



Practice of Epidemiology

Challenges in the Use of Literature-based Meta-Analysis to Examine Gene-Environment Interactions

Luigi Palla*, Julian P. T. Higgins, Nicholas J. Wareham, and Stephen J. Sharp

* Correspondence to Dr. Luigi Palla, MRC Epidemiology Unit, Institute of Metabolic Science, Addenbrooke's Hospital, Hills Road, CB2 0QQ Cambridge, United Kingdom (e-mail: luigi.palla@mrc-epid.cam.ac.uk).

Initially submitted December 2, 2009; accepted for publication February 23, 2010.

Statistical interactions between genes and environmental exposures with respect to disease outcomes may help to identify biologic mechanisms and pathways and inform behavioral interventions. The number of persons required for a single study to have sufficient statistical power to detect such interactions may be considered prohibitively large, making a meta-analysis of published literature an apparently attractive alternative. However, meta-analysis of gene-environment interactions using published literature is challenging, with the conclusions being likely to suffer from bias and lack of generalizability. The authors highlight these challenges and biases using an illustrative example: meta-analysis of interactions between the Pro12Ala variant of the peroxisome proliferator-activated receptor γ ($PPAR_{\gamma}$) gene and various diet and lifestyle factors in the risk of diabetes. The authors conclude that literature-based meta-analysis conducted to examine gene-environment interactions is unlikely to provide a meaningful quantitative conclusion. Alternative strategies are required, including analyses in scientific consortia established to assess main genetic effects, where individual participant data can be shared, allowing both greater power and consistency of analysis methods. However, these consortia are likely to be limited by lack of standardization of the measures of environmental factors. This issue may ultimately only be resolvable by the de novo establishment of large single or multicenter cohorts using comparable methods.

bias (epidemiology); diabetes mellitus, type 2; epidemiologic research design; genotype-environment interaction; meta-analysis; PPAR gamma

Abbreviations: GEI, gene-environment interaction; $PPAR_{\gamma}$, peroxisome proliferator-activated receptor γ .

Gene-environment interaction (GEI) has been of interest in scientific research for nearly a century (1). Early GEI studies were carried out in experimental settings and were often aimed at assessing, within experimental populations, how the effect of some environmental condition on a phenotype varied as a function of the genotype at a given locus (2). Only in the last decade have large-scale investigations of interactions between genes and environmental exposures become possible, due to the decreasing cost of genotyping technology. GEI findings may be useful in identifying biologic mechanisms and pathways (3), or they may inform the targeting of new lifestyle and behavioral interventions.

A large sample size is needed if a study is to have sufficient statistical power to detect even small interaction effects, particularly for complex diseases for which the

genetic main effects are small. Smith and Day (4) reported that if the aim of a study is to detect interactions, the study would need to be at least 4 times larger than a study aiming to detect main effects of the same magnitude. Prohibitively large sample sizes may lead researchers to attempt to address the question of GEI via a meta-analysis of the published literature, as has been done for main genetic effects (5) and biomarkers (6). However, the challenges involved in such an endeavor are numerous and may prevent any meaningful results being obtained.

In this paper, we discuss the potential challenges of meta-analyzing GEIs using published literature, exemplifying some of the issues using a systematic review of the potential interactions between an extensively studied genetic variant associated with diabetes and a variety of lifestyle factors. We

structured the paper around the major steps of a systematic review (7), addressing in turn the processes of setting objectives and eligibility criteria for studies; searching for studies; collecting data; assessing risks of bias in the included studies; considering potential biases (e.g., publication bias) in the review as a whole; and synthesizing results across studies. To provide an overview of experience in literature-based reviews targeted at GEI, we searched PubMed (August 14, 2009) using the strategy “meta-analysis *and* (interaction [title] *or* joint effect* [title]) *and* (gene [title] *or* genes [title] *or* genetic [title]).” This retrieved 13 papers (plus a comment), of which 6 were literature-based meta-analyses targeting GEIs in human populations. We supplemented the 6 systematic reviews and meta-analyses with an unpublished report from a colleague (Cosetta Minelli, EURAC Research (Bolzano, Italy), personal communication, 2009). The resulting 7 papers are described in Table 1.

