BRIEF REPORT

Genetic Epidemiology

OFFICIAL JOURNAL
INTERNATIONAL GENETIC EPIDEMIOLOGY SOCIETY
www.geneticepi.org

The Case-Only Test for Gene–Environment Interaction is Not Uniformly Powerful: An Empirical Example

Chen Wu,[1,2,3] Jiang Chang,[2,3] Baoshan Ma,[1] Xiaoping Miao,[4] Yifeng Zhou,[5] Yu Liu,[2,3] Yun Li,[6] Tangchun Wu,[4] Zhibin Hu,[7] Hongbing Shen,[7] Weihua Jia,[8,9] Yixin Zeng,[8,9] Dongxin Lin,[2,3]* and Peter Kraft[1]*

[1]Program in Molecular and Genetic Epidemiology, Harvard School of Public Health, Boston, Massachusetts; [2]State Key Laboratory of Molecular Oncology, Cancer Institute and Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China; [3]Department of Etiology and Carcinogenesis, Cancer Institute and Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China; [4]Key Laboratory for Environment and Health (Ministry of Education), School of Public Health, Huazhong University of Sciences and Technology, Wuhan, Hubei, China; [5]Laboratory of Cancer Molecular Genetics, Medical College of Soochow University, Suzhou, Jiangsu, China; [6]Department of Genetics,Department of Biostatistics, University of North Carolina, Chapel Hill, North Carolina; [7]Department of Epidemiology and Biostatistics,Cancer Center, Nanjing Medical University, Nanjing, Jiangsu, China; [8]State Key Laboratory of Oncology in Southern China, Sun Yat-Sen University Cancer Center, Guangzhou, Guangdong, China; [9]Department of Experimental Research, Sun Yat-Sen University Cancer Center, Guangzhou, Guangdong, China

Received 14 December 2012; Revised 14 January 2013; accepted revised manuscript 5 February 2013.
Published online 13 March 2013 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/gepi.21713

ABSTRACT: The case-only test has been proposed as a more powerful approach to detect gene–environment (G × E) interactions. This approach assumes that the genetic and environmental factors are independent. Although it is well known that Type I error rate will increase if this assumption is violated, it is less widely appreciated that G × E correlation can also lead to power loss. We illustrate this phenomenon by comparing the performance of the case-only test to other approaches to detect G × E interactions in a genome-wide association study (GWAS) of esophageal squamous-cell carcinoma (ESCC) in Chinese populations. Some of these approaches do not use information on the correlation between exposure and genotype (standard logistic regression), whereas others seek to use this information in a robust fashion to boost power without increasing Type I error (two-step, empirical Bayes, and cocktail methods). G × E interactions were identified involving drinking status and two regions containing genes in the alcohol metabolism pathway, 4q23 and 12q24. Although the case-only test yielded the most significant tests of G × E interaction in the 4q23 region, the case-only test failed to identify significant interactions in the 12q24 region which were readily identified using other approaches. The low power of the case-only test in the 12q24 region is likely due to the strong inverse association between the single nucleotide polymorphism (SNPs) in this region and drinking status. This example underscores the need to consider multiple approaches to detect G × E interactions, as different tests are more or less sensitive to different alternative hypotheses and violations of the G × E independence assumption.
Genet Epidemiol 37:402–407, 2013. © 2013 Wiley Periodicals, Inc.

KEY WORDS: case-only test; gene–environment interaction; esophageal squamous-cell carcinoma; alcohol drinking; SNP

Introduction

The genome-wide association study (GWAS) has emerged as a powerful and successful tool to identify common disease alleles by using high-throughput genotyping technology. It interrogates a large number of tagging single nucleotide polymorphisms (SNPs) that serve as surrogates for untested common SNPs across the genome. However, some true associations might not be detected by GWAS without accounting for environmental risk factors, because some susceptibility loci might act in an environment-responsive manner [Garcia-Closas et al., 2005; Kilpelainen et al., 2011; Wu et al., 2011].

