Misconceptions about gene–environment interactions in psychiatry

Stanley Zammit,1,2 Michael J. Owen,1 Glyn Lewis2

Misunderstandings surrounding the study of gene–environment interactions are very common. Given the large increase in the number of studies examining interactions in recent years, this raises serious concerns about the value of time and resources spent in these endeavours. In this article we discuss why, despite frequent claims to the contrary, studies of gene–environment interactions are very unlikely to enhance our understanding of disease aetiology or prevention.

PSYCHIATRIC DISORDERS ARE OF MULTIFACTORIAL COMPLEX AETIOLOGY

For multifactorial complex disorders, risk factors are neither necessary nor sufficient to cause disease. In other words, for any given case of disease, multiple risk factors will have acted to bring about the onset of disease in that person. Many different combinations of risk factors will exist for a disease in a population, as described in the sufficient-component-cause model of disease.1 Some of these risk factors are likely to be genetic, and some non-genetic, and both types almost certainly play a contributory role in causing most, if not all, cases of disease. This is sometimes referred to as gene–environment interaction (definition 1) and few scientists would dispute that this is a real phenomenon.

However, this is not the same definition of gene–environment interaction as that reported by studies following analysis of data, which always refers to a measure of statistical interaction. Here data are compared to predictions from statistical models in order to study how exposure to two risk factors in combination affects disease risk. Statistical interaction occurs when the risk of disease if exposed to both the gene (A) and the environmental exposure (B) is different from that predicted by the statistical model being used (definition 2; henceforth referred to as G×E).

However, these predictions can be modelled on either additive or multiplicative scales, and it is important to appreciate that interactions under these models mean very different things (see Appendix 1). For example, statistical interaction between A and B under an additive model (eg, when studying risk differences) means that the risk of disease if both A and B are present is not equal to the additive effect (sum) of the risk if exposed to A only and that if exposed to B only (ie, risk from joint exposure could be greater or less than this). Similarly, statistical interaction between A and B under a multiplicative model (eg, studying risk ratios using logistic regression) means that the risk of disease if both A and B are present is not equal to the multiplicative effect (product) of the risk if exposed to A only and that if exposed to B only (ie, risk from joint exposure could be greater or less than this).

This highlights an important, but often overlooked issue. If both A and B are associated with the disease, then assuming a study is adequately powered, one can always find evidence of statistical interaction by testing both additive and multiplicative models (eg, looking at both risk differences and risk ratios within the same data). In this case, simply providing statistical evidence of G×E has no intrinsic value. Furthermore, hypothesizing that an environmental exposure will have a ‘greater effect’ in the presence of a risk allele compared to absence of that allele is also ambiguous unless the statistical model is specified. For example, the hypothetical data in table 1 show that, under an additive model, the effect of E on risk of disease is greater when G is also present compared to G being absent. However, under a multiplicative model, the same data show that the effect of E on risk of disease is smaller when G is also present. Both models show evidence of G×E, but with opposite conclusions.

If it is not previously known whether A and B are risk factors for a disease, then evidence of interaction, if well replicated (see later), is potentially of interest, as both A and B must be associated with the disease for an interaction to have been observed. However, statistical power to find evidence of interaction will be substantially enhanced compared to power to find main effects only in conditions that are likely to be very uncommon for multifactorial complex diseases (see later).

Therefore, debate over whether gene–environment interactions ‘exist’ is not helpful. The more useful questions are as follows: Can we reliably describe the patterns of disease risk for joint exposure to two (or more) risk factors? And, most importantly, is doing so likely to help us understand more about disease aetiology or prevention? We will deal with the second of these two questions first.

QUESTION 1: WHY STUDY INTERACTIONS?

There are three main reasons put forward as to why evidence of interaction may be helpful.

(I) First, it is often argued that study of interactions may allow us to find evidence of risk factors for disease (either G or E) that would not be found if only main effects of exposures were examined. However, the patterns of interaction that would be required for power to find an interaction to be substantially greater than power to find a main effect are likely to be very uncommon for multifactorial complex diseases.2 For example, the largest gain in power from studying interactions over main effects only will occur when the effect of one exposure (eg, E) has opposite effects on disease risk according to whether another exposure (G) is present or not. These are known as qualitative or cross-over interactions. An example of a qualitative interaction is 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.3 4 Such relationships, although possible, seem rather implausible for biological exposures such as genetic variants and are, therefore, likely to be rare in studies of G×E.

