

Commentary: Thoughts on assessing evidence for gene by environment interaction

Clarice R Weinberg

MD A3-03, National Institute of Environmental Health Sciences, PO Box 12233, Research Triangle Park, NC 27709, USA.
E-mail: weinber2@niehs.nih.gov

Accepted 2 March 2012

The commentary in this issue by Boffetta *et al.*¹ addresses a challenging problem: given that a genetic factor, G, and an environmental factor, E, might both influence risk of cancer, how can we systematically assess the accumulated published evidence for the existence of interaction between them? Most complex diseases, not just cancer, probably arise through a web of factors that are both genetic and environmental, so this question arises as one of central relevance to the epidemiological study of aetiology.

Boffetta *et al.* begin with a comprehensive review of the literature on approaches and sources of bias for assessments of interaction, giving an overview that should be highly useful to readers. They then suggest a scoring system to combine prior biological knowledge related to plausibility, for example knowledge that both the exposure and the genetic factor are involved in the same pathway, with available evidence for main effects and for interaction. Whereas the authors deserve credit for taking on a daunting problem and paying due attention to the biology, I have some issues with the proposed enterprise. I will start by offering some supplemental points and end with a bigger-picture conceptual issue.

First, the recommendations provided were represented as specific to cancer, but the same overall strategies should apply to other complex phenotypes as well. Given that readers will want to apply the proposed principles more broadly, it seems worthwhile to broaden our consideration of them. The authors mention some designs other than the case–control design, including the case–parent trio design, which they say is ‘rarely used’ but can be used for testing. It is true that case–parent designs are not typically appropriate in cancer research, because most diagnoses occur in later life. Nevertheless, the case–parent triad design has been extensively used in studies of non-cancerous young-onset conditions. An ongoing ‘Two Sister Study’ uses a family design for young-onset breast cancer. Whereas a case–parent approach cannot assess main effects of exposures (unless unaffected siblings are also studied), the design is in some ways ideal for the

study of birth defects, pregnancy complications and phenotypes such as asthma, schizophrenia and autism, which are diagnosed at a young age. As its inference relies on transmissions of alleles from parents to offspring, under a case–parent design relative risks associated with inherited genotypes are estimable without concern for genetic population stratification. However, this is not its major selling point, as good statistical methods for genomic control are available for case–control studies (provided ancestry-informative markers are available). Family designs also resist bias due to self-selection of cases and controls (no population-based controls are needed). They also resist confounding bias due to prenatal experiences that are caused or modified by the maternal genome, a mechanism that can be particularly important for young-onset conditions and can distort inference based on case–control designs, including inference related to G×E interaction. With data from nuclear families, one can also identify and study causative genes that are subject to parent-of-origin effects, a phenomenon that cannot be probed using a case–control design. Such a gene was recently reported for breast cancer² based on Icelandic family data.

A word of caution is needed. Whereas family-based designs offer some advantages over a case–control design for studying conditions with onset early in life, we have recently shown^{3,4} that, although for genotype main effects they fully resist bias due to population stratification, neither case–control designs nor designs based on nuclear families are fully robust for assessments of G×E interaction. Bias can arise if the population stratification involves the exposure. The problem is that exposures can track with genes because cultural practices tend to be passed on along with genes. What can happen with assessments of G×E is that, e.g. in a genome-wide interaction study, one is typically not directly examining a causative single nucleotide polymorphism (SNP), but a marker SNP that is related to risk only through linkage disequilibrium (LD) with an untyped causative locus. If there are incompletely mixed genetic subpopulations,

and haplotype prevalences and exposure prevalences both vary across those subpopulations, then the exposure can act as a surrogate for the degree of LD between the SNP that was typed and a causative SNP, producing spurious evidence for (or suppression of) multiplicative interaction. Thus, control for population stratification, or inclusion of an unaffected sibling⁴ can be important when assessing interaction.

