and developmental disorders, *Nature Neuroscience,* 19(4), 571-577.

Smith DJ, Escott-Price V, Davies G, Bailey ME, Colodro-Conde L, Ward J, Vedernikov A, Marioni R, Cullen B, Lyall D, Hagenaars SP, Liewald DC, Luciano M, Gale CR, Ritchie SJ, Hayward C, Nicholl B, Bulik-Sullivan B, Adams, M, Couvy-Duchesn B, Graham N, Macka D, Evan J, Smith BH, Porteous DJ, Medland SE, Martin NG, Holmans P, McIntosh AM, Pell JP, Deary IJ, O'Donovan MC (2016). Genome-wide analysis of over 106 000 individuals identifies 9 neuroticism-associated loci, *Molecular Psychiatry,* 21(6), 749-757.

Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, Liu B, Matthews P, Ong G, Pell J, Silman A, Young A, Sprosen T, Peakman T, Collins R (2015). UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age, *PLoS Med,* 12(3), e1001779

Sullivan PF, Kendler KS, Neale MC (2003). Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies, *Archives of General Psychiatry,* 60(12), 1187-1192.

Suvisaari JM, Haukka JK, Lonnqvist JK (2004) No association between season of birth of patients with schizophrenia and risk of schizophrenia among their siblings, *Schizophrenia Research,* 66(1), 1-6

Svensson AC, Lichtenstein P, Sandin S, Oberg S, Sullivan PF, Hultman CM (2012). Familial aggregation of schizophrenia: the moderating effect of age at onset, parental immigration, paternal age and season of birth, *Scandinavial Journal Public Health,* 40(1), 43-50.

Thapar A, Rice F, Hay D, Boivin J, Langley K, van den Bree M, Rutter M and Harold G (2009). Prenatal smoking might not cause attention-deficit/hyperactivity disorder: evidence from a novel design, *Biological Psychiatry,* 66(8), 722-727.

Torrey EF, Torrey BB, Peterson MR (1977). Seasonality of schizophrenic births in the United States, *Archives General Psychiatry,* 34(9), 1065-1070.

Vaucher J, Keating BJ, Lasserr, AM, Gan W, Lyall DM, Ward J, Smith DJ, Pell JP, Sattar N, Pare G, Holmes MV (2017). Cannabis use and risk of schizophrenia: a Mendelian randomization study, *Molecular Psychiatry* doi:10.1038/mp.2016.252

Table 1. Comparison of schizophrenia PRS of individuals in the UK Biobank sample, split by season of births. The schizophrenia PRS are generated using schizophrenia risk SNPs at different thresholds for association. The baseline category of the analysis is winter birth and therefore a negative B indicates decrease in risk of schizophrenia in those born in the season, which is shown in the header row, compared to winter.

SNP selection threshold	NSNPs	Spring		Summer		Autumn	
		B	p-value	B	p-value	B	p-value
10^{-4}	1749	-0.014	0.070	0.004	0.648	-0.001	0.899
0.001	4517	-0.009	0.222	0.008	0.303	0.004	0.607
0.01	13700	-0.005	0.506	0.010	0.199	0.007	0.385
0.05	32576	-0.005	0.538	0.003	0.737	0.003	0.746
0.1	48188	-0.010	0.190	-0.002	0.766	0.000	0.951
0.2	72075	-0.012	0.109	-0.004	0.559	0.001	0.893
0.3	90443	-0.014	0.070	-0.008	0.313	-0.001	0.883
0.4	105255	-0.014	0.059	-0.008	0.284	-0.003	0.731
0.5	117618	-0.014	0.070	-0.008	0.271	-0.002	0.786

Legend: The first column shows the p-value threshold for SNP selection from the GWAS discovery. The numbers of SNPs, which passed the selection criterion, are shown in the second column. The following three sections of the table present effect sizes (B) and p-values comparing each season with winter, estimated simultaneously in a single nominal regression model for each SNP selection threshold. The row in bold represents the primary analysis (P threshold for risk SNPs $p \leq 0.05$); other rows are exploratory.

Table 2. Comparison of schizophrenia PRS of individuals in the UK Biobank sample, born in winter-spring vs summer-autumn ("Whole sample" section) and bottom vs top deciles of schizophrenia PRS of individuals in the UK Biobank sample ("Bottom vs Top deciles of schizophrenia PRS" section).