Study of interactions may also be informative for identifying novel genetic risk variants if G has a strong effect when E is also present, but has no effect if E is
### Table 1: Hypothetical 1-year cumulative incidence of disease (per 10,000) in relation to exposure to environmental (E) and genetic (G) exposures

<table>
<thead>
<tr>
<th>E−</th>
<th>E+</th>
</tr>
</thead>
<tbody>
<tr>
<td>G−</td>
<td>1</td>
</tr>
<tr>
<td>G+</td>
<td>2</td>
</tr>
</tbody>
</table>

**Additive model:** Risk difference for E is 4 where G is absent (5−1), and is 6 where G is present (8−2); the effect of E on disease is greater when G is also present.

**Multiplicative model:** Risk ratio for E is 5 where G is absent (5/1), and is 4 where G is present (8/2); the effect of E on disease is smaller when G is also present.

absent, particularly if E is rare. However, such relationships imply great specificity between E and G (ie, that no other environmental exposures co-participate with G in any causal model of disease). Such relationships seem rather unlikely for multifactorial complex diseases, though perhaps are more plausible for studies of adverse effects in pharmacogenetic studies.

As the patterns of interaction become less extreme, gains in power from studying interaction over that of studying main effects decreases, and under most scenarios, simulations indicate that gains from studying G×E over studying main effects only are likely to be minimal, if any. Even where gains in power can occur from studying G×E, this has to be offset against the potential problem of multiple testing and Type I errors. A major problem in the field of genetics has been that of multiple testing, particularly now with the availability of vast arrays of genotypes from genome-wide studies. Examining not only main effects of genes but also their interactions (potentially under different statistical and genetic models) with a multitude of environmental exposures, or on a number of different outcomes, means that spurious results are likely to be extremely common. This problem of multiple testing is further exacerbated by studies reporting interactions in sub-group analyses, or after testing 3-way, 4-way, or more complex interactions.

The second reason put forward for studying interactions is that this may lead to an increased understanding of biological mechanisms underlying disease aetiology. This may be true where qualitative interactions occur, that is, where one exposure has completely opposite effects depending upon the presence or absence of another factor, but is unlikely to be true for other patterns of interaction.

Under the sufficient-component-cause model of disease it can be shown⁷ that where two risk factors co-participate as causal components in any one causal model of disease, the risk of disease if both risk factors are present will be greater than additive (note that where a joint effect is greater than additive, this is also compatible with lack of evidence, or evidence in support of, interaction under a multiplicative model (Appendix 1)). However, evidence of statistical interaction between A and B under an additive model tells us only that both risk factors A and B played a causal role in some cases of disease and that some individuals only experienced the disease because they were exposed to both risk factors A and B (ie, they would not have developed the disease if either A or B were absent).

Although G×E under an additive model, therefore, tell us that both risk factors are component causes of a causal model of disease, they tell us nothing about underlying biological mechanisms beyond what can be gleaned from main effects. For example, if one found a greater than additive relationship between stressful life events (SLEs) and variation within the serotonin transporter locus (5-HTTLPR), 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 one could obtain evidence for this from other studies, but it would be incorrect to deduce this on the basis of an additive interaction between SLEs and 5-HTTLPR.

Variants within many genes (eg, HLA-DQA1, ZNF804a, and others) now show apparently robust association with schizophrenia,⁸ and cannabis use is also associated with this disorder. Assuming these associations are all real effects, and that we have an adequately powered dataset, we can predict with certainty that we will find evidence of G×E between cannabis and these variants by examining joint effects under either additive or multiplicative models. Similarly, the effects of all other environmental exposures that show association with schizophrenia (eg, victimization, urbanicity etc) will show effects that are conditional on these genes in either additive or multiplicative models. Therefore, unless any of these are qualitative patterns of interaction, it is unlikely that our understanding of biological mechanisms underlying schizophrenia aetiology would be advanced by such evidence of G×E, despite much enthusiasm for adopting this approach.¹⁰

It is possible that more extreme interactions (eg, joint effects that are much greater than 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. Such interactions are important to find if present, though evidence to date suggests they are very uncommon.

**QUESTION 2: CAN WE RELIABLY DESCRIBE PATTERNS OF RISK FOR JOINT EXPOSURE (REPLICATION)??**

Nowhere are the misunderstandings surrounding the study of G×E more apparent than in the literature claiming replication of reported interactions. Interpreting replication of interactions is much more complex than that for main effects.
An essential requirement for claiming replication of G×E is that the pattern of disease risk where there is joint exposure to both risk factors compared to where there is exposure to only one or no risk factors is the same as in the original study. Unfortunately, this is frequently ignored, and unjustified claims of replication are often made simply on the basis of finding evidence of G×E.