*Correspondence to: Dongxin Lin, State Key Laboratory of Molecular Oncology, Cancer Institute and Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China and Peter Kraft, Program in Molecular and Genetic Epidemiology, Harvard School of Public Health, Boston, Massachusetts, USA. E-mail: lindx72@cicams.ac.cn; pkraft@hsph.harvard.edu

Many statistical methods have been proposed for investigation of statistical gene–environment (G × E) interactions in the context of case-control GWAS, and the relative effectiveness of these methods remains an area of active investigation [Cornelis et al., 2012; Hsu et al., 2012; Khoury and Wacholder, 2009; Kraft et al., 2007; Mukherjee et al., 2012; Murcray et al., 2011, 2009; Thomas et al., 2012]. Throughout this paper, we define "gene–environment interaction" as a departure from a multiplicative odds ratio model for the joint effect of genotype and exposure, that is, $OR(G,E) \neq OR(G)\ OR(E)$. We focus on these so-called multiplicative interactions because they are widely studied but note that other definitions of statistical G × E interaction exist (and depending on context may be more relevant), and we note that definitions of statistical and biological interactions are quite distinct [Kraft and Hunter, 2010; Siemiatycki and Thomas, 1981].

The standard test for G × E interaction, based on the coefficient of the product of the genetic and environmental

exposures in a logistic regression, remains a popular method to analyze case-control data, but is known to have low power [Hein et al., 2008; Hunter, 2005]. The case-only approach (GE-CO) has been proposed as a potentially more powerful method. This test leverages the fact that under the assumption that the genetic and environmental factors are independent in the general population, a G × E interaction will induce an association between genotype and exposure in the cases. In the simple case of a binary exposure, the interaction odds ratio, $OR_{GE} = OR(G,E)/[OR(G)\cdot OR(E)]$ is equivalent to the odds ratio from a logistic regression of exposure on genotype (holding all other risk factors constant). However, it is well known that violations of the G × E independence assumption can lead to increased Type I error rate for the case-only test [Albert et al., 2001; Mukherjee et al., 2012; Piegorsch et al., 1994; Satten and Epstein, 2004]. It is arguably less well known that the case-only test and other methods that leverage the G × E independence assumption can lose power when the G × E interaction effect and the correlation between genotype and exposure go in opposite directions. Several recent papers have illustrated this hypothetical situation through simulations [Li and Conti, 2009; Mukherjee et al., 2012]. Here we provide an empirical example of the phenomenon using a case-control study of esophageal squamous-cell carcinoma (ESCC). In particular, we give a practical example where the case-only test fails to identify makers associated with ESCC risk and participating in G × E interactions although other methods could identify these markers.

A number of methods have been proposed that seek to leverage the power boost of the case-only test while retaining control of the Type I error rate when the G × E independence assumption is violated [Hsu et al., 2012; Li and Conti, 2009; Mukherjee et al., 2008; Murcray et al., 2011, 2009]. Some of these methods have been shown to lose power relative to the standard case-control analysis when the G × E odds ratio and G × E association go in opposite directions [Mukherjee et al., 2012]. We also illustrate the performance of these methods in this situation using the ESCC data.

ESCC ranks as the tenth most prevalent cancer in the world, with marked regional variation and a particularly high incidence in regions of China. Molecular epidemiological studies using the candidate gene approach have established that a set of genetic variations primarily related to alcohol metabolism confer susceptibility to ESCC [Brooks et al., 2009; Cui et al., 2009; Hashibe et al., 2008; Hiyama et al., 2007]. We previously conducted a GWAS of ESCC and detected two previously identified regions in the alcohol metabolism pathway that showed strong evidence for G × E interaction: 12q24 harboring *ALDH2* and 4q23 harboring a cluster of seven genes encoding alcohol dehydrogenase (ADH) family (5′*ADH7-ADH1C-ADH1B-ADH1A-ADH6-ADH4-ADH5*-3′) [Wu et al., 2011, 2012]. ADHs oxidize alcohol to acetaldehyde, a likely important carcinogen in the etiology of alcohol-related cancers. *ALDH2* encodes aldehyde dehydrogenase-2, which detoxifies acetaldehyde to acetate. Individuals who carry an *ALDH2*2* allele, which slows this detoxification process, typically experience an unpleasant flushing reaction to alcohol due to the buildup of acetaldehyde, and are less likely to drink regularly or heavily [Brooks et al., 2009]. Thus, it is biologically plausible that alleles in *ADH* and *ALDH2* might be associated with an increase in the effect of alcohol intake on the risk of ESCC and, at the same time, be associated with decreased alcohol intake—consistent with the pattern previous studies have observed [Cui et al., 2009; Lewis and Smith 2005].

In this report, we illustrate the effect of analytic strategy on the ability to identify these two loci known to be involved in a G × E interaction.

