The study of gene–environment interactions in psychiatry: limited gains at a substantial cost?

S. Zammit[1,2]*, N. Wiles[2] and G. Lewis[2]

[1] *Department of Psychological Medicine, School of Medicine, Cardiff University, Cardiff, UK*
[2] *Academic Unit of Psychiatry, University of Bristol, Bristol, UK*

There is an ever-increasing body of literature examining gene–environment interactions in psychiatry, reflecting a widespread belief that such studies will aid identification of novel risk factors for disease, increase understanding about underlying pathological mechanisms, and aid identification of high-risk groups for targeted interventions. In this article we discuss to what extent studies of gene–environment interactions are likely to lead to any such benefits in the future.

Introduction

The processes underlying most psychiatric disorders are probably extremely complex. Studying how the combined effects of genes and environment impact upon risk of disease seems an intuitively attractive approach, and there is an ever-increasing body of literature examining gene–environment interactions (G × E) in psychiatry.

There are three main arguments put forward as to why studies of G × E may be helpful: (1) they may make it easier to identify novel genetic or environmental risk factors for disease, (2) they may increase our understanding about underlying pathological mechanisms of disease, and (3) they may aid identification of high-risk groups that might benefit from targeted interventions. In this article we discuss whether or not the study of G × E is likely to lead to any of these potential benefits. Although we focus mainly on gene–environment interactions, the arguments we present hold equally well for studies of gene–gene or environment–environment interactions.

Measuring interaction

Within epidemiology, the term interaction is used to describe the situation where the association between one exposure (risk factor) and disease varies according to the presence or absence of another exposure.

* Address for correspondence: Dr S. Zammit, Department of Psychological Medicine, School of Medicine, Cardiff University, Heath Park, Cardiff CF14 4XN, Wales, UK.

(Email: zammits@cardiff.ac.uk)

Where one exposure has an opposite effect on disease risk according to the presence or absence of another exposure (a *qualitative* interaction), then the discovery of such an interaction is important as this can indeed lead to all of the potential advantages described above. For example, it has been reported that, in the presence of high paternal antisocial personality (ASP) traits, the risk of child conduct problems increases the more time the father lives with the child, but with an opposite effect if paternal ASP traits are low, such that the risk of child conduct problems decreases the more time the father lives with the child (Jaffee *et al.* 2003; Blazei *et al.* 2008). Findings such as this can potentially have important implications for finding risk factors, understanding aetiology and identifying high-risk groups. However, such qualitative interactions have only rarely been observed in medicine and are not the main focus of the discussions that follow. Most interactions that have been described in the psychiatric literature are less extreme ones, and the potential advantages from identifying such *quantitative* rather than qualitative interactions are much less clear.

To study how exposure to two risk factors in combination affects disease risk, we compare data to predictions from statistical models. Statistical interaction occurs when the risk of disease if exposed to both factors A and B is different from that predicted by the statistical model being used. These predictions are usually modelled on either additive or multiplicative scales, and it is important to appreciate that interactions under these models mean different things. Statistical interaction is therefore model dependent and, as demonstrated below, it is meaningless to speak

Table 1. *Examples of statistical models used to study interactions*

Statistical model	Null hypothesis
Additive[a]	Risk (A and B) = Risk (A only) + Risk (B only) − Risk (neither A nor B)
Multiplicative	Risk ratio (A and B) = Risk ratio (A only) × Risk ratio (B only)

[a] Although beyond the scope of this article, methods have been developed to calculate a more valid measure of synergism (Darroch, 1997) that are also modelled on an additive scale.

Table 2. *Hypothetical 1-year cumulative incidence of schizophrenia (per 10 000) given an additive relationship between heavy cannabis use and COMT rs4680 valine allele on risk of schizophrenia*

	Risk of schizophrenia	
	No cannabis	Cannabis
Valine absent	1	10
Valine present	5	14

Risk difference (RD) for cannabis use is 9 where valine is absent (10 − 1), and is also 9 where valine is present (14 − 5); RDs equal ≡ no additive interaction. However, the risk ratio (RR) for cannabis use is 10 where valine is absent (10/1) but is 2.8 where valine is present (14/5); RRs unequal ≡ multiplicative interaction. Note that where RDs for cannabis are the same whether people have the valine allele or not, the RRs will be different.

