

The Impact of Phenotypic and Genetic Heterogeneity on Results of Genome Wide Association Studies of Complex Diseases

Mirko Manchia¹✉, Jeffrey Cullis¹✉, Gustavo Turecki², Guy A. Rouleau³, Rudolf Uher¹, Martin Alda^{1*}

1 Department of Psychiatry, Dalhousie University, Halifax, Nova Scotia, Canada, **2** McGill Group for Suicide Studies, Douglas Mental Health University Institute, Montréal, Quebec, Canada, **3** Montreal Neurological Institute, Department of Neurology and Neurosurgery, McGill University, Montréal, Canada

Abstract

Phenotypic misclassification (between cases) has been shown to reduce the power to detect association in genetic studies. However, it is conceivable that complex traits are heterogeneous with respect to individual genetic susceptibility and disease pathophysiology, and that the effect of heterogeneity has a larger magnitude than the effect of phenotyping errors. Although an intuitively clear concept, the effect of heterogeneity on genetic studies of common diseases has received little attention. Here we investigate the impact of phenotypic and genetic heterogeneity on the statistical power of genome wide association studies (GWAS). We first performed a study of simulated genotypic and phenotypic data. Next, we analyzed the Wellcome Trust Case-Control Consortium (WTCCC) data for diabetes mellitus (DM) type 1 (T1D) and type 2 (T2D), using varying proportions of each type of diabetes in order to examine the impact of heterogeneity on the strength and statistical significance of association previously found in the WTCCC data. In both simulated and real data, heterogeneity (presence of “non-cases”) reduced the statistical power to detect genetic association and greatly decreased the estimates of risk attributed to genetic variation. This finding was also supported by the analysis of loci validated in subsequent large-scale meta-analyses. For example, heterogeneity of 50% increases the required sample size by approximately three times. These results suggest that accurate phenotype delineation may be more important for detecting true genetic associations than increase in sample size.

Citation: Manchia M, Cullis J, Turecki G, Rouleau GA, Uher R, et al. (2013) The Impact of Phenotypic and Genetic Heterogeneity on Results of Genome Wide Association Studies of Complex Diseases. PLoS ONE 8(10): e76295. doi:10.1371/journal.pone.0076295

Editor: Andreas Reif, University of Wuerzburg, Germany

Received: February 6, 2013; **Accepted:** August 26, 2013; **Published:** October 11, 2013

Copyright: © 2013 Manchia et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This study was in part funded by grant 64410 from Canadian Institute of Health Research (CIHR) to MA. No additional external funding was received for this study. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: malda@dal.ca

✉ These authors contributed equally to this work.

Introduction

Phenotypic misclassification reduces substantially the power to detect association, particularly in case-control studies [1–6]. Typically these analyses were restricted to rates of misclassification on the order of 1–5%. However, it is conceivable that complex traits may be heterogeneous with respect to genetic susceptibility and disease pathophysiology, and that the effect of phenotypic or genetic heterogeneity (henceforth referred to as “heterogeneity”) is of a larger magnitude than that of phenotypic misclassification. This is particularly relevant for psychiatric disorders [7]. The diagnosis of mental illness is based primarily on descriptive clinical criteria and is typically made in absence of laboratory diagnostic tests or other clinically relevant knowledge of individual pathophysiology. It is possible that depression, psychosis, bipolar disorder (BD), or substance abuse each represent a common phenotypic manifestation of an underlying polygenic diathesis. But it is also conceivable that these syndromes encompass diverse conditions, each with a distinct genetic basis and little overlap with the others [8]. Several such subgroups of major psychiatric disorders, including lithium responsive BD [9] and mood

incongruent psychosis [10] have been proposed based on clinical, familial and biological criteria.

It is also possible that heterogeneity contributes to the discrepancy between the heritability estimates of complex diseases and the proportion of phenotypic variance explained by the identified loci in genome-wide association studies (GWAS). Indeed, the estimated effect sizes of genetic associations in GWAS of complex traits are significantly reduced by imprecise phenotyping [11,12]. This discrepancy appears to be of larger magnitude in psychiatric disorders than in other complex traits [13]. For instance, the heritability of BD has been estimated to be as high as 85% [14,15]. But only a smaller fraction of BD heritability is accounted for by loci identified through GWAS, even when considering all GWAS polymorphisms simultaneously [16–19].

Little attention has been given so far to the extent of the effect of heterogeneity on genetic association findings in common complex diseases. Testing the impact of heterogeneity on GWAS findings requires extensive knowledge of the pathophysiology and genetic architecture of the common trait under study. This assumption makes psychiatric disorders unsuitable for such analysis. On the other hand, another common disease such as diabetes mellitus (DM) is suitable for testing the impact of heterogeneity. The two

types of DM were distinguished only in the late 1930s [20], each characterized by distinct pathophysiology [21] and genetic architecture. If a diagnosis of DM was based on high blood glucose alone, DM type 1 (T1D) cases could not be differentiated from DM type 2 (T2D) cases. The state of DM classification prior to 1930 may well approximate the state of knowledge about psychiatric disorders in the beginning of the 21st century.