Methods

Genome-Wide G × E Interaction Studies of ESCC

Details of the GWAS analyzed here have been reported in previous papers [Wu et al., 2011, 2012]. Briefly, 2,031 ESCC cases and 2,044 controls were recruited from the Han Chinese population in Beijing region, China. Demographic characteristics including age, sex, smoking status, and drinking status were obtained from patient's medical records. Control subjects were selected from those undergoing physical examination at primary care clinics in the Beijing region and frequency matched for age and sex to ESCC cases. For this study, alcohol drinking status was assessed by a detailed questionnaire. For the present analysis, individuals were classified as drinkers if they reported drinking any form of alcohol at least twice a week; otherwise, they were defined as nondrinkers. Individuals who reported smoking more than 100 cigarettes in their life or smoking tobacco in a pipe more than 100 times were defined as smokers; all others were defined as nonsmokers. At recruitment, informed consent was obtained from each subject, and the study was approved by the institutional review boards of the Chinese Academy of Medical Sciences Cancer Institute, Peking University.

Statistical Tests

We compared six statistical tests to identify G × E interactions in case-control data (Table 1). Five of the tests use

Table 1. Interaction tests considered in this report

	Explicitly considers gene–environment correlation?	Robust to population-level gene–environment correlation?[a]	Reference
Standard case-control test	N	Y	[Cornelis et al., 2012; Mukherjee et al., 2012] and many others
Case-only test	Y	N	[Piegorsch et al., 1994]
Empirical Bayes test	Y	Y	[Mukherjee et al., 2008]
Hybrid two-step approach	Y	Y	[Murcray et al., 2011]
Cocktail 1	Y	Y	[Hsu et al., 2012]
Cocktail 2	Y	Y	[Hsu et al., 2012]

[a] Assuming genotype and exposure are measured without error. For the effects of measurement error, see [Garcia-Closas et al., 1998; Lindstrom et al., 2009].

G × E correlation to increase power to detect G × E interactions, and four of these are designed to control the Type I error rate when genotype and exposure are correlated in the general population.

We applied standard logistic regression and case-only tests to markers in regions on 4q23 and 12q24, spanning the *ADH* gene cluster and *ALDH2*, respectively. For these analyses, we assumed a log-additive mode of inheritance, that is, we coded genotype as a count of minor alleles. These analyses also adjusted for smoking, age, and gender. Additional analyses adjusting for the top three principal components of genetic variation did not appreciably change these results [Wu et al., 2011, 2012]. For additional analyses of rs11066015 and rs3805322, the markers at 12q24 and 4q23 that showed strongest evidence for G × E interaction from the standard logistic regression test, we used a dominant coding for the nonreference allele (A and G, respectively) when applying all the methods listed in Table 1. We did this for two reasons. First, rs11066015 is in strong linkage equilibrium (r^2 = 0.71 in the Han Chinese in Beijing, China + Japanese in Tokyo, Japan panel of the 1000 Genomes Project Pilot 1 data [http://www.broadinstitute.org/mpg/snap/ldsearchpw.php, accessed November 1, 2012]) with the *ALDH2*2* (rs671) allele, which is known to act in a dominant fashion on the conversion of acetaldehyde to acetate [Brooks et al., 2009]. Second, we are using a formulation of the empirical Bayes test that assumes the genotype and exposure have a binary coding [Mukherjee et al., 2008]; to be consistent, we adopted this coding for the other tests. The hybrid two-step and cocktail methods both involve a screening step where only the markers that show some evidence for marginal association with disease and/or G × E correlation in the full sample are tested for G × E interaction. These methods require the user to specify *P*-value thresholds to use in the screening step. We used $\alpha_A = \alpha_M = 0.0001$ and $\rho = 0.5$ for the hybrid test and a screening *P*-value of 0.001 for the cocktail methods (with c = 0.001 for Cocktail 1), as recommended in previous publications [Hsu et al., 2012; Murcray et al., 2011].

In all analyses the E term, drinking status, was modeled as a binary trait: drinker and nondrinker. (This binary coding was chosen in part to guard against inflated Type I error in tests of interaction when the main effect of environment is misspecified [Cornelis et al., 2012; Tchetgen Tchetgen and Kraft, 2011].) Analyses were conducted by using PLINK, SAS, version 9.3, and R.

Results

Characteristics of 2,032 cases and 2,044 controls have been presented in previous papers [Wu et al., 2011]. Drinking status (being a drinker) was significantly associated with risk of ESCC with an OR (95% CI) of 1.63 (1.44–1.85). Quantile-quantile plots did not suggest any large-scale systemic bias due to population stratification or differential genotyping error for the marginal or case-only tests. Using marginal logistic regression adjusted by sex, age, smoking status, and drinking status, three SNPs (rs11066015, rs11066280, and rs2074356) on 12q24 and eight SNPs (rs1042026, rs3805322, rs17033, rs17028973, rs1614972, rs1229977, rs1789903, and rs1893883) on 4q23were identified associated with risk of ESCC with *P* values from 1.23×10^{-5} and 5.07×10^{-12} [Wu et al., 2011, 2012].