Problems of multiple testing in studies of G×E are further compounded by a temptation for authors to conduct subgroup analyses, or to test 3-way or 4-way interactions in what often appear to be increasingly desperate efforts to find evidence of G×E and thereby (incorrectly) claim to replicate an original finding. We illustrate the problems associated with replication of reported G×E by using three high-profile examples.

5-HTT x SLE on depression
An interaction reported between SLEs and variation within a polymorphism in the serotonin transporter gene (5-HTTLPR) on risk of depression received substantial media attention and has undoubtedly been influential in the large increase in studies examining G×E over recent years. In the original report, presence of SLEs was associated with increased risk of depression in subjects homozygous for the s-allele (OR for depression per SLE = 1.68, 95% CI 1.25 to 2.30), with a weaker effect in heterozygotes (OR = 1.48, 95% CI 1.24 to 1.76), and especially l-allele homozygotes (OR = 1.17, 95% CI 0.91 to 1.51). Risk of depression when exposed to both SLEs and 5HTT s-allele was greater than multiplicative (and hence also greater than additive).

Of 34 studies that have examined this relationship since then (see online supplement), 17 have examined whether this relationship is indeed greater than multiplicative. Of these, only 2 find a pattern of interaction consistent with that in the originally study (1 of these in a sub-group only), 3 find evidence of interaction but with patterns inconsistent with that originally reported (4 of these in the opposite direction) and 7 fail to find evidence of any interaction (5 of these with sample sizes substantially larger than the original). Of 15 studies that examined additive models, 6 reported interactions with patterns consistent with the original study (2 in sub-groups only), 3 found patterns that were inconsistent with the original (1 opposite) and 6 failed to find evidence of interaction, though sample sizes were on average, 3-times smaller than the original study. Two of these 6 studies reported G×E in subsequent subgroup analyses.

**MAOA x maltreatment on antisocial or violent behaviour**
An interaction between variation in the monoamine oxidase A gene (MAOA) and childhood maltreatment on antisocial and violent behaviour in males was reported, such that the effect of maltreatment was greater with the low-activity variants of MAOA compared to those with high activity. Again, this relationship was greater than multiplicative.

Of 11 studies that have since examined this in males (see online supplement), only 4 have examined whether this relationship was indeed greater than multiplicative. Of these, 3 found no evidence for such a relationship, while 1 found support present only when reporting a 1-tailed p value. Of 10 studies that examined additive models, 5 reported evidence of a greater than additive effect (2 of these in subgroup analyses only), 4 found no evidence and 1 found evidence of an interaction but in the opposite direction to that originally reported. In samples that also included females, 1 study examined a multiplicative relationship and found an interaction in the opposite direction. Of 6 studies that examined additive models, 2 found evidence of a greater than additive effect, 3 failed to do so and 1 found an effect in the opposite direction. Studies finding, or failing to find evidence of interaction were equable in size.

**COMT x cannabis on psychosis**
An interaction between variation in COMT Val108Met (rs4680) and cannabis use on psychosis was reported, such that the effect of cannabis was greater in individuals with the valine (high activity) variant compared to those with the methionine allele. Again, this relationship was more than multiplicative, with cannabis use associated with a 10-fold increase in psychosis risk in valine homozygotes, but with no increase in risk in methionine homozygotes. If such an interaction effect size was correct, then we would expect main effects to be observed for both exposures without the need to study interactions given the frequencies of the valine allele and of cannabis use in the population; nevertheless, even in large meta-analyses the association between COMT Val108Met and schizophrenia remains uncertain, despite more than adequately sized samples to identify a main effect of COMT if an interaction with cannabis existed as described above. Three studies that have subsequently examined this relationship (see online supplement) found no evidence to support a relationship as originally reported. In one of these studies, an interaction was reported but only from a subgroup (3-way interaction) analysis.

Overall, as can be seen, there is little consistent evidence to support the patterns of disease risk described by any of the original three studies discussed previously. A recent meta-analysis, for example, failed to support a greater than multiplicative relationship between 5-HTT and SLEs on risk of depression. Furthermore, there was no evidence to support a main effect of 5-HTT in this or another recent meta-analysis, though both had adequate power to detect one if the G×E exists as originally reported.