Here we investigate the impact of heterogeneity on the statistical power of GWAS. To do so, we first performed a computer simulation study. Next, we analyzed the Wellcome Trust Case-Control Consortium (WTCCC) data for DM T1D and T2D. We combined varying proportions of individual data from each of the diabetes subtypes to examine the impact of heterogeneity on the strength and statistical significance of association and compared these results with those previously found in the WTCCC study. Heterogeneity reduced the statistical power to detect genetic association and greatly decreased the estimates of risk attributed to genetic variation.

Materials and Methods

Simulation of Case-Control Association Analysis Under Heterogeneity

To study the impact of heterogeneity on GWAS findings, we simulated case-control data assuming increasing (from 10% to 90% in 10% steps) phenotypic admixture. Admixture (indicated as β) was the proportion of “non-cases” (i.e. controls) in the case group. Specific genetic models of disease prevalence were used in order to simulate genotypes for case and control populations. We used two basic single-locus genetic models: dominant [22] and multiplicative (or log additive) [23]. The model parameters were the following: population prevalence for the disorder (0.001, 0.01, 0.05, 0.1), minor allele frequency (MAF) (0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5), and the relative risk of the disorder for the minor allele (1.1, 1.2, 1.3, 1.5, 2, 5, 10). We used a Monte Carlo procedure to simulate cases and controls under all combinations of the parameters listed above. Genotypes were assigned using one of two probability distributions, according to the group (case or control) that each individual came from. Using a script in R software (version 2.13.2), we performed two sets of simulations. First, we determined the minimum sample size needed to detect an association in 90% of trials at a significance level set at $p < 5 \times 10^{-8}$. The sample size obtained at each iteration of the simulation was specified according to a binary search that terminated once 90% power (over 10,000 replicates) was achieved, or when the sample size limit of 1,000,000 was reached. In the second set of simulations, we studied the impact of β on the statistical power to detect association at $p < 5 \times 10^{-8}$. To this end we generated 10,000 cases and controls, replicated 1,000 times, for each combination of parameters. For each genetic model (dominant and multiplicative) we obtained case-control frequency tables. Using χ^2 and the Cochran-Armitage-Trend-Test (CATT) module implemented in R we calculated p-values and odds ratios (ORs) for increasing levels of β .

Analysis of WTCCC Type 1 and Type 2 Diabetes Data Under Heterogeneity

We accessed and downloaded all of the genotypic data available for the two control populations (the 1958 Birth Cohort (58C) and the UK National Study (NBS)) and for the T1D and T2D cohorts from the WTCCC website.

A number of quality control (QC) steps were performed on this data in the original WTCCC GWAS study [24]. Individuals and SNPs that were retained had passed each of the following

exclusion criteria: 1) missing data rate $> 3\%$ per sample across all SNPs; 2) heterozygosity $> 30\%$ or $< 23\%$; 3) discrepancies in ID information; 4) ancestry control (outliers after multi-dimensional scaling); 5) duplicated samples (identity $> 99\%$); 6) relatedness (86%–96% identity); 7) missing data rate $> 5\%$ per SNP; 8) missing data rate $> 1\%$ when $MAF < 5\%$; 9) Hardy-Weinberg exact p-value $< 5.7 \times 10^{-7}$ (in 2,938 controls) [24].

Since the QC analysis in the original study was performed across $\sim 14,000$ subjects (compared to the $\sim 7,000$ available for the present study), we also applied exclusion lists supplied in the WTCCC data in addition to in-house QC. The scripts used for QC are available online.

Two recoded phenotype files were created each containing 2,938 control individuals (the combination of all the filtered 58C and NBS individuals) and the 1,963 T1D or the 1,924 T2D cases, respectively. Next, we performed the WTCCC standard case-control association analysis (χ^2) to confirm the validity of the QC steps using PLINK [25]. Results of this analysis are outlined in Table 1.

To investigate the effect of heterogeneity, we examined the 20 SNPs most significantly associated with T1D or T2D, respectively. These SNPs are independent and in linkage equilibrium to each other, with the exception of rs9939609 and rs7193144 within the FTO gene which are in strong linkage disequilibrium ($R^2 = 0.97$, $D' = 1$) and are both associated with T2D. Further, we studied 21 and 16 polymorphisms found associated in subsequent large-scale meta-analyses for T1D [26–29] and T2D [30], respectively. The selected SNPs were genotyped in the WTCCC study and were required to have $p < 10^{-6}$ in at least one meta-analysis study. We created alternate phenotype files in order to simulate heterogeneity. Let N_1 and N_2 denote the respective totals of T1D and T2D cases. At each 10% increment of β , we removed $N_1 * \beta$ subjects randomly selected from the T1D population and replaced them with $N_1 * \beta$ subjects randomly chosen from the T2D population. Similarly, we analyzed T2D data, replacing $N_2 * \beta$ T2D with T1D cases. For each of T1D and T2D, 100 alternate samples were created at each 10% increment of β . Association analyses were then run using the χ^2 test in PLINK [25] on each alternate phenotype file. We took the $-\log(p\text{-value})$ from each of the 100 association analyses at each β level, and then calculated the arithmetic mean of the $-\log(p\text{-value})$. The scripts developed for the analyses presented in the study are available at the following URL: <http://web.cs.dal.ca/~cullis/heterogeneity/>.