MATERIALS AND METHODS

Example: PPAR γ and type 2 diabetes

To illustrate the challenges of meta-analyzing GEIs using published literature, we describe our experience of undertaking a systematic review and meta-analysis oriented around the peroxisome proliferator-activated receptor γ (PPAR γ) gene, which has been extensively studied in relation to type 2 diabetes (8). PPAR γ is a ligand-activated transcription factor involved in lipid and glucose metabolism, fatty acid transport, and adipocyte differentiation (9). A missense mutation (substitution of alanine for proline) at codon 12 of the PPAR γ gene (10) is present in several human populations. The minor allele of this Pro12Ala polymorphism in Caucasians has a frequency of approximately 12%. Large genetic association studies (>1,000 persons) have consistently shown a protective effect of the minor allele (odds ratio = 0.8), leading to a population attributable risk as high as 0.25 for type 2 diabetes (8). At least 50 genetic association studies have been performed to assess the main effect of the Pro12Ala variant on the risk of diabetes, and 2 meta-analyses have also been published (11, 12). Our interest was in estimating interactions between the Pro12Ala variant and lifestyle factors such as diet and physical activity. The project was motivated by a previous broad-ranging review of gene-lifestyle interactions in diabetes that identified the PPAR γ gene as one with strong biologic plausibility for evidence of GEI (13). To our knowledge, no meta-analysis of these GEIs has been attempted to date.

Setting eligibility criteria

The first step in any systematic review or meta-analysis is to define a clear set of objectives for the review and then formulate specific eligibility criteria for inclusion of studies, based on these objectives. The objectives for a meta-analysis of GEIs include definition of at least 4 components: 1) the disease or outcome, 2) the gene variant(s), 3) the environmental exposure(s), and 4) the types of studies deemed appropriate for examining the interaction. The eligibility criteria can then be set accordingly. As with all reviews, a tradeoff must usually be made between making

these criteria narrow (addressing very specific questions, but with the risk that not many studies will be included) and making these criteria broad (allowing more studies to be included, but with the risk that the questions addressed become unfocused). The examples shown in Table 1 illustrate a variety of approaches. For example, Raimondi et al. (14) considered any gene (examining interaction with smoking in risk of colorectal cancer) but then restricted their attention to those genes that had been investigated 5 or more times. On the other hand, Zeiger et al. (15) sought only studies addressing the *TaqI* polymorphism in the transforming growth factor α (*TGFA*) gene (interacting with maternal smoking in risk of oral clefts).

In general, environmental exposures are likely to be more difficult to specify than genetic variants, since they can be defined and measured in many different ways (16). For a specific exposure such as smoking, this may not pose a major problem, but more complex and difficult-to-measure exposures, such as dietary intake, will require dissecting, either at the stage of determining eligibility criteria for the review or within the review itself.

Eligible study designs for investigating GEI might include case-control studies, cohort studies, cross-sectional studies, and case-only studies, the latter having been devised especially for the study of GEIs under the assumption that genotype and environmental exposure are independent in the general population (17). Comprehensive reviews of GEI methods in epidemiology are presented in papers by Khoury et al. (18) and Yang and Khoury (19). Ideally the eligible study designs should be prespecified, although few of the articles we reviewed clearly stated this. Availability of data may be an eligibility criterion and certainly will determine whether a study is included in a meta-analysis. In their review, Zeiger et al. (15) considered only studies reporting case and control counts for the 2-way cross-classification of the genetic variant with smoking. In contrast, Minelli et al. (unpublished manuscript) sought all relevant studies irrespective of data availability, although they decided eventually that meta-analysis would not be an option given the incomplete reporting of suitable data.

Experiences from PPAR γ review

For our review of PPAR γ , lifestyle factors, and type 2 diabetes, we were highly specific about addressing only the Pro12Ala variant of the PPAR γ gene, but we had broader eligibility criteria for environmental exposures, outcomes, and study designs. We sought studies investigating any lifestyle variables related to physical activity and diet, broadly interpreted to include obesity and intake of specific components such as fatty acids. As outcomes, we included type 2 diabetes (present or not) as well as related quantitative phenotypes such as fasting glucose level and fasting insulin level. We considered any type of study design.