SNP selection threshold	Whole sample		Bottom vs Top deciles of schizophrenia PRS	
	OR	p-value	OR	p-value
1.00E-04	1.009	0.114	1.034	0.415
0.001	1.011	0.041	1.013	0.759
0.01	1.011	0.040	1.051	0.229
0.05	1.005	0.325	1.075	0.082
0.1	1.004	0.433	1.101	0.020
0.2	1.005	0.350	1.047	0.266
0.3	1.003	0.546	1.035	0.401
0.4	1.003	0.622	1.019	0.646
0.5	1.003	0.638	1.011	0.787

Legend: The baseline of the analysis is winter-spring birth combined. The schizophrenia PRS are generated using schizophrenia risk SNPs at different thresholds for association (column 1). OR (column 2) is the exponentiation of the *B*-coefficient provided by logistic regression.

Figure 1. Frequency of birth by month in the UK Biobank population.

Figure 2. Distribution of schizophrenia polygenic risk scores (PRS) of 136,538 UK Biobank participants with respect to their season of birth. Schizophrenia polygenic risk scores were constructed using schizophrenia risk SNPs with association p-value≤0.05.

Figure 3. Distribution of schizophrenia polygenic risk scores of 136,538 UK Biobank participants with respect to their month of birth. Schizophrenia polygenic risk scores were constructed using schizophrenia risk SNPs with association p-value≤0.05.

Supplementary Table 1. Comparison of schizophrenia PRS of individuals in the UK Biobank sample, split by month of birth. The schizophrenia PRS are generated using schizophrenia risk SNPs at different association p-value thresholds (top row). The base line of the analysis is January and therefore a negative B indicates decrease in risk of schizophrenia in those born in the month, which is shown in the first column as compared to January.

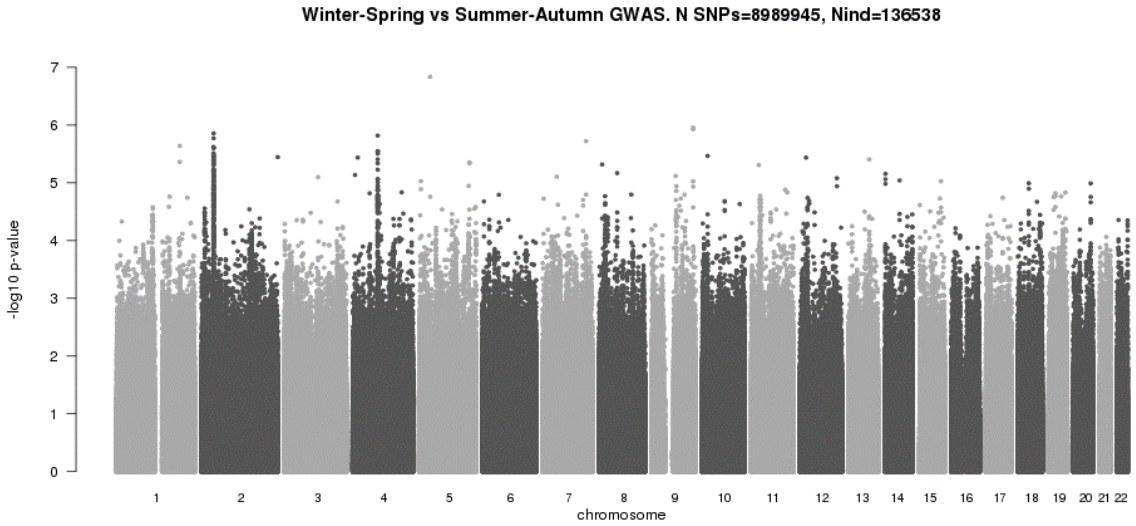
Month	SNP selection p-value threshold	10^{-4}	0.001	0.01	**0.05**	0.1	0.2	0.3	0.4	0.5
February	B	0.001	0.001	-0.013	**-0.011**	-0.011	-0.011	-0.010	-0.009	-0.011
	p-value	*0.915*	*0.951*	*0.345*	***0.392***	*0.429*	*0.392*	*0.473*	*0.500*	*0.418*
March	B	-0.016	-0.002	-0.013	**-0.018**	-0.023	-0.024	-0.024	-0.022	-0.024
	p-value	*0.228*	*0.848*	*0.313*	***0.172***	*0.078*	*0.068*	*0.064*	*0.085*	*0.068*
April	B	0.005	0.007	0.004	**0.001**	-0.005	-0.010	-0.010	-0.012	-0.012
	p-value	*0.726*	*0.606*	*0.737*	***0.918***	*0.708*	*0.467*	*0.444*	*0.358*	*0.376*
May	B	-0.024	-0.025	-0.015	**-0.009**	-0.011	-0.016	-0.018	-0.019	-0.020
	p-value	*0.066*	*0.050*	*0.247*	***0.481***	*0.406*	*0.217*	*0.169*	*0.139*	*0.114*
June	B	-0.004	0.004	0.001	**-0.010**	-0.014	-0.020	-0.023	-0.022	-0.023