Results

Simulated Case-Control Association Analysis Under Heterogeneity

We simulated case-control data assuming increasing β . We present the results for a disease prevalence of 0.01. This value is relevant to complex traits such as BD or schizophrenia. Results for other prevalence values can be found in the Table S1 and S2. In all analyses, the admixture level significantly affected the sample size required to reach 90% power at the significance level of $p < 5 \times 10^{-8}$. Specifically, there was a substantial loss of statistical power that was disproportionately larger than the degree of heterogeneity. An increase in the proportion of “non-cases” resulted in a non-linear increase of the sample size needed to achieve 90% of statistical power. As shown in Figure 1, this effect was evident both in the dominant and multiplicative models and for different relative risks ($RR = 1.2, 1.5$ and 2). For instance, with β at 50%, the sample size needed to achieve the same statistical power without admixture was three times larger, particularly with genetic effect sizes equal to or larger than 1.5.

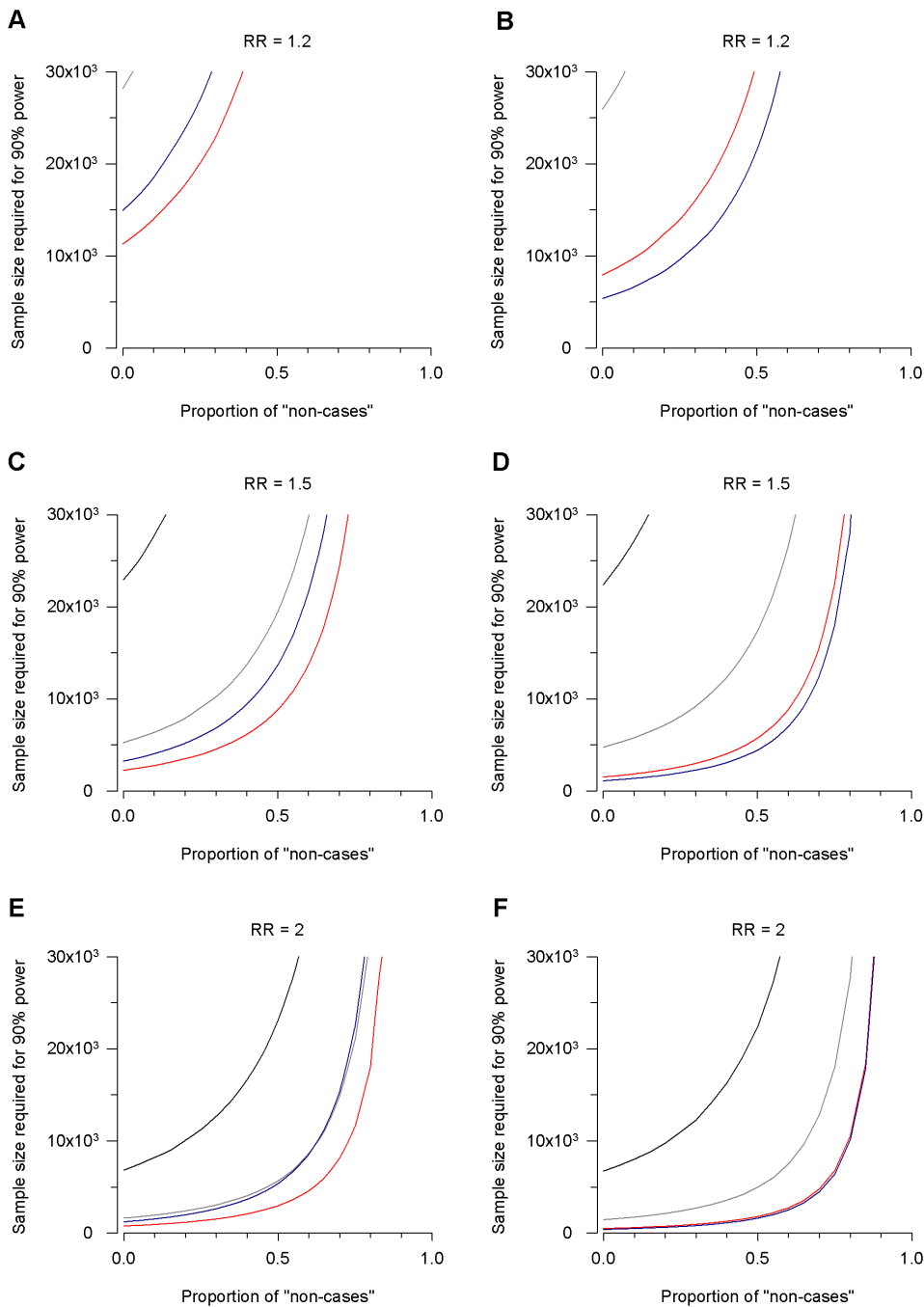


Figure 1. The impact of heterogeneity on the sample size (cases and controls) required for 90% of statistical power. The minimum sample size to achieve to detect association was calculated in simulated case-control data with increasing proportion of “non-cases” considering a disease prevalence of 0.01. Data are reported for minor allele frequencies (MAF) of 0.01 (black), 0.05 (grey), 0.2 (red) and 0.5 (blue). The results are reported for dominant (panels A, C, and E) and multiplicative (panels B, D, and F) genetic models. RR=relative risk. doi:10.1371/journal.pone.0076295.g001

Heterogeneity also resulted in a marked reduction of the estimated effect size which is reflected here as a decrease in the estimated ORs. This reduction in effect size was found to be non-linear in relation to the degree of β and was more pronounced for larger relative risk values (Figure 2).