One problem that must be confronted in case-control approaches is that valid tests of GxE interaction (regardless of the choice of null model) require that the main effects of both G and E be correctly specified. G effects can be saturated simply through assuming a co-dominant model. In contrast, if E is a continuous exposure, getting the E main effects model right is challenging. With a case-parents design, this issue is partially bypassed by modelling the transmission of the variant allele as possibly depending on E under GxE alternatives.^{5,6} This approach permits valid inference for a test of multiplicative GxE without the need to correctly specify the E main effects. Categorization of E provides another practical solution, which should work for a case-control design or a case-parent design,⁶ but the choice of cut-points becomes somewhat arbitrary, as discussed by Boffetta *et al.*¹

Boffetta *et al.* quite sensibly, in my opinion, propose that a marginal main effect of G and main effect of E will likely be discernable if the two interact. Although one can invent examples where neither has a marginal effect on risk, but nevertheless the two interact, such examples are highly artificial. The exposure would have to be beneficial in those with a particular genotype and detrimental in those with a different genotype. Consequently, it seems sensible to include assessments of main effects as important to establishing plausibility that two factors interact.

Consider the evidence for a main effect of G. One must remember that a genome-wide association study (GWAS) with a million SNPs may still not have typed the causative SNP; consequently, associations will often be attenuated because we are typing more or less weak surrogates for the aetiologically relevant genotype. Moreover, the LD that GWASs rely on to produce evidence for a main effect for G may be different in different populations, and flip-flops can even occur, where the direction of effect switches due to different ancestry in the different populations.⁷ Thus, a finding of heterogeneity in the apparent G effect across studies should not necessarily be taken as weakening the evidence for a genotype's main effect. Another related point is that several SNPs in loci that are in high LD may simultaneously need to be present (e.g. if the true effect depends on a haplotype). Similarly, there can be 'epistasis': several unlinked variant SNPs may be jointly needed to disable a causally relevant pathway. Such multi-SNP effects are subject to huge attenuation under SNP-by-SNP analysis.⁸

Finally, interaction itself, as a phenomenon that might either exist or not exist, is not a well-defined

concept. Do E and G interact? Mathematically, we already know the answer. If they both are risk factors under some coding, then the answer must be yes; for some selected no-interaction null hypothesis, they statistically must 'interact'. If their true joint effect is multiplicative then they interact compared with an additive no-interaction null. If their true joint effect is additive, then they interact compared with a multiplicative no-interaction null. If their joint effect is something else entirely, neither additive nor multiplicative, then they interact compared with both nulls. So, in the event that both main effects are present, if we make the existence of interaction our primary question, we either agree that there is always some kind of interaction, or we condemn ourselves to an ongoing round of arguments over which of the possible null-interaction models is the right one. I^{9,10} and others have argued for additivity of hazards as the best no-interaction null model, because hazard additivity captures biological effects that are separate and probabilistically independent, which feels like what one should want biologic independence to mean. This is also how toxicologists have historically defined 'simple independent action' for the joint effect of toxic exposures.¹¹ But if two factors act at separate stages, e.g. in a carcinogenic cascade, then we know that biologic independence can alternatively produce a multiplicative joint effect.¹² In my view, the arguments over which is the best null model for no interaction are interesting but ultimately unresolvable. Perhaps the question itself, the one about whether G and E interact, needs to be reframed. Suppose evidence suggests that both G and E are risk factors. If our goal is to better understand how risk factors work together to increase risk, the real aetiological challenge is to devise and understand the best multi-predictor model for their combined effect.¹⁰ Having developed a revelatory model, we could then probe it for insights into the underlying biologic processes.¹³ This project is admittedly difficult. But ultimately, I think what will serve us better than a well-crafted scoring scheme is multi-predictor models that reflect both the published data and prior biological knowledge, and offer testable insights into the underlying pathogenic processes.

Conflict of interest: None declared.

References

- Boffetta P, Winn DM, Ioannidis JP *et al.* Recommendations and proposed guidelines for assessing the cumulative evidence on joint effects of genes and environments on cancer occurrence in humans. *Int J Epidemiol* 2012;**41**:686–704.
- Kong A, Steinthorsdottir V, Masson G *et al.* Parental origin of sequence variants associated with complex diseases. *Nature* 2009;**462**:868–74.