	p-value	*0.735*	*0.740*	*0.950*	***0.428***	*0.279*	*0.131*	*0.083*	*0.101*	*0.084*
July	B	0.009	0.013	0.010	**-0.003**	-0.003	-0.008	-0.010	-0.011	-0.013
	p-value	*0.481*	*0.323*	*0.451*	***0.805***	*0.799*	*0.547*	*0.457*	*0.409*	*0.328*
August	B	0.011	0.012	0.010	**0.010**	0.002	0.001	-0.001	-0.003	-0.004
	p-value	*0.396*	*0.350*	*0.463*	***0.467***	*0.898*	*0.915*	*0.911*	*0.825*	*0.739*
September	B	0.009	0.016	0.010	**0.000**	-0.004	-0.007	-0.009	-0.012	-0.012
	p-value	*0.522*	*0.239*	*0.437*	***0.995***	*0.744*	*0.576*	*0.479*	*0.388*	*0.374*
October	B	-0.013	-0.014	-0.015	**-0.016**	-0.014	-0.010	-0.011	-0.011	-0.012
	p-value	*0.328*	*0.284*	*0.261*	***0.230***	*0.279*	*0.467*	*0.423*	*0.410*	*0.371*
November	B	0.007	0.017	0.016	**0.012**	0.009	0.008	0.006	0.004	0.003
	p-value	*0.590*	*0.202*	*0.223*	***0.360***	*0.514*	*0.567*	*0.643*	*0.758*	*0.818*
December	B	0.004	0.005	0.003	**-0.001**	0.001	-0.002	-0.002	-0.002	-0.004
	p-value	*0.761*	*0.696*	*0.810*	***0.946***	*0.932*	*0.883*	*0.902*	*0.876*	*0.746*

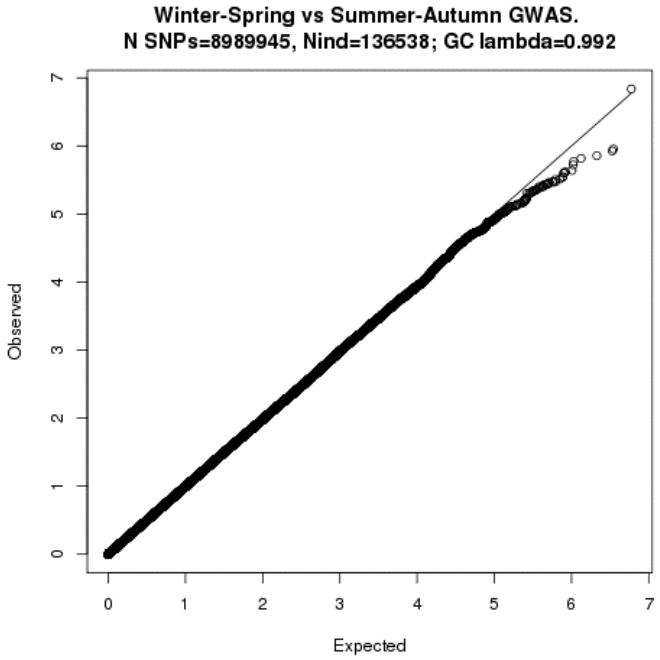
Legend: The first row shows the p-value threshold for SNP selection from the discovery GWAS of schizophrenia. The numbers of SNPs, which passed the selection criterion, are the same as in Supplementary Table 1. The following 11 sections of the table present effect sizes (B) and p-