RESULTS

Searching for studies

General considerations for GEIs. A literature-based meta-analysis requires searches of bibliographic databases,

Table 1. Results From 7 Meta-Analyses of Gene-Environment Interactions

Study, Year (Reference No.)	Disease	Gene(s)	Environmental Factor(s)	No. of Studies	Findings	Synthesis Method(s) Used
Minelli et al., unpublished manuscript ^a	Respiratory function	Antioxidant genes	Air pollution	13	Heterogeneous designs, populations, and environmental exposures. Overall suggestion that interactions were present, although finding could have been spurious.	Despite the authors' intentions, no meta-analysis was conducted because of heterogeneity and incomplete reporting.
Raimondi et al., 2009 (14)	Colorectal cancer	Any gene variant examined in 5 or more studies (5 genes were included: <i>GSTM1</i> , <i>GSTT1</i> , <i>mEH3</i> , <i>mEH4</i> , and <i>NAT2</i>)	Smoking	27 case-control or nested case-control studies (12 presented data on <i>GSTM1</i> , 8 on <i>GSTT1</i> , 7 on <i>mEH3</i> and <i>mEH4</i> , and 10 on <i>NAT2</i>)	Weak suggestion of interaction with 1 gene variant.	<i>Q</i> test, <i>I</i> ² , meta-regression, random-effects meta-analysis; pooled <i>P</i> value for interaction by stratified logistic regression (log-likelihood test).
Risch et al., 2009 (22)	Depression	Serotonin transporter gene (<i>5-HTTLPR</i>)	Stressful life events	14 eligible studies, 10 included in meta-analysis	No interaction detected.	Logistic regression within studies, combined using random-effects meta-analysis.
Sanderson et al., 2007 (33)	Bladder cancer	<i>NAT1</i> and <i>NAT2</i>	Smoking	36 (1 study of 3-way interaction, 13 studies of <i>NAT2</i> , and 2 studies of 2-way interaction for <i>NAT1</i>)	Evidence of joint effects.	Novel methods developed to deal with difference subsets of full cross-classification.
Kim-Cohen et al., 2006 (39)	Antisocial behavior	<i>MAOA</i>	Familial adversity in childhood	5	Evidence of interaction.	Correlations between exposure and outcome within genotype categories, <i>Z</i> -transformed. <i>Q</i> , <i>I</i> ² , fixed- and random-effects meta-analysis. Interactions by difference in Fisher <i>Z</i> values within studies.
Zeiger et al., 2005 (15)	Oral clefts	<i>TGFA</i>	Maternal smoking	5	Interaction found, but not robust to exclusion of 1 influential study (strongest effect).	Mantel-Haenszel methods used within smoking categories. Interaction addressed using logistic regression adjusting for study; case-only analysis.
Marcus et al., 2000 (23)	Bladder cancer	<i>NAT2</i>	Smoking	16	Evidence of interaction.	Case-only odds ratio. Logistic regression within studies, combined using fixed- and random-effects meta-analysis.

Abbreviations: *GSTM1*, glutathione *S*-transferase M1; *GSTT1*, glutathione *S*-transferase T1; *MAOA*, monoamine oxidase A; *mEH3*, microsomal epoxide hydrolase 3; *mEH4*, microsomal epoxide hydrolase 4; *NAT1*, *N*-acetyltransferase 1; *NAT2*, *N*-acetyltransferase 2; *TGFA*, transforming growth factor α .