We applied the standard case-control and case-only tests for G × E interaction to 340 SNPs in a 2 Mb window around rs11066015 at 12q24 and 304 SNPs in a 2 Mb window around rs1042026 at 4q23.Using the standard case-control test, three SNPs (rs11066015, rs11066280 and rs2074356) on 12q24 showed highly statistically significantly interactions with alcohol drinking to promote ESCC risk ($P_{GE} < 10^{-16}$). Two SNPs, rs1042026 and rs3805322, on 4q23 also showed significant interactions, although P_{GE} values did not reach genome-wide significance levels ($P = 0.0052$ and 0.0002, respectively).

In contrast, the case-only test did not identify any genome-wide significant associations between SNPs at 12q24 and ESCC ($P > 0.0003$), whereas *P* values for the case-only test of G × E interaction for rs2074356 and rs3805322 at 4q23 were much smaller than those from the standard case-control test of interaction ($P_{GE\text{-}CO} = 4.68 \times 10^{-6}$ and 6.78×10^{-7}, respectively). To further illustrate these discrepancies, we present detailed results on rs11066015 at 12q24 and rs3805322 at 4q23 (Table 2).

The case-only estimate of the G × E interaction odds ratio for carriers of the A allele at rs11066015 was 1.06 (95% CI: 0.88–1.26, $P = 0.55$), as compared to 3.71 (2.84–4.86, $P < 10^{-16}$) from the standard G × E test. Nondrinkers who carried the minor allele of rs11066015 had a nonsignificant 0.84-fold decrease in odds of ESCC relative to noncarriers, whereas drinkers who carried the minor allele had a 3.12-fold

Table 2. Association between rs11066015 (12q24) and rs3805322 (4q23) and esophageal squamous-cell carcinoma (ESCC), stratified by drinking status

	12q24: rs11066015 (A)		
	Noncarrier	Carrier	
	Nondrinkers		
Control	615	523	
Case	513	367	
	$OR_{G\|D=0}$	0.84 (0.70, 1.00)	$P = 0.06$
	Drinkers		
Control	728	176	
Case	650	491	
	$OR_{G\|D=1}$	3.12 (2.55, 3.82)	$P < 1 \times 10^{-16}$
	$OR_{G\times E\text{-standard}}$	3.71 (2.84, 4.86)	$P < 1 \times 10^{-16}$
	$OR_{G\times E\text{-case-only}}$	1.06 (0.88,1.26)	0.55
	4q23: rs3805322 (G)		
	Noncarrier	Carrier	
	Nondrinkers		
Control	847	291	
Case	620	260	
	$OR_{G\|D=0}$	1.22 (1.00, 1.49)	$P = 0.05$
	Drinkers		
Control	652	252	
Case	652	489	
	$OR_{G\|D=1}$	1.94 (1.61, 2.34)	$P < 3.4 \times 10^{-12}$
	$OR_{G\times E\text{-standard}}$	1.59 (1.21, 2.09)	$P < 8.1 \times 10^{-4}$
	$OR_{G\times E\text{-case-only}}$	1.79 (1.48, 2.15)	$P = 9.6 \times 10^{-10}$

increase in odds (Table 2). At the same time, controls who carried the minor allele were less likely to drink: the odds of a carrier being a drinker were 0.28 times that of noncarriers (95% CI 0.23–0.35; $P < 10^{-16}$). (The association between rs11066015 and alcohol intake did not change appreciably after adjusting for age, gender, smoking status, or the top three principal components of genetic variation [Wu et al., 2012].) This is consistent with previous cross-sectional studies in east Asian populations that showed carriers of the minor allele of *ALDH2*2* (rs671) are more likely to experience a flushing reaction to alcohol and less likely to drink regularly [Brooks et al., 2009].

On the other hand, the case-only estimate of the G × E interaction odds ratio for carriers of the G allele at rs3805322 was 1.79 (1.48–2.15; $P = 6.9 \times 10^{-10}$), compared to 1.59 (1.21–2.09; $P = 8.1 \times 10^{-4}$) for the standard case-control analysis. There was no evidence of association between rs3805322 and drinking status among controls, consistent with previous population-based studies of the *ADH* gene cluster [Cui et al., 2009; Hashibe et al., 2008].This suggests that the discrepancy between the case-only and standard tests for rs11066015 is due to a negative correlation between the deleterious exposure and the risk allele, consistent with previous theoretical results.