Table 3. *Hypothetical 1-year cumulative incidence of schizophrenia (per 10 000) given a multiplicative relationship between heavy cannabis use and COMT rs4680 valine allele on risk of schizophrenia*

	Risk of schizophrenia	
	No cannabis	Cannabis
Valine absent	1	10
Valine present	5	50

Risk difference (RD) for cannabis use is 9 where valine is absent (10−1), but is 45 where valine is present (50−5); RDs unequal ≡ additive interaction. However, the risk ratio (RR) for cannabis use is 10 where valine is absent (10/1) and is also 10 where valine is present (50/5); RRs equal ≡ no multiplicative interaction. Note that where RRs for cannabis are the same whether people have the valine allele or not, the RDs will be different.

of interactions without specifying the statistical model to which they refer.

For example, in a study of risk differences or differences between means, the model is on an additive scale. The null hypothesis for statistical interaction in this model is that the joint effect of being exposed to both A and B is additive (Table 1). Departure from this (less than or greater than an additive effect) will provide evidence of statistical interaction. A hypothetical example of an additive relationship is given in Table 2.

However, if ratio measures (such as risk ratios or odds ratios) are used to study the association between exposures and disease, these are modelled on a multiplicative scale, for example using logistic regression. The null hypothesis for statistical interaction in this model is that the joint effect of being exposed to both A and B is multiplicative (Table 1). Departure from this (super- or submultiplicativity) will provide evidence of statistical interaction. A hypothetical example of a multiplicative relationship is given in Table 3. Note that the data in Table 2 show evidence of statistical interaction if risk ratios are examined (multiplicative model), although there is no statistical interaction if risk differences are studied (additive model). Table 3, however, shows no evidence of statistical interaction if risk ratios are examined but shows statistical interaction if risk differences are studied.

Statistical interaction is therefore both present and absent in each of the two hypothetical examples given in Tables 2 and 3, depending on whether the data in each example are analysed under an additive or a multiplicative model. With the exception of qualitative interactions, statistical interaction is therefore model dependent, and does not have any clear biological meaning (Clayton & McKeigue, 2001). If a study is adequately powered, evidence of statistical interaction can always be found by looking at both risk differences and risk ratios within the same data.

Is there a 'correct' model to use?

Under the sufficient-component-cause model of disease it can be argued that risk factors co-participating as causal components in any one causal model of
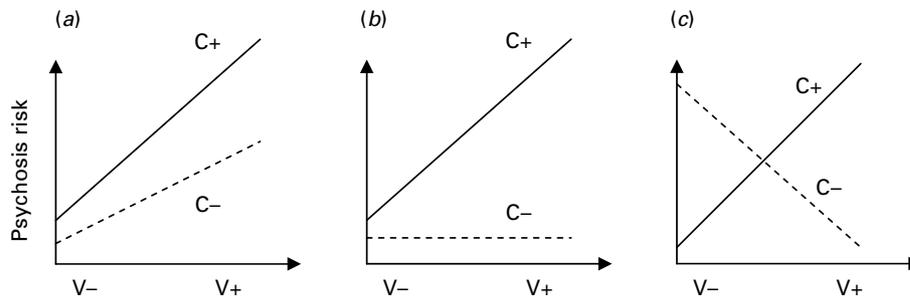
Fig. 1. Hypothetical interactions between *COMT* rs4680 valine allele (V + where valine allele present, V − where absent) and cannabis use (C + where cannabis exposure present, C − where absent) on risk of psychosis.

disease will show departure from additivity (Greenland *et al.* 2008). This concept is sometimes referred to as 'biological interaction' but we will avoid using this term as it can be ambiguous, often being misinterpreted as though it tells us something about pathological mechanisms. It does not; it tells us only that *both* risk factors A and B played a causal role in those particular cases of disease, and that some individuals only experienced the disease because they were exposed to *both* risk factors A and B (i.e. they would not have developed the disease if either A or B were absent) (Greenland *et al.* 2008). It is important to understand that neither interaction under an additive model nor under a multiplicative one is likely to tell us anything about the underlying biology or pathology beyond that from the study of main effects only (except perhaps where qualitative interactions occur) (Thompson, 1991).