^a Cosetta Minelli, EURAC Research (Bolzano, Italy), personal communication, 2009.

and the strategies used for conducting these searches should relate directly to the eligibility criteria. The principal issue in a meta-analysis of GEI is whether to seek studies that prominently report interaction or to seek a larger set of studies and investigate the full texts of the articles for reports of GEI. A paper whose focus is reporting main genetic effects might include data relevant to GEI only within the results, tables, or text, so important results on GEI would not be identified by a search based on titles, abstracts, and keywords. Any interaction from such papers that is mentioned in an abstract could well represent an “exciting” finding, and results identified in this way could easily bias the results of a meta-analysis. Only a detailed examination of the full text of many papers would avoid the possibility of such bias due to selective reporting. For example, Raimondi et al. (14) used a collection of studies that had been identified for a previous systematic review of associations between smoking and colorectal cancer, and they examined each for information on interaction with genes.

In most of the papers in our selection of examples, investigators specified the search terms used but did not describe how the search strategy was structured, so we were unable to determine what they did from their published reports. Minelli et al. (unpublished manuscript) targeted reports of interaction, conducting a search of the form “(gene *or* synonyms) and (respiratory illness *or* synonyms) and (air pollution *or* synonyms).” In the 4 most recent of the 7 papers in our selection, the investigators searched multiple databases; in the oldest 2, they searched MEDLINE alone; and in the intermediate paper, they did not report their sources. Most investigators searched reference lists, but none reported major attempts to unearth less accessible studies.

Experiences from PPAR γ review. Like Raimondi et al. (14), we had access to results from a comprehensive search for a “parent” collection of association studies from a systematic review of association between PPAR γ Pro12Ala and type 2 diabetes (12). We examined the full text of each paper included in this review for reports of interaction between the genetic variant and lifestyle factors. We supplemented this list with additional searches for papers in 2 ways:

1. A search for GEIs involving PPAR γ and type 2 diabetes within HuGE Navigator (Human Genome Epidemiology Network) (20).
2. A broad PubMed search of the form “(PPAR γ *or* synonyms) and (Pro12Ala *or* synonyms) and (type 2 diabetes *or* hyperglycemia *or* synonyms)” and examination of the resulting titles and abstracts for mention of GEI.

We did not restrict any searches to terms related to specific study designs. The search and subsequent selection process yielded 33 reports from 30 independent studies.

Data collection

General considerations for GEIs. Information on GEI may be presented in a variety of ways, including an estimate of an interaction parameter from a regression model with a confidence interval or *P* value and estimates of the effect of exposure stratified by genotype. In general, the models estimating interactions may be very different depending on

the type of outcome (possibilities include linear regression, logistic regression, unweighted or weighted Cox regression, mixed-effects models, and generalized estimating equations) and also depending on what other covariates have been included in the model and how these covariates have been measured.

In addition to the variety of models that could be used to estimate an interaction, there are many ways in which interactions might be reported. Knol et al. (21) have studied this empirically, based on a sample of 150 case-control studies and 75 cohort studies published in leading journals between 2001 and 2007. They concluded that although a majority of the studies addressed possible interactions between exposures, in about half of these studies the information provided was unsatisfactory (21). Examples of “unsatisfactory” reporting included making statements about interactions without presenting any supporting data, presenting only a *P* value, and presenting stratum-specific estimates with no meaningful comparison between strata.

Given the variety of ways that interactions may be reported, only thorough scrutiny of all of the results, tables, and figures in each paper will ensure that all relevant information is collected for the review or meta-analysis. Beyond this, contact with the authors of the primary studies may be required, and in 5 of the 7 example meta-analyses, the investigators did this. Risch et al. (22) sought raw data from all included studies and obtained them for 10 of their 14 studies. In contrast, Marcus et al. (23) reported substantial attempts to retrieve usable data, without success, and changed their strategy as a result to examine only case groups rather than full case-control studies.

Experiences from PPAR γ review. We used the classification system proposed by Knol et al. (21) to categorize the 33 reports we found according to how the interactions were reported (see Table 2). In 8 (24%) of the 33 papers, investigators reported interactions at level 1 (only a *P* value or statement of statistical significance), and in 14 (42%) they reported them at level 2 (effect estimates and confidence intervals provided across strata of exposure for each genotype without a direct test for interaction). Therefore, in 67% of the papers, investigators did not present results from a meaningful comparison of stratum-specific effect estimates. In 11 (33%) papers, investigators reported interactions at level 3 (effect estimates and confidence intervals across strata and a *P* value or statement of statistical significance based on an interaction test), but in no papers did investigators report interactions at level 4 (sufficient information to allow interpretation of the interaction on both an additive and a multiplicative scale). In most cases, the conclusion about interaction was based on a finding of a statistically significant effect in 1 group (e.g., body mass index (weight (kg)/height (m)²) \geq 30) but no significant effect in the other group (e.g., body mass index < 30). These results are consistent with those reported by Knol et al. (21), and the heterogeneity in reporting undermines the potential application of meta-analysis in this instance.