Table 3 shows how this pattern of G × E correlation affected other tests that use G × E correlation when testing for G × E interaction. All of the recently proposed tests that use information from both the case-only and the standard case-control test were able to detect the interaction on chromosome 12. For the chromosome 4 interaction, the case-only test yielded the strongest evidence for interaction ($P = 9.6 \times 10^{-10}$). The standard case-control and empirical Bayes tests failed to achieve genome-wide significance ($P = 8.1 \times 10^{-4}$ and $P = 5.4 \times 10^{-5}$, respectively). Because it uses the standard case-control test at the second stage, the hybrid test would fail to identify this interaction in a genome-wide screen. Both cocktail methods use the empirical Bayes test at the second stage for this locus, which is significant after accounting for the number of markers that make it through the initial screening steps.

Table 3. Genome-wide significance of tests for gene–environment interaction for rs11066015 (12q24) and rs3805322 (4q23)

	Genome-wide significant? ($\alpha = 5 \times 10^{-8}$)	
	rs11066015[a]	rs3805322[b]
Standard case-control test	Yes	No
Case-only test	No	Yes
Empirical Bayes test	Yes	No
Hybrid two-step approach	Yes	No
Cocktail 1	Yes	Yes
Cocktail 2	Yes	Yes

[a] Empirical Bayes estimate of $OR_{G\times E} = 3.66$ (2.79,4.80); for the screening stage of the hybrid test, both G × E association and marginal tests were significant with $P_A = 6.0 \times 10^{-14} < \alpha_A$ and $P_M = 7.3 \times 10^{-8} < \alpha_M$, and the standard test of G × E interaction at the second stage was quite significant ($P < 10^{-16}$); for the cocktail methods, $P^{screen} = P_M$ for cocktail 1 and $P^{screen} = P_A$ for cocktail 2, both of these pass the first stage threshold, and the second stage tests (the empirical Bayes test for cocktail 1 and standard case-control test for cocktail 2) are both very significant ($P < 10^{-16}$).

[b] Empirical Bayes estimate of $OR_{G\times E} = 1.70$ (1.36, 2.20), $P = 5.4 \times 10^{-5}$; for the screening stage of the hybrid test, both G × E association and marginal tests were significant with $P_A = 1.1 \times 10^{-9} < \alpha_A$ and $P_M = 9.3 \times 10^{-13} < \alpha_M$, however, the standard test of G × E interaction at the second stage did not meet the second stage threshold ($\approx 4.2 \times 10^{-4}$); for the cocktail methods, $P^{screen} = P_M$ for cocktail 1 and 2, which passes the first stage threshold, and the second stage test (the empirical Bayes test for both) meets the second stage threshold ($\approx 4.2 \times 10^{-4}$).

Discussion

Many methods have been developed to investigate G × E interactions in GWAS, but none of these approaches have been shown to be consistently most effective [Mukherjee et al., 2008; Murcray et al., 2011]. Although the standard logistic regression test is widely used, the power of this method is limited. Taking our precious GWAS of ESCC as an example, two regions, 12q24 and 4q23, were identified to be marginally associated with risk of ESCC. In an expanded sample (10,123 cases and 10,664 controls compared to 2,031 cases and 2,044 controls in the original GWAS), both showed genome-wide significant interactions with drinking status. However, the standard G × E interaction test failed to identify the interaction at 4q23 at genome-wide significance levels in the original GWAS. This suggests that other more powerful methods might identify interactions missed by the standard method.

The case-only test has been proposed as a powerful approach to exploit G × E independence. Although it is widely appreciated that the case-only method can have increased Type I error in the presence of G × E correlation, it is perhaps less widely appreciated that this method can have lower power when the G × E interaction odds ratio and G × E correlation are in opposite directions. This study gives a practical example of this phenomenon. We also show that other, newer methods that leverage the G × E independence association but retain control of the Type I error retain sufficient power to detect this strong interaction at a genome-wide significance level.