Consider an example (taken from Rothman & Greenland, 2005) of an elderly person who falls and fractures their hip, where other factors contributing to the fall include disturbed balance resulting from a childhood head injury, the icy weather, the type of shoes they were wearing, their body weight, and the strong wind at the time. If any of these component causes had been absent, that person would not have broken their hip, and there will be interaction on an additive scale between all of these factors. Within epidemiology, such non-additivity is perhaps what we might expect for most risk factors acting on any disease of complex multifactorial aetiology, where risk factors are neither necessary nor sufficient to bring about disease. As Greenland *et al.* (2008) note, we would rarely expect to observe additivity between risk factors as it is very unlikely that risk factors never co-participate in any causal models of disease.

In the absence of any strong theoretical grounds upon which to base our null hypothesis for interaction, this expectation that most risk factors combine non-additively seems a reasonable assumption to make. Indeed, empirical evidence suggests that, where the combined effect of two risk factors on disease has been examined, findings in general tend to support a multiplicative (or at least greater than additive) pattern of joint risk, both in other fields of medicine (Godsland *et al.* 2000; Morgan *et al.* 2004; JBS, 2005; Wraith & Mengersen, 2007) and in psychiatry (Cougnard *et al.* 2007; van Os *et al.* 2008; Clarke *et al.* 2009). We recently examined the patterns of schizophrenia risk for pair-wise combinations of five risk factors for schizophrenia (cannabis use, lower IQ score, other psychiatric diagnoses, poor social adjustment, and disturbed behaviour) in the 1969 Swedish conscript cohort (personal communication). Of 10 possible combinations, we observed statistical evidence of super-additivity in seven of these, with weak evidence in two others.

These findings, in conjunction with observations from other areas of medicine, suggest that multiplicative models may represent a better fit than additive models for data describing the combination of risk factors on disease. However, irrespective of the choice of the most appropriate model to use, perhaps the important question that remains is: how useful is it to study interactions?

Why study interactions?

Let us now consider the three main reasons put forward as arguments as to why evidence of interaction may be helpful.

First, study of interactions may allow us to find evidence of risk factors for disease that would not be found if only main effects of exposures were examined. As an example, let us imagine that we want to examine whether the valine (Val) allele of Val[158]Met (rs4680) within the catechol-*O*-methyl transferase gene (*COMT*) is a risk factor for psychosis, and that this genetic effect varies according to whether or not someone uses cannabis.

Some of the hypothetical patterns of risk whereby interactions may occur are depicted in Fig. 1. Note that

in Fig. 1*a*, the Val allele is a risk factor for psychosis in both cannabis users and non-users, although it is more strongly associated in users. In this example, the average effect in the whole population (regardless of cannabis use status) will be that the Val allele increases psychosis risk, and therefore the study of interactions here will not greatly benefit identification of *COMT* variation as a risk factor for psychosis. In Fig. 1*b*, however, the average effect of the Val allele on psychosis risk in the population may be minimal if most people in the population do not use cannabis. In this case, stratification of the effect of *COMT* variation by cannabis use may allow identification of the Val allele as a risk factor for psychosis in cannabis users. In Fig. 1*c* (a qualitative, cross-over interaction), the average effect of the Val allele on psychosis risk in the population may be zero if there are equal numbers of cannabis users and non-users in the population. In this case, stratification of the effect of *COMT* variation by cannabis use would allow identification of the Val allele as a risk factor for psychosis in cannabis users, but a protective factor in non-users.

In the original report of this putative interaction (Caspi *et al.* 2005), cannabis use was associated with a 10-fold increase in risk of psychosis in Val homozygotes, but with no increase in risk in those without the valine allele (pattern similar to Fig. 1*b*). If such an interaction effect size was correct, however, then we might expect main effects to be observed for both exposures without the need to study interactions given the frequencies of the Val allele and of cannabis use in the population.