It is possible to introduce bias in meta-analyses by only including studies from which data can be extracted, while debate about inclusion criteria for meta-analyses, as well as about potential reasons for non-replication, often centre on the use of different measures or different sample characteristics. Measures that are less likely than others to accurately reflect an environmental exposure or outcome may reduce the power to observe main effects and interactions, although we wouldn’t expect the pattern of interaction to be substantially different. Similarly, sample characteristics such as age might result in weaker evidence of effects (eg, if time between E and disease increases with age), but is unlikely to substantially alter the pattern of risk for joint exposure to E and G.

Reviews and meta-analyses of putative G×E can be informative if done adequately. However, it is a weak argument to conclude that there is evidence of replication of a specific relationship between two risk factors based merely on the number of studies showing G×E, without also considering the patterns of interaction reported.

**SUPPORTING EVIDENCE**
Finally, it has been argued, quite reasonably, that in assessing the putative G×Es reported, other lines of evidence should be taken into account. However, conclusions from reports of G×E in, for example, animal models, are equally limited. For example, evidence of greater than additive relationships between 5-HTT and stressful events does not tell us anything about the underlying biology of depression in animals other than both stress and 5-HTT play contributory roles in aetiology.

**5-HTT x SLE on depression**
An interaction reported between SLEs and variation within a polymorphism in the serotonin transporter gene (5-HTTLPR) on risk of depression received substantial media attention and has undoubtedly been influential in the large increase in studies examining G×E over recent years. In the original report, presence of SLEs was associated with increased risk of depression in subjects homozygous for the s-allele (OR for depression per SLE = 1.68, 95% CI 1.25 to 2.30), with a weaker effect in heterozygotes (OR = 1.48, 95% CI 1.24 to 1.76), and especially l-allele homozygotes (OR = 1.17, 95% CI 0.91 to 1.51). Risk of depression when exposed to both SLEs and 5HTT s-allele was greater than multiplicative (and hence also greater than additive).

Of 34 studies that have examined this relationship since then (see online supplement), 17 have examined whether this relationship is indeed greater than multiplicative. Of these, only 2 find a pattern of interaction consistent with that in the originally study (1 of these in a sub-group only), 3 find evidence of interaction but with patterns inconsistent with that originally reported (4 of these in the opposite direction) and 7 fail to find evidence of any interaction (5 of these with sample sizes substantially larger than the original). Of 15 studies that examined additive models, 6 reported interactions with patterns consistent with the original study (2 in sub-groups only), 3 found patterns that were inconsistent with the original (1 opposite) and 6 failed to find evidence of interaction, though sample sizes were on average, 3-times smaller than the original study. Two of these 6 studies reported G×E in subsequent subgroup analyses.

**MAOA x maltreatment on antisocial or violent behaviour**
An interaction between variation in the monoamine oxidase A gene (MAOA) and childhood maltreatment on antisocial and violent behaviour in males was reported, such that the effect of maltreatment was greater with the low-activity variants of MAOA compared to those with high activity. Again, this relationship was greater than multiplicative.

Of 11 studies that have since examined this in males (see online supplement), only 4 have examined whether this relationship was indeed greater than multiplicative. Of these, 3 found no evidence for such a relationship, while 1 found support present only when reporting a 1-tailed p value. Of 10 studies that examined additive models, 5 reported evidence of a greater than additive effect (2 of these in subgroup analyses only), 4 found no evidence and 1 found evidence of an interaction but in the opposite direction to that originally reported. In samples that also included females, 1 study examined a multiplicative relationship and found an interaction in the opposite direction. Of 6 studies that examined additive models, 2 found evidence of a greater than additive effect, 3 failed to do so and 1 found an effect in the opposite direction. Studies finding, or failing to find evidence of interaction were equable in size.

**COMT x cannabis on psychosis**
An interaction between variation in COMT Val108Met (rs4680) and cannabis use on psychosis was reported, such that the effect of cannabis was greater in individuals with the valine (high activity) variant compared to those with the methionine allele. Again, this relationship was more than multiplicative, with cannabis use associated with a 10-fold increase in psychosis risk in valine homozygotes, but with no increase in risk in methionine homozygotes. If such an interaction effect size was correct, then we would expect main effects to be observed for both exposures without the need to study interactions given the frequencies of the valine allele and of cannabis use in the population; nevertheless, even in large meta-analyses the association between COMT Val108Met and schizophrenia remains uncertain, despite more than adequately sized samples to identify a main effect of COMT if an interaction with cannabis existed as described above. Three studies that have subsequently examined this relationship (see online supplement) found no evidence to support a relationship as originally reported. In one of these studies, an interaction was reported but only from a subgroup (3-way interaction) analysis.