Application to WTCCC Type 1 and Type 2 Diabetes Data

To further test the validity of the results identified in the computer simulation we analyzed the WTCCC T1D and T2D

data [24]. In both datasets we examined the 20 most significantly associated SNPs. Of these, 7 loci for T1D and 3 for T2D reached genome-wide significance and an additional 13 and 17 loci, respectively, showed moderate association. The results of this analysis are outlined in Figure 3. The strength of the association for T1D decreased substantially as the proportion of T2D cases in the sample was increased. Most of the significantly associated SNPs became equivocal at a relatively low degree of β (between 20% and 30%). Only the association of *HLA-DRB1* with T1D [31]

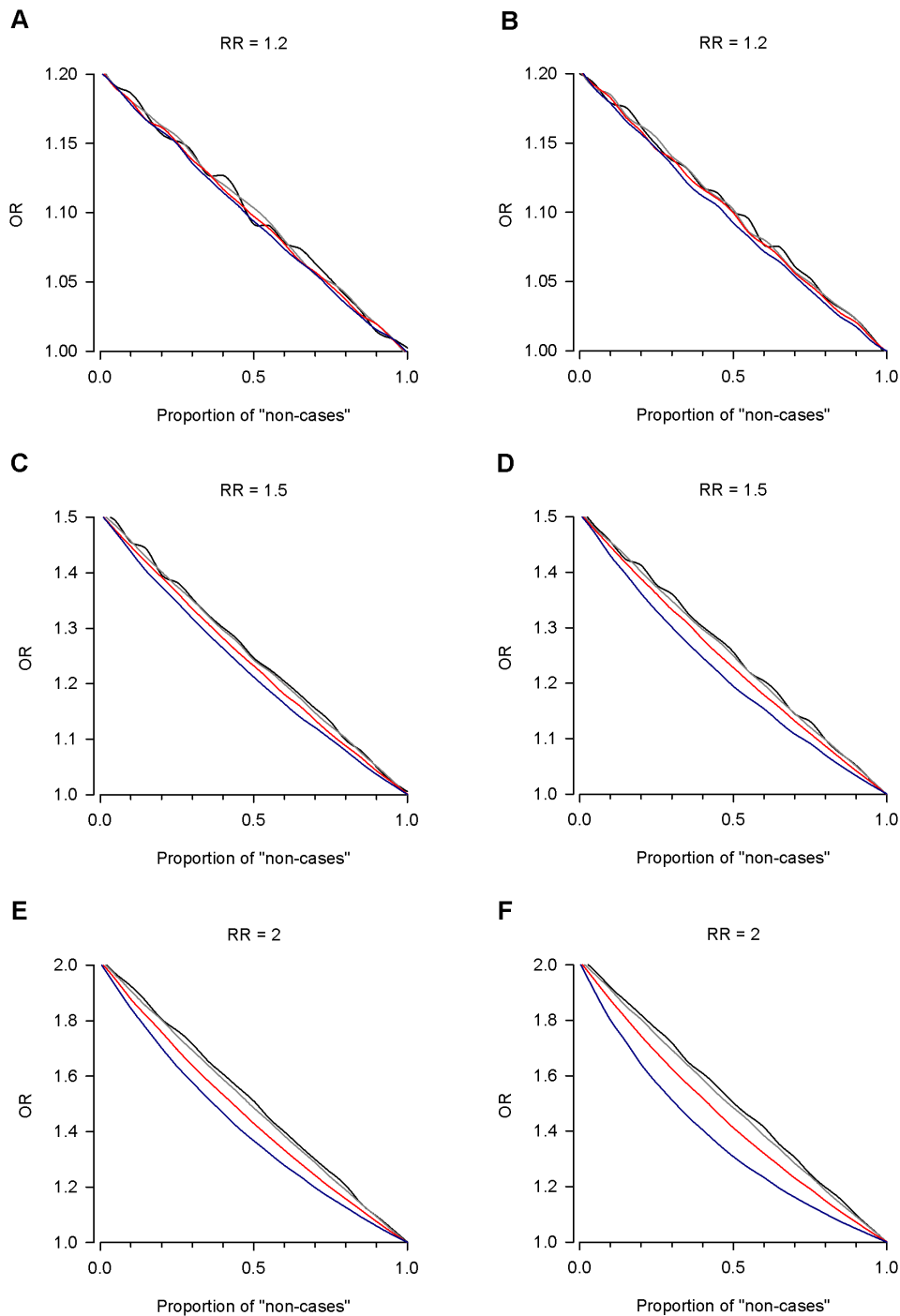


Figure 2. The impact of heterogeneity on the estimation of the genetic effect size. Odds ratios from simulated case-control data were calculated for each step of admixture. Data are reported for minor allele frequencies (MAF) of 0.01 (black), 0.05 (grey), 0.2 (red) and 0.5 (blue). The results are reported for dominant (panels A, C, and E) and multiplicative (panels B, D, and F) genetic models. RR=relative risk; OR=odds ratio. doi:10.1371/journal.pone.0076295.g002

was robust to the effects of heterogeneity, with a significant effect present with up to 90% of the sample made up of T2D cases. Similarly, the strength of the association signals for T2D progressively diminished as T1D cases gradually replaced the “true” cases of T2D. Further support came from the analysis of loci found associated in subsequent meta-analyses of T1D and T2D (Figures S1 and S2). Again, with increasing degree of β the magnitude of association of the SNPs declined substantially.