Of the three patterns described above, Fig. 1*a* seems the most plausible biologically, whereas Fig. 1*c* seems the least plausible (though not impossible). Patterns of interaction as seen in Fig. 1*b* can be seen in single gene disorders (e.g. phenylketonuria), but it seems unlikely such a pattern would be observed for epidemiological studies of multifactorial complex diseases given that risk factors are neither necessary nor sufficient to cause disease. Given the rather extreme interactions required therefore for studies of G × E to help identify novel risk factors for disease [i.e. patterns similar to Fig. 1(*b* or *c*)], it seems unlikely that studies of interactions will contribute substantially to the identification of novel risk factors for disease (Munafo *et al.* 2009).

The second reason put forward for studying interactions is that this may lead to an increased understanding of disease aetiology. However, as discussed above, statistical interaction under an additive model (and lack of statistical interaction under a multiplicative model is compatible with this) simply tells us that both risk factors are component causes of a causal model of disease, and nothing about pathological mechanisms (Thompson, 1991; Greenland *et al.* 2008). For example, if a greater than additive relationship between stressful life events (SLEs) and variation within the serotonin transporter locus (*5-HTTLPR*) were found, it would be incorrect to interpret this as evidence that SLEs increase depression risk through effects on the serotonergic system. In reality, this might be true, and evidence for this could be obtained from other studies, but it would be incorrect to deduce this on the basis of an additive (or indeed a multiplicative) interaction between SLEs and *5-HTTLPR*.

Studies of environment–environment interactions (E × E) have been ongoing for decades within epidemiology, and as yet there has been little evidence that studying interactions has benefited understanding of the pathogenesis of the diseases examined (Clayton & McKeigue, 2001). It is possible that more extreme interactions (e.g. super-multiplicative) may be informative about disease aetiology, but in the main it is probably only qualitative interactions that will increase understanding of biological mechanisms beyond that gained from studying main effects only (Thompson, 1991). Such interactions are important to find, if present, although evidence to date suggests they are very uncommon.

The third reason put forward for studying interactions is that they may allow for specific targeting of interventions in high-risk groups. Targeted interventions are often an inefficient approach at a population level, given that most individuals who develop a disease are not usually in the targeted high-risk groups (Rose, 2005), but such a strategy can be important at an individual level.

In fact, where the relationship between two risk factors is greater than additive, the largest reduction in absolute risk of disease will always be obtained from interventions targeted at those exposed to both factors. Evidence of additive interaction would therefore support an approach of targeting high-risk groups, but seems unnecessary given that, as discussed earlier, the relationships between risk factors for multifactorial complex diseases are unlikely to be truly additive at an epidemiological level. Indeed, the programme of interventions aimed at reducing cardiovascular disease by specific targeting of high-risk groups is based upon the assumption of multiplicative models of combined effects on risk (JBS, 2005). Where very strong interactions occur (e.g. patterns similar to Fig. 1*b*), the case for selective interventions may be strengthened, but it is where qualitative interactions occur that clearly has the most important implications for targeting high-risk groups.

The practical implications, therefore, from studying G × E are likely to be limited for epidemiological

studies, although they might be greater for studies of pharmacokinetics and more direct studies of cell or system biology.

Costs of studying interactions

The possible advantages of studying interactions have to be weighed up against the potential costs. The main problem of studying interactions is that of multiple testing (Colhoun *et al.* 2003; Hunter, 2005; Sullivan, 2007; Flint & Munafo, 2008), particularly now with the availability of vast arrays of genotypes from genome-wide studies that can potentially be used in studies of $G \times E$. Even where studies aim to replicate tentative interactions reported, multiple testing is a common problem as tests are often carried out under different genetic models, using multiple measures of environmental exposures, using multiple outcomes, or examining subgroups. This problem could be reduced if: (*a*) all interactions reported were treated as highly tentative and most likely to be due to chance (regardless of strength of statistical evidence or biological plausibility); (*b*) strict criteria for claiming replication were observed, with use of the same statistical and genetic models, the same (or very similar) environmental measure, and the same outcome, without addition of further (e.g. three-way or four-way) interactions (i.e. avoiding further subgroup analyses); and (*c*) all interaction tests performed were reported and published to reduce publication bias.