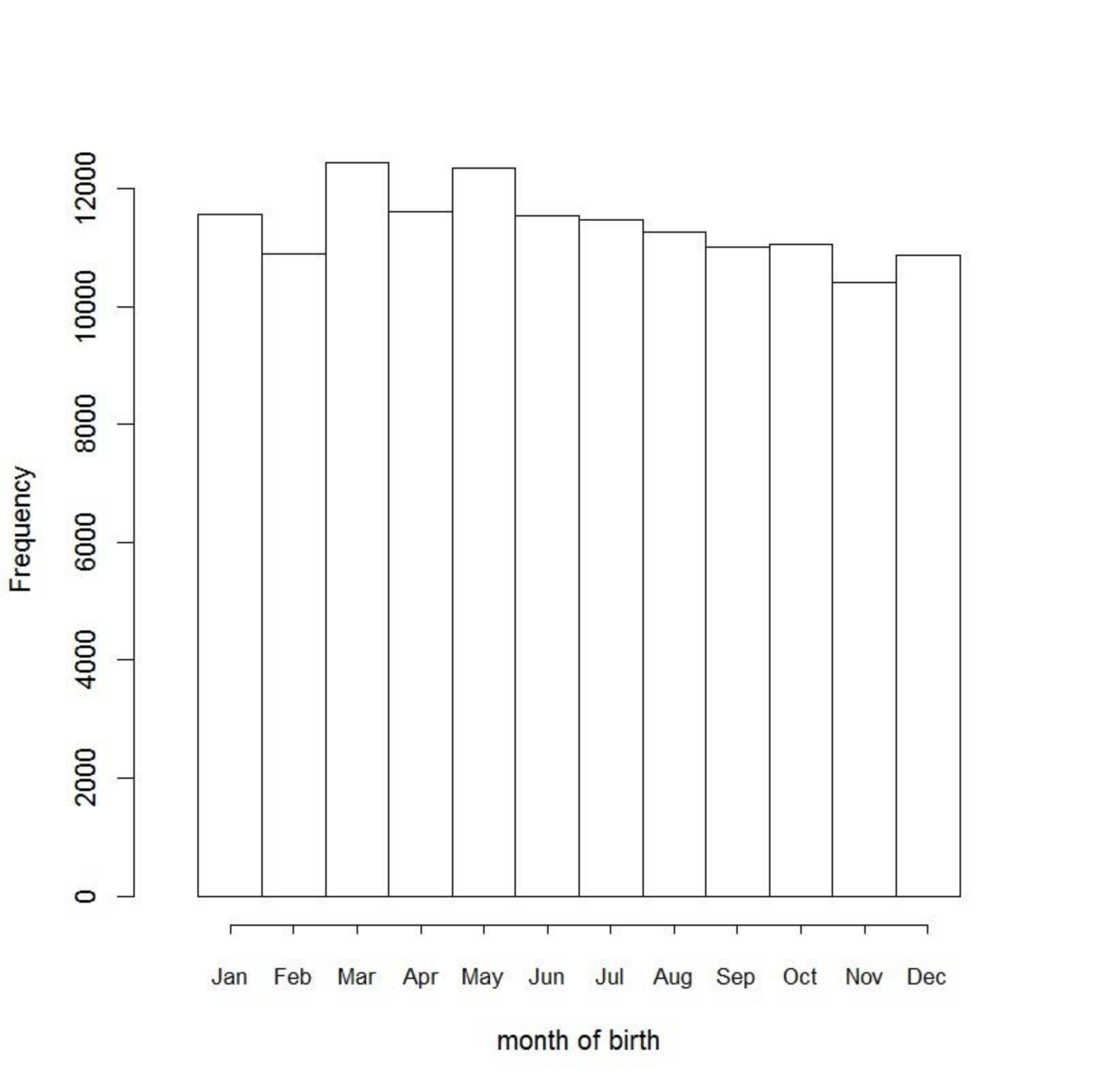
values comparing each month with January, estimated simultaneously in a single nominal regression model for each SNP selection threshold. The column in bold represents the primary analysis (P threshold for risk SNPs $p \leq 0.05$); other columns are exploratory.