Discussion

The purpose of this study was to quantify the impact of heterogeneity in the analysis and interpretation of GWAS findings. We showed that the presence of heterogeneity (presence of “non-cases”) reduced both the statistical power as well as the observed risks attributed to susceptibility alleles or genotypes. These findings were supported by the analysis of both simulated case-control and

T1D			WTCCC p-values					β						
Reported SNP	Trend p-value	Genotypic p-value	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	
rs9272346	133.62	133.26	140.30	113.30	89.65	68.25	51.03	36.63	24.42	15.28	8.19	3.28	0.65	
rs6679677	25.93	25.27	25.49	21.29	17.86	14.48	11.48	9.08	6.35	4.35	2.70	1.37	0.46	
rs17696736	14.66	13.82	14.51	12.16	9.97	8.13	6.51	4.62	3.41	2.41	1.49	0.80	0.28	
rs11171739	10.94	10.01	11.00	9.15	7.33	5.75	4.47	3.27	2.12	1.42	0.77	0.31	0.02	
rs9746695	8.09	7.31	8.24	6.85	5.25	4.06	3.13	2.25	1.45	0.88	0.50	0.15	0.14	
rs12708716	7.03	6.31	7.27	5.96	4.45	3.31	2.45	1.70	1.00	0.53	0.22	0.14	0.42	
rs17388568	6.30	5.49	6.34	5.84	5.36	4.75	4.17	3.95	3.28	2.90	2.69	2.23	1.90	
rs2542151	5.72	4.94	5.67	4.59	3.57	2.61	1.76	1.21	0.71	0.30	0.15	0.33	0.70	
rs2639703	5.07	4.76	5.23	4.92	4.60	4.30	3.81	3.34	3.02	2.89	2.47	2.19	1.92	
rs12061474	4.90	4.58	5.11	4.59	4.01	3.48	2.96	2.70	2.45	2.02	1.52	1.26	0.96	
rs1520703	4.97	4.21	5.07	4.37	3.66	2.88	2.37	1.94	1.36	1.13	0.76	0.45	0.20	
rs2269241	4.91	4.17	5.02	4.09	3.50	2.74	2.27	1.67	1.20	0.90	0.53	0.22	0.02	
rs2104286	5.10	4.36	4.99	4.30	3.64	3.06	2.44	2.11	1.66	1.18	0.91	0.58	0.36	
rs2544677	5.08	4.35	4.98	4.35	3.89	3.29	2.86	2.31	1.92	1.61	1.18	0.85	0.59	
rs7722135	4.77	4.10	4.82	4.24	3.84	3.27	2.87	2.34	2.00	1.68	1.26	0.94	0.68	
rs1025039	4.80	4.05	4.71	3.98	3.30	2.85	2.27	1.86	1.40	1.01	0.70	0.40	0.18	
rs10221582	4.63	4.02	4.69	4.29	3.67	3.38	2.99	2.80	2.28	2.00	1.60	1.37	1.10	
rs13096142	4.63	4.61	4.66	4.19	3.57	3.04	2.74	2.30	1.87	1.48	1.22	0.92	0.63	
rs7221109	4.66	4.01	4.65	4.04	3.60	3.09	2.45	2.06	1.76	1.23	1.04	0.75	0.49	
rs2666236	4.67	4.17	4.63	3.90	3.39	2.82	2.29	1.96	1.40	1.12	0.81	0.49	0.26	

T2D			WTCCC p-values					β						
Reported SNP	Trend p-value	Genotypic p-value	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	
rs4506565	12.25	11.30	11.98	9.21	6.89	4.85	3.29	1.93	1.06	0.34	0.23	0.76	1.62	
rs7193144	7.84	7.32	7.31	5.95	4.52	3.43	2.48	1.75	1.10	0.51	0.21	0.13	0.43	
rs9939609	7.28	6.72	7.19	5.86	4.49	3.42	2.50	1.80	1.15	0.56	0.24	0.11	0.35	
rs1495377	5.88	5.19	5.87	5.01	4.19	3.44	3.00	2.21	1.67	1.23	0.83	0.53	0.27	
rs9465871	5.99	6.48	5.71	5.13	4.53	4.22	3.40	2.84	2.54	2.12	1.81	1.39	1.08	
rs6718526	5.62	4.94	5.37	4.48	3.46	2.79	2.12	1.64	1.15	0.82	0.44	0.21	0.02	
rs9326506	5.11	4.52	5.06	4.21	3.42	2.88	2.12	1.72	1.24	0.75	0.50	0.26	0.05	
rs2930291	5.11	4.36	5.05	4.28	3.63	3.24	2.54	2.02	1.67	1.21	0.96	0.59	0.33	
rs2903265	5.02	4.30	5.05	4.51	3.95	3.42	3.01	2.57	2.15	1.81	1.45	1.10	0.87	
rs17248501	4.86	4.12	5.01	4.43	4.02	3.63	3.25	2.72	2.38	1.96	1.60	1.29	1.04	
rs2099106	4.76	4.53	4.87	4.60	4.19	3.82	3.39	3.23	2.87	2.63	2.24	1.93	1.75	
rs11178531	5.00	4.43	4.83	3.98	3.20	2.40	2.08	1.37	0.96	0.63	0.28	0.12	0.10	
rs1665901	4.61	3.86	4.70	4.44	3.95	3.63	3.23	3.16	2.78	2.59	2.09	1.95	1.69	
rs11688935	4.67	4.08	4.59	4.13	3.48	3.07	2.61	2.29	1.81	1.52	1.16	0.88	0.62	
rs6691406	4.47	3.74	4.53	3.91	3.64	2.97	2.59	2.23	1.88	1.63	1.19	0.98	0.73	
rs8012854	4.40	3.67	4.50	3.95	3.52	3.06	2.62	2.20	1.85	1.47	1.14	0.86	0.65	
rs440646	4.35	3.65	4.31	3.68	3.17	2.81	2.20	1.77	1.34	1.02	0.72	0.54	0.29	
rs769097	4.14	3.51	4.25	4.03	3.68	3.47	3.00	2.80	2.48	2.40	2.14	1.91	1.65	
rs6846031	4.25	4.01	4.17	3.97	3.76	3.74	3.57	3.43	3.41	3.20	2.92	2.82	2.67	
rs9391949	4.20	3.61	4.16	3.94	3.73	3.65	3.38	3.34	2.91	2.72	2.57	2.31	2.14	