Studies of $G \times E$ in psychiatry frequently violate these criteria. Perhaps this is best exemplified by claims of replicated evidence of an interaction between *5-HTT* and SLE on risk of depression (Moffitt *et al.* 2005; Rutter *et al.* 2006; Uher & McGuffin, 2008). Most of the apparent replication studies use different statistical models, only report findings in subgroup analyses, or report qualitatively different patterns of interaction to the original findings (Zammit & Owen, 2006; Munafo *et al.* 2009; Risch *et al.* 2009).

Problems of studying interactions also include lower statistical power (Munafo *et al.* 2009) and reduced precision compared to study of main effects, and problems related to misclassification of exposures and confounders that are more complex than those for the study of main effects (Greenland, 1993).

Lack of appreciation of the different interpretation of interaction results under different statistical models and the erroneous assumption that evidence of interactions tells us anything about underlying pathological mechanisms have led to much misunderstanding of how findings in this field should be interpreted. Any gains from studying interactions are therefore not only likely to be somewhat limited but also to come at a cost; studying interactions in the context of the common errors and misunderstanding highlighted above wastes valuable (and limited) time and resources. Problems of poor methodology and incorrect interpretation of results are of course common to all areas of research. However, these are likely to be more common in the study of interactions because of the increased complexity involved. Furthermore, and of greater concern, is that this lack of understanding means that reports of putative interactions have the potential to start impacting, inappropriately, upon clinical practice (Wilhelm *et al.* 2009).

Conclusions

Although providing evidence of interaction without specifying the underlying model is clearly meaningless, the choice of which is the most appropriate model to use is less clear. Empirical evidence supports the view that it seems unlikely that risk factors for multifactorial complex diseases will combine additively and not ever co-participate in any causal models of disease. Given this, it follows that subjects with co-exposure to risk factors will benefit the most from targeted interventions, and studies of interactions will rarely add anything of value to such an approach. Furthermore, irrespective of the model used, it is unlikely that evidence of interaction will increase understanding of pathogenesis, often advocated as one of the benefits potentially arising from studying interactions. The only exception is where qualitative interactions occur, and such interactions have only rarely been described in epidemiology. Other putative benefits, such as increased power to detect novel risk factors, are also likely to occur only in uncommon circumstances, and any potential gains need to be balanced against the substantial problems such as multiple testing and publication bias inherent in such studies.

There has been a recent increase in the development of research projects around $G \times E$ and funding of research in this area. Gains from such studies will only accrue if studies are set within a clear understanding of what is being studied, and where results from statistical interactions examined are interpreted appropriately.

Declaration of Interest

None.

References

Blazei RW, Iacono WG, McGue M (2008). Father–child transmission of antisocial behavior: the moderating role of father's presence in the home. *Journal of the American Academy of Child and Adolescent Psychiatry* **47**, 406–415.

Caspi A, Moffitt TE, Cannon M, McClay J, Murray R, Harrington H, Taylor A, Arseneault L, Williams B, Braithwaite A, Poulton R, Craig IW (2005). Moderation of the effect of adolescent-onset cannabis use on adult psychosis by a functional polymorphism in the catechol-*O*-methyltransferase gene: longitudinal evidence of a gene × environment interaction. *Biological Psychiatry* **57**, 1117–1127.

Clarke MC, Tanskanen A, Huttunen M, Whittaker JC, Cannon M (2009). Evidence for an interaction between familial liability and prenatal exposure to infection in the causation of schizophrenia. *American Journal of Psychiatry* **166**, 1025–1030.

Clayton D, McKeigue PM (2001). Epidemiological methods for studying genes and environmental factors in complex diseases. *Lancet* **358**, 1356–1360.

Colhoun HM, McKeigue PM, Davey Smith G (2003). Problems of reporting genetic associations with complex outcomes. *Lancet* **361**, 865–872.