In our study, the case-only method failed to identify three SNPs at 12q24 that are involved in a strong G × E interaction. The minor alleles of these SNPs at the 12q24 region are in linkage disequilibrium with the *ALDH2*2* allele, which increases the risk of ESCC because the carriers have a decreased rate of detoxifying acetaldehyde to acetate [Hashibe et al., 2008; Lewis and Smith, 2005; McKay et al., 2011]. However, nondrinkers carry a remarkably higher frequency of this risk allele than the drinkers among controls included in this study. The individuals with this risk allele are unable to degrade aldehyde efficiently and tend to develop malaise, flushing reaction and other uncomfortable symptoms when drinking alcohol. These conditions make individuals with inactive *ALDH2* less likely to consume alcohol [Li et al., 2006]. Therefore, the G × E correlation and interaction effects for SNPs in *ALDH2* region on 12q24 are in the opposite direction and decrease the power of case-only test.

On the other hand, the case-only test identified two SNPs on 4q23 with genome-wide significant G × E interactions, whereas the *P* values from the standard G × E test and

empirical Bayes test were five orders of magnitude larger. There was no compelling evidence in this study or previous studies that alleles in this region are associated with alcohol consumption. These results are consistent with the theoretical prediction that the case-only test will be much more powerful than the standard test when the G × E independence assumption holds.

Of the two-stage approaches, the cocktail methods were able to identify the interaction at 4q23, whereas the hybrid test was not. In certain situations (specifically, when the screening *P*-value [P^{screen}] is equal to the *P*-value from the marginal test of gene-disease association [P_M]), the cocktail method uses the empirical Bayes test at the second stage, rather than the standard logistic regression test (which the hybrid test always uses). Because the empirical Bayes test leverages evidence for G × E independence, it had smaller *P* values than the standard interaction test at this locus.

We have given an example involving a polymorphism (*ALDH2*2*) that has a very strong and direct association with a known environmental risk factor for disease (alcohol intake). The strength of this association is arguably exceptionally large—large GWAS of behavioral "exposures" such as alcohol, caffeine, and tobacco cigarette intake have not identified associations with similarly strong associations [Tobacco and Genetics Consortium, 2010; Bierut et al., 2010; Cornelis et al., 2011; Liu et al., 2010; Thorgeirsson et al., 2010]—but in very large sample sizes (such as those needed to reliably identify modest interaction effect), smaller G × E interaction associations may produce the biases seen here. Of particular note, differences in exposure frequencies (including differences in unmeasured confounders) that are correlated with genetic ancestry may lead to pervasive G × E correlation, and appropriate adjustment or control for population stratification should be applied. This may be of special concern in admixed populations, such as African Americans or Latinos.

In summary, our empirical results are consistent with previous theoretical studies and suggest that multiple analytic approaches should be used when screening for G × E interactions, as no one method is universally powerful. Different approaches will be sensitive to different alternative hypotheses and G × E correlation patterns. Because they combine the strengths of several approaches while maintaining appropriate control of Type I error rate, recently developed two-step tests for G × E interaction like the hybrid method of Murcray et al. [2011] and the cocktail methods of Hsu et al. [2012] may be broadly useful. In general, the choice of analytic approach(es) will depend on the primary research question and the context of the setting [Cordell, 2009; Kraft and Hunter, 2010; Thomas, 2010]. For example, if the study aims to characterize the association between a particular marker and exposure and disease risk, then a method that provides unbiased and precise parameter estimates will be preferred to methods that provide an interaction test but do not provide estimates of all parameters of interest (which will typically include both main effect and interaction parameters). Several recent papers have considered the performance of many tests for G × E interaction across a wide range of hypothetical scenarios (strength of main and interaction effects, presence and size of G × E correlations) [Hsu et al., 2012; Li and Conti, 2009; Mukherjee et al., 2008; Murcray et al., 2011, 2009]. It remains an open question, which of these hypothetical situations are most relevant in practice? We have provided an empirical, cautionary example of one of the more interesting hypothetical situations.

Acknowledgments

This work was funded by the National High-Tech Research and Development Program of China (2009AA022706 to D.L.), the National Basic Research Program of China (2011CB504303 to D.L. and W.T.), the National Natural Science Foundation of China (30721001 to D.L., Q.Z., and Z.L.), and U.S. National Institutes of Health: U19 CA148065-01 and R21 CA165920-01. The authors have no conflict of interest to declare.

References

Tobacco and Genetics Consortium. 2010. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat Genet* 42(5):441–447.

Albert PS, Ratnasinghe D, Tangrea J, Wacholder S. 2001. Limitations of the case-only design for identifying gene-environment interactions. *Am J Epidemiol* 154(8):687–693.