Overall, as can be seen, there is little consistent evidence to support the patterns of disease risk described by any of the original three studies discussed previously. A recent meta-analysis, for example, failed to support a greater than multiplicative relationship between 5-HTT and SLEs on risk of depression. Furthermore, there was no evidence to support a main effect of 5-HTT in this or another recent meta-analysis, though both had adequate power to detect one if the G×E exists as originally reported.

It is possible to introduce bias in meta-analyses by only including studies from which data can be extracted, while debate about inclusion criteria for meta-analyses, as well as about potential reasons for non-replication, often centre on the use of different measures or different sample characteristics. Measures that are less likely than others to accurately reflect an environmental exposure or outcome may reduce the power to observe main effects and interactions, although we wouldn’t expect the pattern of interaction to be substantially different. Similarly, sample characteristics such as age might result in weaker evidence of effects (eg, if time between E and disease increases with age), but is unlikely to substantially alter the pattern of risk for joint exposure to E and G.

Reviews and meta-analyses of putative G×E can be informative if done adequately. However, it is a weak argument to conclude that there is evidence of replication of a specific relationship between two risk factors based merely on the number of studies showing G×E, without also considering the patterns of interaction reported.

**SUPPORTING EVIDENCE**
Finally, it has been argued, quite reasonably, that in assessing the putative G×Es reported, other lines of evidence should be taken into account. However, conclusions from reports of G×E in, for example, animal models, are equally limited. For example, evidence of greater than additive relationships between 5-HTT and stressful events does not tell us anything about the underlying biology of depression in animals other than both stress and 5-HTT play contributory roles in aetiology.

Assessment
Relationship Interaction
This could be greater than additive,

Additive Risk (A and B) = Risk (A only) + Risk (B only) – Risk (Neither A nor B)

No (Null hypothesis) Yes
This is also less than multiplicative greater than additive

Greater than additive
Risk (A and B) > Risk (A only) + Risk (B only) – Risk (Neither A nor B)

Yes
This could be greater than multiplicative, additive, or less than additive

Less than additive
Risk ratio (A and B) < Risk ratio (A only) x Risk ratio (B only)

No (Null hypothesis) Yes
This is also greater than additive

Greater than multiplicative
Risk ratio (A and B) > Risk ratio (A only) x Risk ratio (B only)

Yes
This could be greater than additive, additive, or less than additive

Less than multiplicative
Risk ratio (A and B) < Risk ratio (A only) x Risk ratio (B only)

Yes
This is also greater than additive

Multiplicative

Greater than additive
Risk ratio (A and B) = Risk ratio (A only) x Risk ratio (B only)

No (Null hypothesis) Yes
This is also greater than additive

Less than additive
Risk ratio (A and B) < Risk ratio (A only) x Risk ratio (B only)

Yes
This could be greater than additive, additive, or less than additive

Statistical models used to study interactions between two risk factors, A and B


REFERENCES

Misconceptions about gene–environment interactions in psychiatry

Stanley Zammit, Michael J Owen and Glyn Lewis

_Evid Based Mental Health_ 2010 13: 65-68 originally published online June 8, 2010
doi: 10.1136/ebmh1056

Updated information and services can be found at:
[http://ebmh.bmj.com/content/13/3/65](http://ebmh.bmj.com/content/13/3/65)

These include:

**Supplementary Material**
Supplementary material can be found at:
[http://ebmh.bmj.com/content/suppl/2011/08/11/13.3.65.DC1](http://ebmh.bmj.com/content/suppl/2011/08/11/13.3.65.DC1)

**References**
This article cites 20 articles, 4 of which you can access for free at:
[http://ebmh.bmj.com/content/13/3/65#BIBL](http://ebmh.bmj.com/content/13/3/65#BIBL)

**Email alerting service**
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

**Topic Collections**
Articles on similar topics can be found in the following collections

- Editor's choice (98)

**Notes**

To request permissions go to:
[http://group.bmj.com/group/rights-licensing/permissions](http://group.bmj.com/group/rights-licensing/permissions)

To order reprints go to:
[http://journals.bmj.com/cgi/reprintform](http://journals.bmj.com/cgi/reprintform)

To subscribe to BMJ go to:
[http://group.bmj.com/subscribe/](http://group.bmj.com/subscribe/)