< 5×10 ⁻⁸	< 10 ⁻⁷	< 10 ⁻⁶	< 10 ⁻⁵	< 10 ⁻⁴	< 10 ⁻³	< 10 ⁻²	< 10 ⁻¹	< 1
----------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	-----

Figure 3. Genome-wide analysis of the Wellcome Trust Case-Control Consortium (WTCCC) data for diabetes type 1 (T1D) and type 2 (T2D) under heterogeneity: twenty most significant associations [-log(p-values)]. SNP = single nucleotide polymorphism; β = admixture. doi:10.1371/journal.pone.0076295.g003

GWAS data from WTCCC T1D and T2D cohorts [24]. We also tested our hypothesis by analyzing loci replicated and validated in large-scale meta-analyses irrespective of their association strength in the WTCCC study, as these are more likely to represent true associations. The same pattern was observed for these loci: the magnitude of association declined substantially with increasing heterogeneity.

Several findings deserve comment. First, irrespective of the various parameters and genetic models tested, the absence of

heterogeneity in samples allows for a power to detect association that would require a far larger sample size if a medium or high level of heterogeneity were present. A recent study examining the impact of diagnostic misclassification in genetic studies by Wray et al. [32] has reached similar results with a different method. In the Supplementary Information to their paper, Wray et al. [32] provide an analytical solution that allows the calculation of statistical power under misclassification without computer simulation. These observations are consistent with the findings of

pharmacogenomic GWAS. Indeed, the robustness of the phenotypic measure of treatment response given by specific biomarkers (for example identifiable in serum) allowed the identification of significant association signals even with relatively small sample sizes [33]. In addition, pharmacogenetic traits have not been subject to selection and so larger effect sizes may exist compared to complex traits.

In the absence of more reliable measures (for instance diagnostic tests) in more heterogeneous complex traits – including many psychiatric and neurological diseases – researchers have obtained promising results by selecting more homogeneous subgroups of patients such as responders to lithium treatment in BD [34], or early illness onset in Alzheimer's disease [35]. This strategy can significantly decrease the number of cases to collect for adequately powered genetic studies.

We have also shown that the presence of heterogeneity strongly influences the estimates of the effect sizes. This result suggests that the effect sizes of marginally significant variants identified in GWAS might be larger if they were instead examined in more homogeneous samples. This is in agreement with previous studies pointing to the importance of accurate and stringent phenotypic definition in GWAS data [11,12]. Interestingly, van der Sluis et al. [12] employed simulation studies to show that accurate modelling of phenotypic information improved the estimation of the genetic variants effect size. Further, Evangelou et al. [11] found empirical support for importance of phenotypic definition in the analysis of GWAS data in HIV-1 infected subjects. They showed that the observed genetic effects in HIV-1 seroconverters (a more stringent phenotype) were larger than in seroprevalent subjects. Using simulation and real data both studies consistently showed how accurate phenotyping increased the power to detect association signals and to estimate correctly their effect size.

Here we observed that as the degree of heterogeneity increased, so did the minimum sample size required to achieve sufficient statistical power. This effect was particularly evident for genotypic relative risk values on the order of 1.2 and 1.5 and for common variants ($MAF > 0.05$). Interestingly, for a complex trait such as BD, most of the reported genetic risk variants have an associated effect size of approximately 1.5 or below [24,36–42]. At least in clinically heterogeneous disorders, it is conceivable that even collecting large sample sizes could only partially compensate for the loss of statistical power in GWAS. In this context it appears crucial to focus on the steps preceding the GWAS. Careful clinical history from all available sources, consensus diagnosis, validity of the phenotypic measures used, evaluation of the inter-rater agreement and reliability and use of prospective design could all help in overcoming the issue of phenotypic heterogeneity.