- ³ Shi M, Umbach DM, Weinberg CR. Family-based gene-by-environment interaction studies: revelations and remedies. *Epidemiology* 2011;**22**:400–07.
- ⁴ Weinberg CR, Shi M, Umbach DM. A sibling-augmented case-only design for assessing multiplicative gene-environment interaction. *Am J Epidemiol* 2011;**174**:1183–89.
- ⁵ Kistner EO, Shi M, Weinberg CR. Using cases and parents to study multiplicative gene-by-environment interaction. *Am J Epidemiol* 2009;**170**:393–400.
- ⁶ Umbach DM, Weinberg CR. The use of case-parent triads to study joint effects of genotype and exposure. *Am J Hum Genet* 2000;**66**:251–61.
- ⁷ Lin PL, Vance PM, Pericak-Vance MA, Martin ER. No gene is an island: the flip-flop phenomenon. *Am J Hum Genet* 2007;**80**:531–38.
- ⁸ Shi M, Weinberg CR. How much are we missing in SNP-by-SNP analyses of GWAS? *Epidemiology* 2011;**22**:845–47.
- ⁹ Weinberg CR. Applicability of the simple independent action model to epidemiologic studies involving two factors and a dichotomous outcome. *Am J Epidemiol* 1986;**123**:162–73.
- ¹⁰ Weinberg CR. Interaction and exposure modification: are we asking the right questions? *Am J Epidemiol* 2012; doi:10.1093/aje/kwr495.
- ¹¹ Finney DJ. *Probit Analysis*. 3rd edn. New York: Cambridge University Press, 1971.
- ¹² Siemiatycki J, Thomas D. Biological models and statistical interactions: an example from multistage carcinogenesis. *Int J Epidemiol* 1981;**10**:383–87.
- ¹³ Wacholder S, Han S, Weinberg C. Inference from a multiplicative model of joint genetic effects on ovarian cancer risk. *J Natl Cancer Inst* 2011;**103**:82–83.

Commentary: Reporting and assessing evidence for interaction: why, when and how?

David Clayton

Wellcome Trust and Juvenile Diabetes Research Foundation, Diabetes and Inflammation Laboratory, Department of Medical Genetics, University of Cambridge, UK. E-mail: dc208@cam.ac.uk

Accepted 28 March 2012

Thompson¹ noted in 1991 that, although more than a decade had passed since the first discussion of ‘interaction’ in the epidemiological literature, debate had by then subsided and few clear conclusions had emerged. He concluded:

Unfortunately, choice among theories of pathogenesis is enhanced hardly at all by the epidemiological assessment of interaction... What few causal systems can be rejected on the basis of observed results would provide decidedly limited etiologic insight.

This conclusion probably represented the consensus of opinion at that time. Yet 20 years later, discussion and reports of interactions pervade the literature. In this journal alone, two papers giving guidelines and recommendations for reporting and assessing interactions have appeared within weeks of one another. Knol and VanderWeele² considered only the reporting of interactions. Boffetta *et al.*³ primarily discuss the

assessment of evidence, but, since they are much concerned with the ‘cumulative’ evidence following meta-analysis or systematic review, they inevitably touch on the reporting of interactions.

One could point to several developments that have contributed to this reawakening of interest. The first of these has arisen from recent work by the statisticians in ‘causal modelling’, which has led to new, deterministic, definitions of mechanistic interaction. One such approach derives from the earlier work of Rothman⁴ on component causes, whereas the other is based on the currently popular focus on counterfactual outcomes. In the former approach, interaction between two causal factors is defined as the presence of a sufficient cause, which involves both factors as component causes, whereas from the latter viewpoint, interaction implies the existence of people in the population who would not have developed disease unless they had been exposed to both factors. These viewpoints have been shown to be mutually consistent^{5–7} and, subject to certain assumptions, are also