Bierut LJ, Agrawal A, Bucholz KK, Doheny KF, Laurie C, Pugh E, Fisher S, Fox L, Howells W, Bertelsen S, and others. 2010. A genome-wide association study of alcohol dependence. *Proc Natl Acad Sci USA* 107(11):5082–5087.

Brooks PJ, Enoch MA, Goldman D, Li TK, Yokoyama A. 2009. The alcohol flushing response: an unrecognized risk factor for esophageal cancer from alcohol consumption. *PLoS Med* 6(3):e50.

Cordell HJ. 2009. Detecting gene-gene interactions that underlie human diseases. *Nat Rev Genet* 10(6):392–404.

Cornelis MC, Monda KL, Yu K, Paynter N, Azzato EM, Bennett SN, Berndt SI, Boerwinkle E, Chanock S, Chatterjee N, and others. 2011. Genome-wide meta-analysis identifies regions on 7p21 (AHR) and 15q24 (CYP1A2) as determinants of habitual caffeine consumption. *PLoS Genet* 7(4):e1002033.

Cornelis MC, Tchetgen EJ, Liang L, Qi L, Chatterjee N, Hu FB, Kraft P. 2012. Gene-environment interactions in genome-wide association studies: a comparative study of tests applied to empirical studies of type 2 diabetes. *Am J Epidemiol* 175(3):191–202.

Cui R, Kamatani Y, Takahashi A, Usami M, Hosono N, Kawaguchi T, Tsunoda T, Kamatani N, Kubo M, Nakamura Y, and others. 2009. Functional variants in ADH1B and ALDH2 coupled with alcohol and smoking synergistically enhance esophageal cancer risk. *Gastroenterology* 137(5):1768–1775.

Garcia-Closas M, Thompson WD, Robins JM. 1998. Differential misclassification and the assessment of gene-environment interactions in case-control studies. *Am J Epidemiol* 147(5):426–433.

Garcia-Closas M, Malats N, Silverman D, Dosemeci M, Kogevinas M, Hein DW, Tardon A, Serra C, Carrato A, Garcia-Closas R, and others. 2005. NAT2 slow acetylation, GSTM1 null genotype, and risk of bladder cancer: results from the Spanish Bladder Cancer Study and meta-analyses. *Lancet* 366(9486):649–659.

Hashibe M, McKay JD, Curado MP, Oliveira JC, Koifman S, Koifman R, Zaridze D, Shangina O, Wunsch-Filho V, Eluf-Neto J, and others. 2008. Multiple ADH genes are associated with upper aerodigestive cancers. *Nat Genet* 40(6):707–709.

Hein R, Beckmann L, Chang-Claude J. 2008. Sample size requirements for indirect association studies of gene-environment interactions (G x E). *Genet Epidemiol* 32(3):235–245.

Hiyama T, Yoshihara M, Tanaka S, Chayama K. 2007. Genetic polymorphisms and esophageal cancer risk. *Int J Cancer* 121(8):1643–1658.

Hsu L, Jiao S, Dai JY, Hutter C, Peters U, Kooperberg C. 2012. Powerful cocktail methods for detecting genome-wide gene-environment interaction. *Genet Epidemiol* 36(3):183–194.

Hunter DJ. 2005. Gene-environment interactions in human diseases. *Nat Rev Genet* 6(4):287–298.

Khoury MJ, Wacholder S. 2009. Invited commentary: from genome-wide association studies to gene-environment-wide interaction studies—challenges and opportunities. *Am J Epidemiol* 169(2):227–230; discussion 234–5.

Kilpelainen TO, Qi L, Brage S, Sharp SJ, Sonestedt E, Demerath E, Ahmad T, Mora S, Kaakinen M, Sandholt CH, and others. 2011. Physical activity attenuates the influence of FTO variants on obesity risk: a meta-analysis of 218,166 adults and 19,268 children. *PLoS Med* 8(11):e1001116.

Kraft P, Hunter D. 2010. The challenge of assessing complex gene–gene and gene–environment interactions. In: Khoury MJ, Bedrosian S, Gwinn M, Higgins JPT, Ioannidis JPA, Little J, editors, *Human Genome Epidemiology: Building the Evidence for Using Genetic Information to Improve Health and Prevent Disease* (2nd ed.). New York: Oxford University Press p 165–187.

Kraft P, Yen YC, Stram DO, Morrison J, Gauderman WJ. 2007. Exploiting gene-environment interaction to detect genetic associations. *Hum Hered* 63(2): 111–119.

Lewis SJ, Smith GD. 2005. Alcohol, ALDH2, and esophageal cancer: a meta-analysis which illustrates the potentials and limitations of a Mendelian randomization approach. *Cancer Epidemiol Biomarkers Prev* 14(8):1967–1971.