Study quality and biases

General considerations for GEIs. In addition to the difficulties of identifying consistently reported GEI information

Table 2. Distribution of 33 Studies Containing Information About the Interaction Between the Peroxisome Proliferator-Activated Receptor γ (*PPAR γ*) Pro12Ala Variant and Lifestyle Factors for Type 2 Diabetes and Related Metabolic Traits, by Outcome-Exposure Combination and Study Design

Outcome	Exposure									
	Physical Activity		Diet		BMI ^a /Obesity		Hypertension		Other	
	All Designs	CS Design	All Designs	CS Design	All Designs	CS Design	All Designs	CS Design	All Designs	CS Design
Diabetes	2	0	3	0	5	0	1	0	0	0
Insulin ^b	2	0	4	3	7	3	1	0	3	0
Glucose ^c	1	0	2	2	6	2	1	0	2	0
BMI/anthropometry ^d	2	0	4	2	4	1	0	0	4	1
Other	0	0	0	0	1	1	0	0	2	0

Abbreviations: BMI, body mass index; CS, cross-sectional; HOMA-IR, homeostasis model assessment of insulin resistance.

^a Weight (kg)/height (m)².

^b Includes fasting insulin level, HOMA-IR, and HOMA75 (75th percentile of HOMA-IR distribution).

^c Includes fasting glucose level and hemoglobin A1c level.

^d Includes BMI, waist:hip ratio, and waist circumference.

from published papers, there are numerous biases that may affect a meta-analysis—both biases within individual studies and bias that may arise in the meta-analysis itself.

In general, studies in which GEIs are reported may be prone to all of the usual biases inherent in observational studies. This is one of the substantial complications in extending a meta-analysis of genetic association studies to examine GEI, since in genetic association studies Mendelian randomization of genetic variants at conception means that concerns about confounding are typically restricted to population stratification (when the studies are conducted in different ethnic groups or in admixed populations (24)). Therefore, there may be substantial bias arising from failure to control adequately for other confounding factors once environmental exposures are analyzed. There may also be bias arising from measurement error in both the environmental exposure and the confounders, as well as from genotype misclassification (25) and selection bias (26). These issues are typical of attempts to meta-analyze results from nongenetic, observational epidemiology studies (27, 28). The combined effect of all of these biases on the estimate of an interaction would be very difficult to predict (29).

Sanderson et al. (30) have reviewed tools for assessing risk of bias in observational studies which potentially could be adapted to also address GEI. Minelli et al. (31) have reviewed the quality of meta-analyses of genetic association studies and issued some recommendations for reporting it, with a view that study quality should be investigated as a cause of heterogeneity and highlighting that this is rarely done, partly because the important biases are not fully understood (7). It is perhaps not surprising, therefore, that in most of the example analyses listed in Table 1, investigators did not conduct a quality assessment of the included studies. Minelli et al. (unpublished manuscript) extensively assessed risks of bias, and Marcus et al. commented that “the validity of our results may be affected by confounding, misclassification, or other limitations of our data” (23, p. 465), but investigators in the other 5 meta-analyses apparently took the findings from the included studies at face value, although

in 2 of them the researchers collected data on some study characteristics that could be related to quality.

Experiences from PPAR γ review. In the *PPAR γ* review, we found 30 independent studies covering 7 study designs, including 11 cross-sectional studies, 6 case-control studies, 6 cohort studies, 3 clinical trials, 3 family-based studies, and 1 meta-analysis. No study used a case-only design. We did not undertake a formal quality assessment of the studies we identified, largely because we found very heterogeneous evidence, with little prospect of being able to draw useful conclusions even if all of the studies were free of limitations.