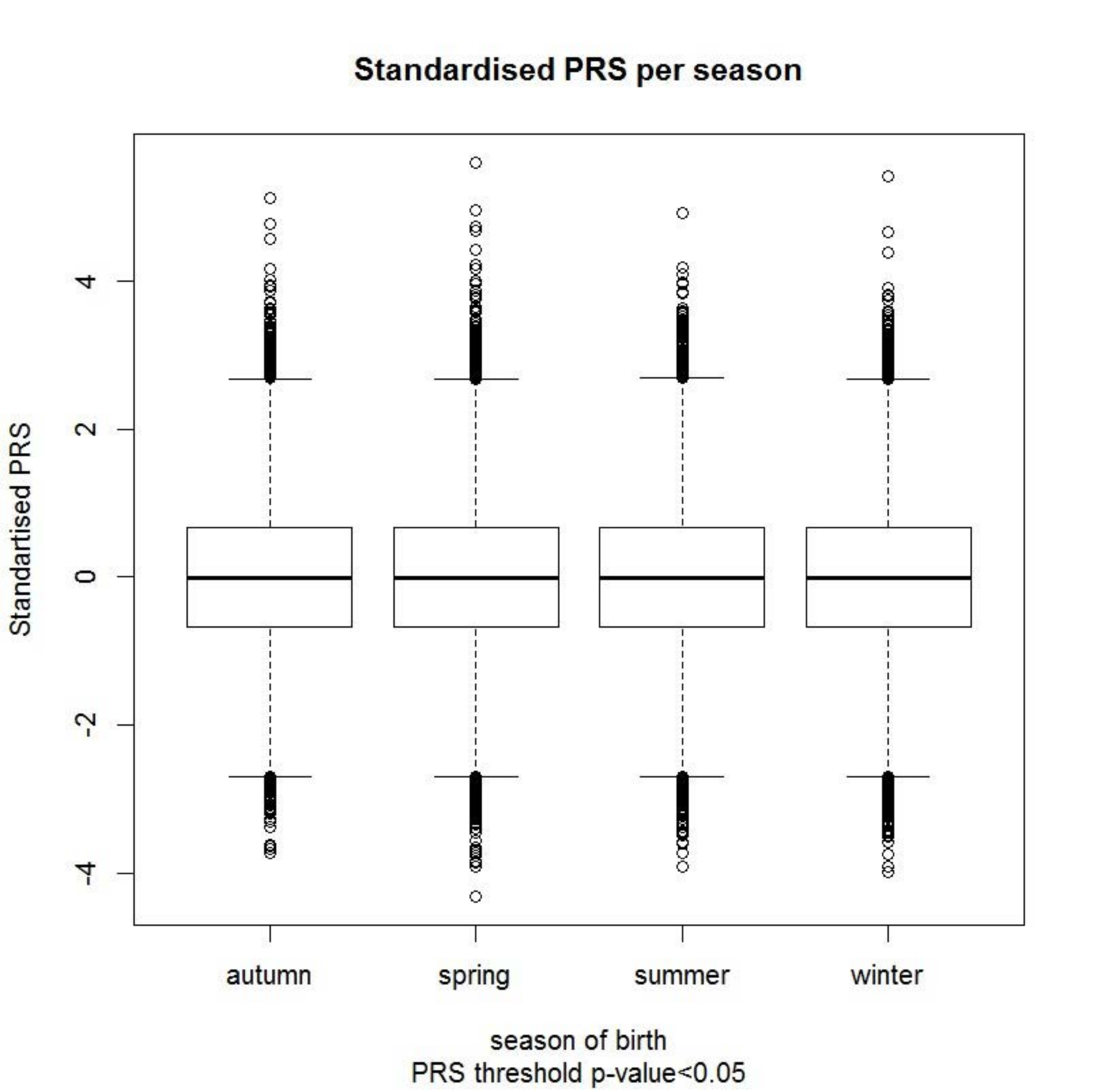
Supplementary Figure 1. Manhattan plot of genome-wide association study of season of birth, comparing allele frequencies of people born in winter-spring with those born in summer-autumn.