Li D, Conti DV. 2009. Detecting gene-environment interactions using a combined case-only and case-control approach. *Am J Epidemiol* 169(4):497–504.

Li Y, Zhang D, Jin W, Shao C, Yan P, Xu C, Sheng H, Liu Y, Yu J, Xie Y, and others. 2006. Mitochondrial aldehyde dehydrogenase-2 (ALDH2) Glu504Lys polymorphism contributes to the variation in efficacy of sublingual nitroglycerin. *J Clin Invest* 116(2):506–511.

Lindstrom S, Yen YC, Spiegelman D, Kraft P. 2009. The impact of gene-environment dependence and misclassification in genetic association studies incorporating gene-environment interactions. *Hum Hered* 68(3):171–181.

Liu JZ, Tozzi F, Waterworth DM, Pillai SG, Muglia P, Middleton L, Berrettini W, Knouff CW, Yuan X, Waeber G, and others. 2010. Meta-analysis and imputation refines the association of 15q25 with smoking quantity. *Nat Genet* 42(5):436–440.

McKay JD, Truong T, Gaborieau V, Chabrier A, Chuang SC, Byrnes G, Zaridze D, Shangina O, Szeszenia-Dabrowska N, Lissowska J, and others. 2011. A genome-wide association study of upper aerodigestive tract cancers conducted within the INHANCE consortium. *PLoS Genet* 7(3):e1001333.

Mukherjee B, Ahn J, Gruber SB, Rennert G, Moreno V, Chatterjee N. 2008. Tests for gene-environment interaction from case-control data: a novel study of type I error, power and designs. *Genet Epidemiol* 32(7):615–626.

Mukherjee B, Ahn J, Gruber SB, Chatterjee N. 2012. Testing gene-environment interaction in large-scale case-control association studies: possible choices and comparisons. *Am J Epidemiol* 175(3):177–190.

Murcray CE, Lewinger JP, Gauderman WJ. 2009. Gene-environment interaction in genome-wide association studies. *Am J Epidemiol* 169(2):219–226.

Murcray CE, Lewinger JP, Conti DV, Thomas DC, Gauderman WJ. 2011. Sample size requirements to detect gene-environment interactions in genome-wide association studies. *Genet Epidemiol* 35(3):201–210.

Piegorsch WW, Weinberg CR, Taylor JA. 1994. Non-hierarchical logistic models and case-only designs for assessing susceptibility in population-based case-control studies. *Stat Med* 13(2):153–162.

Satten GA, Epstein MP. 2004. Comparison of prospective and retrospective methods for haplotype inference in case-control studies. *Genet Epidemiol* 27(3): 192–201.

Siemiatycki J, Thomas D. 1981. Biological models and statistical interactions: an example from multistage carcinogenesis. *Int J Epidemiol* 10:383–397.

Tchetgen Tchetgen EJ, Kraft P. 2011. On the robustness of tests of genetic associations incorporating gene-environment interaction when the environmental exposure is misspecified. *Epidemiology* 22:257–261.

Thomas D. 2010. Gene–environment-wide association studies: emerging approaches. *Nat Rev Genet* 11(4):259–272.

Thomas DC, Lewinger JP, Murcray CE, Gauderman WJ. 2012. Invited commentary: GE-Whiz! Ratcheting gene-environment studies up to the whole genome and the whole exposome. *Am J Epidemiol* 175(3):203–207; discussion 208–209.

Thorgeirsson TE, Gudbjartsson DF, Surakka I, Vink JM, Amin N, Geller F, Sulem P, Rafnar T, Esko T, Walter S, and others. 2010. Sequence variants at CHRNB3-CHRNA6 and CYP2A6 affect smoking behavior. *Nat Genet* 42(5): 448–453.

Wu C, Hu Z, He Z, Jia W, Wang F, Zhou Y, Liu Z, Zhan Q, Liu Y, Yu D, and others. 2011. Genome-wide association study identifies three new susceptibility loci for esophageal squamous-cell carcinoma in Chinese populations. *Nat Genet* 43(7):679–684.

Wu C, Kraft P, Zhai K, Chang J, Wang Z, Li Y, Hu Z, He Z, Jia W, Abnet CC, and others. 2012. Genome-wide association analyses of esophageal squamous cell carcinoma in Chinese identify multiple susceptibility loci and gene-environment interactions. *Nat Genet* 44(10):1090–1097.