Biases in the review as a whole

General considerations for GEIs. Meta-analysis is well known for having the potential to suffer from publication bias—the tendency for “exciting” results to be more likely to be published than statistically nonsignificant results. This would seem to be an even greater issue where the interest is in interactions than where the interest is in main effects, given that interactions are not usually the primary hypothesis of a study. It is highly likely that researchers might study a number of interactions as secondary hypotheses but then only report those that are statistically significant. A standard set of tools, including funnel plots and associated statistical tests, is available for examining relations between study size and effect size (32) and these tools are often mistakenly used as tests for publication bias (investigators in 2 of our examples in Table 1 did this). Asymmetry in a funnel plot may be due to publication bias, but this is only 1 of several possible explanations. Among our examples, we observed very few comments about the possibility of publication bias or selective reporting of exciting findings. Minelli et al. (unpublished manuscript) provided an extensive discussion and were suspicious of selective reporting; Marcus et al. (23) acknowledged that publication bias could be responsible for their findings. The only true test of publication bias is to compare results from published studies with

results from comparable unpublished studies. In a literature-based meta-analysis, this is very unlikely to be feasible, so the problem might be regarded as insurmountable.

Experiences from PPAR γ review. Our strategy for addressing reporting biases had 2 aspects. First, by having access to a carefully assembled collection of studies on the association between PPAR γ and type 2 diabetes (12), we were able to examine the full text of most of the papers that could theoretically include information on interaction with diabetes. Ten of the 33 papers retrieved by our search had been included in the review of associations with PPAR γ main effects, and no relevant papers from the PPAR γ main-effects review had been missed by our search. Thus, in this instance, the bibliographic database search would have been sufficient. Second, we acknowledge that selective reporting of statistically significant interactions is likely in the published literature, and we interpret any findings with considerable skepticism.

Combining evidence across different exposures, outcomes, and study designs

General considerations for GEIs. In a typical meta-analysis, investigators proceed by computing a weighted average of comparable point estimates. Ideally, we would have estimates of the same interaction term from each of several studies examining the same genetic variant, the same exposure, and the same outcome, studied in a similar way. However, as we have already mentioned, several different designs may be used to examine interactions; authors may analyze data and report results in different ways; and there may be a variety of different outcomes and exposures. These differences will present challenges when attempting to meta-analyze the results across studies (29). The challenges involved in investigating interaction may be greater even than when addressing association in (nongenetic) epidemiologic studies (27, 28). Our evidence from other meta-analyses of published literature indicates that only in the presence of dichotomous or categorical environmental variables (e.g., smoking) was it possible to perform meta-analysis. With continuous environmental exposures (e.g., air pollution in Minelli et al. (unpublished manuscript)), the measures of environmental exposure were too heterogeneous for a quantitative synthesis to be feasible.

Advanced methods of synthesis can be used to combine information from a more complex data structure. Sanderson et al. (33) developed a Bayesian method specifically for their review to facilitate the meta-analysis of results from studies where interactions have been analyzed and reported in different or incomplete ways, making use of external information about the prevalences of the various unreported exposures in the populations under study.

Experiences from PPAR γ review. We grouped the possible outcomes into 5 classes (diabetes, glucose outcomes, insulin outcomes, anthropometric outcomes, and other) and grouped the environmental exposures similarly (dietary exposures, physical activity exposures, body mass index/obesity exposures, hypertension, and other). Table 2 shows how many of the papers were identified within each combination of outcome and exposure. The most frequent combi-

nation of outcome and exposure was an insulin-related outcome and body mass index/obesity. The most common study design was cross-sectional. When we classified outcome and exposure for the most frequent design class (cross-sectional), we found that each combination of outcome, exposure, and design was virtually unique (Table 2). For example, the combination of insulin, diet, and a cross-sectional design yielded 3 studies. However, even within this category, there were 3 somewhat different environmental exposures, namely polyunsaturated fat:saturated fat ratio, free fatty acids, and fish intake. Thus, the scope for meta-analysis was severely limited, and we did not attempt any.