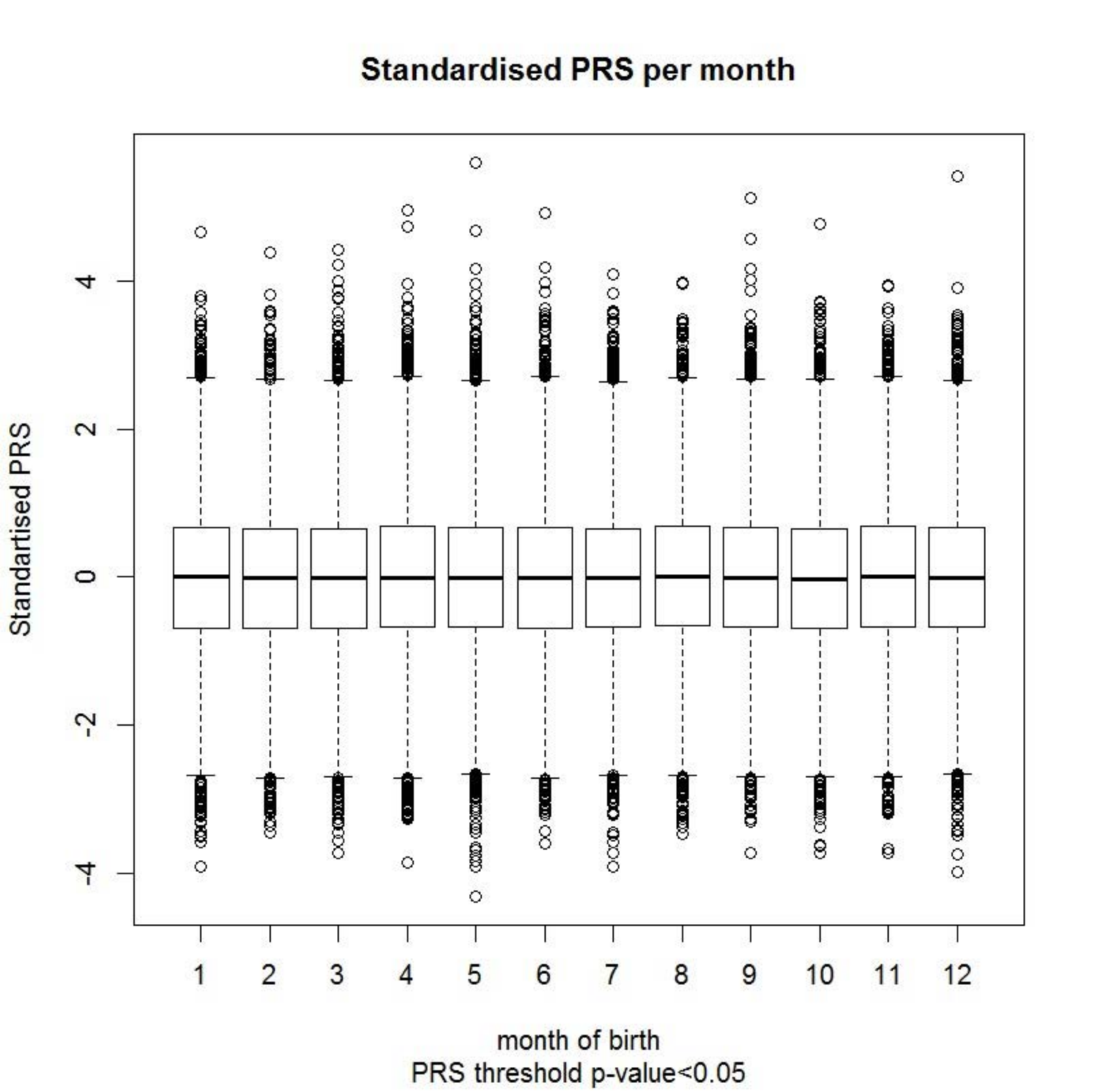


Supplementary Figure 2. QQ plot of genome-wide association study results, comparing allele frequencies of people born in winter-spring with those born in summer-autumn.




Frequency
0
2000
4000
6000
8000
10000
12000
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
month of birth


Standardised PRS per season
Standartised PRS
4
2
0
-2
-4
autumn
spring
summer
winter
season of birth
PRS threshold p-value<0.05


Standardised PRS per month
Standartised PRS
4
2
0
-2
-4
1
2
3
4
5
6
7
8
9
10
11
12
month of birth
PRS threshold p-value<0.05