In conclusion, we demonstrated that heterogeneity can have a major impact on GWAS findings. The extent of its effect is of large magnitude and, unexpectedly, affects significantly even loci robustly associated with the trait under study. We showed that even a relatively low proportion of “non cases” (20%) can dilute

the observed genetic effect size for common variants (with $MAF > 0.05$) targeted in the current GWAS analyses. The partial failure of GWAS to detect a substantial proportion of the heritability of genetic complex diseases could be a consequence of the presence of a high degree of heterogeneity.

Supporting Information

Figure S1 Impact of the admixture of diabetes type 1 (T1D) and type 2 (T2D) on Wellcome Trust Case-Control Consortium (WTCCC) T1D findings confirmed in large scale meta-analyses [$-\log(p\text{-values})$]. SNP = single nucleotide polymorphism; β = admixture; nr = not reported. (TIF)

Figure S2 Impact of the admixture of diabetes type 1 (T1D) and type 2 (T2D) on Wellcome Trust Case-Control Consortium (WTCCC) T2D findings confirmed in large scale meta-analysis [$-\log(p\text{-values})$]. SNP = single nucleotide polymorphism; β = admixture; nr = not reported. *p-value reported in the Zeggini et al. [1] study. Reference: 1. Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, et al. (2007) Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* 316: 1336–1341. (TIF)

Table S1 Results of case-control simulation to determine the minimum sample size needed to detect an association in 90% of trials at a significance level set at $p < 5 \times 10^{-8}$ with increasing levels of admixture (β). Results are presented for dominant and multiplicative genetic models with the following parameters: population prevalence for the disorder (0.001, 0.01, 0.05, 0.1), minor allele frequency (MAF) (0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5), and the relative risk of the disorder for the minor allele (1.1, 1.2, 1.3, 1.5, 2, 5, 10). (XLS)

Table S2 Case-control simulation results of the impact of increasing level of admixture (β) on the statistical power to detect association at $p < 5 \times 10^{-8}$. Results are presented for dominant and multiplicative genetic models with the following parameters: population prevalence for the disorder (0.001, 0.01, 0.05, 0.1), minor allele frequency (MAF) (0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5), and the relative risk of the disorder for the minor allele (1.1, 1.2, 1.3, 1.5, 2, 5, 10). (XLS)

Author Contributions

Conceived and designed the experiments: MA. Performed the experiments: MM JC MA. Analyzed the data: MM JC RU MA. Contributed reagents/materials/analysis tools: GT GAR MA. Wrote the paper: MM JC RU MA.

References

- Gordon D, Yang Y, Haynes C, Finch SJ, Mendell NR, et al. (2004) Increasing power for tests of genetic association in the presence of phenotype and/or genotype error by use of double-sampling. *Stat Appl Genet Mol Biol* 3: Article26.
- Edwards BJ, Haynes C, Levenstien MA, Finch SJ, Gordon D (2005) Power and sample size calculations in the presence of phenotype errors for case/control genetic association studies. *BMC Genet* 6: 18.
- Ji F, Yang Y, Haynes C, Finch SJ, Gordon D (2005) Computing asymptotic power and sample size for case-control genetic association studies in the presence of phenotype and/or genotype misclassification errors. *Stat Appl Genet Mol Biol* 4: Article37.
- Barral S, Haynes C, Stone M, Gordon D (2006) LRTae: improving statistical power for genetic association with case/control data when phenotype and/or genotype misclassification errors are present. *BMC Genet* 7: 24.
- Gordon D, Haynes C, Yang Y, Kramer PL, Finch SJ (2007) Linear trend tests for case-control genetic association that incorporate random phenotype and genotype misclassification error. *Genet Epidemiol* 31: 853–870.
- Buyske S, Yang G, Matise TC, Gordon D (2009) When a case is not a case: effects of phenotype misclassification on power and sample size requirements for the transmission disequilibrium test with affected child trios. *Hum Hered* 67: 287–292.
- Angst J (2007) Psychiatric diagnoses: the weak component of modern research. *World Psychiatry* 6: 94–95.