Cougnard A, Marcelis M, Myin-Germeys I, De Graaf R, Vollebergh W, Krabbendam L, Lieb R, Wittchen HU, Henquet C, Spauwen J, Van Os J (2007). Does normal developmental expression of psychosis combine with environmental risk to cause persistence of psychosis? A psychosis proneness-persistence model. *Psychological Medicine* **37**, 513–527.

Darroch J (1997). Biologic synergism and parallelism. *American Journal of Epidemiology* **145**, 661–668.

Flint J, Munafo MR (2008). Forum: Interactions between gene and environment. *Current Opinion in Psychiatry* **21**, 315–317.

Godsland IF, Winkler U, Lidegaard O, Crook D (2000). Occlusive vascular diseases in oral contraceptive users. Epidemiology, pathology and mechanisms. *Drugs* **60**, 721–869.

Greenland S (1993). Basic problems in interaction assessment. *Environmental Health Perspectives* **101** (Suppl. 4), 59–66.

Greenland S, Lash TL, Rothman KJ (2008). Concepts of interaction. In *Modern Epidemiology* (ed. K. J. Rothman, S. Greenland and T. L. Lash), pp. 71–83. Lippincott Williams & Wilkins: Philadelphia, PA.

Hunter DJ (2005). Gene–environment interactions in human diseases. *Nature Reviews Genetics* **6**, 287–298.

Jaffee SR, Moffitt TE, Caspi A, Taylor A (2003). Life with (or without) father: the benefits of living with two biological parents depend on the father's antisocial behavior. *Child Development* **74**, 109–126.

JBS (2005). JBS 2: Joint British Societies' guidelines on prevention of cardiovascular disease in clinical practice. *Heart* **91** (Suppl. 5), v1–52.

Moffitt TE, Caspi A, Rutter M (2005). Strategy for investigating interactions between measured genes and measured environments. *Archives of General Psychiatry* **62**, 473–481.

Morgan TR, Mandayam S, Jamal MM (2004). Alcohol and hepatocellular carcinoma. *Gastroenterology* **127**, S87–S96.

Munafo MR, Durrant C, Lewis G, Flint J (2009). Gene × environment interactions at the serotonin transporter locus. *Biological Psychiatry* **65**, 211–219.

Risch N, Herrell R, Lehner T, Liang KY, Eaves L, Hoh J, Griem A, Kovacs M, Ott J, Merikangas KR (2009). Interaction between the serotonin transporter gene (5-HTTLPR), stressful life events, and risk of depression: a meta-analysis. *Journal of the American Medical Association* **301**, 2462–2471.

Rose G (2005). *The Strategy of Preventive Medicine*. Oxford University Press: Oxford.

Rothman KJ, Greenland S (2005). Causation and causal inference in epidemiology. *American Journal of Public Health* **95** (Suppl. 1), S144–S150.

Rutter M, Moffitt TE, Caspi A (2006). Gene–environment interplay and psychopathology: multiple varieties but real effects. *Journal of Child Psychology and Psychiatry* **47**, 226–261.

Sullivan PF (2007). Spurious genetic associations. *Biological Psychiatry* **61**, 1121–1126.

Thompson WD (1991). Effect modification and the limits of biological inference from epidemiologic data. *Journal of Clinical Epidemiology* **44**, 221–232.

Uher R, McGuffin P (2008). The moderation by the serotonin transporter gene of environmental adversity in the aetiology of mental illness: review and methodological analysis. *Molecular Psychiatry* **13**, 131–146.

van Os J, Rutten BP, Poulton R (2008). Gene–environment interactions in schizophrenia: review of epidemiological findings and future directions. *Schizophrenia Bulletin* **34**, 1066–1082.

Wilhelm K, Meiser B, Mitchel, PB, Finch AW, Siegel JE, Parker G, Schofield PR (2009). Issues concerning feedback about genetic testing and risk of depression. *British Journal of Psychiatry* **194**, 404–410.

Wraith D, Mengersen K (2007). Assessing the combined effect of asbestos exposure and smoking on lung cancer: a Bayesian approach. *Statistics in Medicine* **26**, 1150–1169.

Zammit S, Owen MJ (2006). Stressful life events, 5-HTT genotype and risk of depression. *British Journal of Psychiatry* **188**, 199–201.