8. Corvin A, Craddock N, Sullivan PF (2010) Genome-wide association studies: a primer. *Psychol Med* 40: 1063–1077.
9. Alda M, Grof P, Rouleau GA, Turecki G, Young LT (2005) Investigating responders to lithium prophylaxis as a strategy for mapping susceptibility genes for bipolar disorder. *Prog Neuropsychopharmacol Biol Psychiatry* 29: 1038–1045.
10. Goes FS, Zandi PP, Miao K, McMahon FJ, Steele J, et al. (2007) Mood-incongruent psychotic features in bipolar disorder: familial aggregation and suggestive linkage to 2p11-q14 and 13q21–33. *Am J Psychiatry* 164: 236–247.
11. Evangelou E, Fellay J, Colombo S, Martinez-Picado J, Obel N, et al. (2011) Impact of phenotype definition on genome-wide association signals: empirical evaluation in human immunodeficiency virus type 1 infection. *Am J Epidemiol* 173: 1336–1342.
12. van der Sluis S, Verhage M, Posthuma D, Dolan CV (2010) Phenotypic complexity, measurement bias, and poor phenotypic resolution contribute to the missing heritability problem in genetic association studies. *PLoS One* 5: e13929.
13. Visscher PM, Brown MA, McCarthy MI, Yang J (2012) Five years of GWAS discovery. *Am J Hum Genet* 90: 7–24.
14. McGuffin P, Rijdsdijk F, Andrew M, Sham P, Katz R, et al. (2003) The heritability of bipolar affective disorder and the genetic relationship to unipolar depression. *Arch Gen Psychiatry* 60: 497–502.
15. Bienvenu OJ, Davydov DS, Kendler KS (2011) Psychiatric ‘diseases’ versus behavioral disorders and degree of genetic influence. *Psychol Med* 41: 33–40.
16. Gershon ES, Alliey-Rodriguez N, Liu C (2011) After GWAS: searching for genetic risk for schizophrenia and bipolar disorder. *Am J Psychiatry* 168: 253–256.
17. Lee SH, Wray NR, Goddard ME, Visscher PM (2011) Estimating missing heritability for disease from genome-wide association studies. *Am J Hum Genet* 88: 294–305.
18. Sklar P, Ripke S, Scott LJ, Andreassen OA, Cichon S, et al. (2011) Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat Genet* 43: 977–983.
19. Visscher PM, Brown MA, McCarthy MI, Yang J (2012) Five years of GWAS discovery. *Am J Hum Genet* 90: 7–24.
20. Himsworth HP (1940) Insulin Deficiency and Insulin Inefficiency. *Br Med J* 1: 719–722.
21. Himsworth HP (2011) Diabetes mellitus: its differentiation into insulin-sensitive and insulin-insensitive types. *Diabet Med* 28: 1440–1444.
22. Lewis CM (2002) Genetic association studies: design, analysis and interpretation. *Brief Bioinform* 3: 146–153.
23. Risch N (1990) Linkage strategies for genetically complex traits. I. Multilocus models. *Am J Hum Genet* 46: 222–228.
24. Wellcome Trust Case Control Consortium (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447: 661–678.
25. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81: 559–575.
26. Cooper JD, Smyth DJ, Smiles AM, Plagnol V, Walker NM, et al. (2008) Meta-analysis of genome-wide association study data identifies additional type 1 diabetes risk loci. *Nat Genet* 40: 1399–1401.
27. Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, et al. (2009) Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat Genet* 41: 703–707.
28. Qu HQ, Bradfield JP, Li Q, Kim C, Frackelton E, et al. (2010) In silico replication of the genome-wide association results of the Type 1 Diabetes Genetics Consortium. *Hum Mol Genet* 19: 2534–2538.
29. Bradfield JP, Qu HQ, Wang K, Zhang H, Sleiman PM, et al. (2011) A genome-wide meta-analysis of six type 1 diabetes cohorts identifies multiple associated loci. *PLoS Genet* 7: e1002293.
30. Morris AP, Voight BF, Teslovich TM, Ferreira T, Segre AV, et al. (2012) Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet* 44: 981–990.
31. Sheehy MJ, Scharf SJ, Rowe JR, Neme de Gimenez MH, Meske LM, et al. (1989) A diabetes-susceptible HLA haplotype is best defined by a combination of HLA-DR and -DQ alleles. *J Clin Invest* 83: 830–835.
32. Wray NR, Lee SH, Kendler KS (2012) Impact of diagnostic misclassification on estimation of genetic correlations using genome-wide genotypes. *Eur J Hum Genet* 20: 668–674.
33. Daly AK (2010) Genome-wide association studies in pharmacogenomics. *Nat Rev Genet* 11: 241–246.
34. Lopez de Lara C, Jaitovich-Groisman I, Cruceanu C, Mamdani F, Lebel V, et al. (2010) Implication of synapse-related genes in bipolar disorder by linkage and gene expression analyses. *Int J Neuropsychopharmacol* 13: 1397–1410.
35. Davies P (1986) The genetics of Alzheimer’s disease: a review and a discussion of the implications. *Neurobiol Aging* 7: 459–466.
36. Sklar P, Smoller JW, Fan J, Ferreira MA, Perlis RH, et al. (2008) Whole-genome association study of bipolar disorder. *Mol Psychiatry* 13: 558–569.
37. Ferreira MA, O’Donovan MC, Meng YA, Jones IR, Ruderfer DM, et al. (2008) Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. *Nat Genet* 40: 1056–1058.
38. Smith EN, Bloss CS, Badner JA, Barrett T, Belmonte PL, et al. (2009) Genome-wide association study of bipolar disorder in European American and African American individuals. *Mol Psychiatry* 14: 755–763.
39. Sklar P, Ripke S, Scott LJ, Andreassen OA, Cichon S, et al. (2011) Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat Genet* 43: 977–983.
40. Smith EN, Koller DL, Panganiban C, Szelinger S, Zhang P, et al. (2011) Genome-wide association of bipolar disorder suggests an enrichment of replicable associations in regions near genes. *PLoS Genet* 7: e1002134.
41. Cichon S, Muhleisen TW, Degenhardt FA, Mattheisen M, Miro X, et al. (2011) Genome-wide association study identifies genetic variation in neurocan as a susceptibility factor for bipolar disorder. *Am J Hum Genet* 88: 372–381.
42. Vassos E, Steinberg S, Cichon S, Breen G, Sigurdsson E, et al. (2012) Replication study and meta-analysis in European samples supports association of the 3p21.1 locus with bipolar disorder. *Biol Psychiatry* 72: 645–650.