DISCUSSION

The challenges of meta-analyzing GEIs using published literature are numerous. They include:

- 1) lack of consistency in both definition and measurement of outcomes and environmental exposures;
- 2) lack of consistency in study designs;
- 3) bias inherent in the observational studies included in the review;
- 4) lack of consistency in statistical analysis models and methods for assessing interactions;
- 5) lack of consistency in how results for interactions are presented and the likelihood that they will not be included in the title or abstract of a paper, thus requiring hand-searching of the full text of all possible papers on the gene of interest; and
- 6) the possibility that nonsignificant interactions have not been reported and that hence any quantitative meta-analysis will suffer from publication bias.

Challenges 1–3 are common to observational epidemiologic studies of main effects, and challenges 4–6 are particularly pertinent to the study of interactions. Collective efforts have been put into place to overcome such challenges; these efforts have particularly focused on guidelines for reporting of observational studies (34) and, more recently, genetic association studies (35). However, the additional challenges (challenges 4–6) entailed by interactions suggest that efforts to meta-analyze GEIs using published literature might be a poor use of time and resources. This is an area where meta-analysis based on the individual participant data from each study might be beneficial, and where prospective harmonization of data collection and analytic strategies is critical (36).

Many scientific consortia have been established to share and meta-analyze data for the assessment of main genetic effects (37), and such consortia are well placed to examine GEIs. Of course, these consortia are often established retrospectively, and because different definitions and measures of the environmental factors will have been used, their standardization will be necessary. Such an approach is likely to be limited to a few highly specific lifestyle behaviors, such as smoking, that are reasonably consistently and precisely measured across studies. The study of gene-lifestyle interaction for other, more difficult-to-measure exposures such

as diet and physical activity is likely to require analysis in consortia where standardization of these exposures has been instigated at the design stage. Moreover, because historical studies utilized relatively imprecise methods for exposure measurement, the study of interaction may also require the establishment of new cohorts with investment in precise measurement of important but difficult-to-measure behaviors.

ACKNOWLEDGMENTS

Author affiliations: MRC Epidemiology Unit, Institute of Metabolic Science, Cambridge University Hospital Foundation Trust, Cambridge, United Kingdom (Luigi Palla, Nicholas J. Wareham, Stephen J. Sharp); and MRC Biostatistics Unit, Institute of Public Health, University of Cambridge, Cambridge, United Kingdom (Julian P. T. Higgins).

This work was supported by funding awarded to the InterAct consortium through the European Union (Integrated Project LSHM-CT-2006-037197 in the Framework Programme 6 of the European Community). Dr. Julian P. T. Higgins was supported by MRC grant U.1052.00.011.

Conflict of interest: none declared.

REFERENCES

- Haldane JBS. The interaction of nature and nurture. *Ann Eugen.* 1946;13:197–205.
- Lynch M, Walsh B. *Genetics and Analysis of Quantitative Traits*. Sunderland, MA: Sinauer Associates, Inc; 1998.
- Vineis P, Kriebel D. Causal models in epidemiology: past inheritance and genetic future [electronic article]. *Environ Health.* 2006;5:21.
- Smith PG, Day NE. The design of case-control studies: the influence of confounding and interaction effects. *Int J Epidemiol.* 1984;13(3):356–365.
- Higgins JP, Little J, Ioannidis JP, et al. Turning the pump handle: evolving methods for integrating the evidence on gene-disease association. *Am J Epidemiol.* 2007;166(8):863–866.
- Hooper L, Ashton K, Harvey LJ, et al. Assessing potential biomarkers of micronutrient status by using a systematic review methodology: methods. *Am J Clin Nutr.* 2009;89(6):1953S–1959S.
- Sagoo GS, Little J, Higgins JP. Systematic reviews of genetic association studies. Human Genome Epidemiology Network [electronic article]. *PLoS Med.* 2009;6(3):e28.
- Florez JC. Phenotypic consequences of the peroxisome proliferator-activated receptor-gamma Pro12Ala polymorphism: the weight of the evidence in genetic association studies. *J Clin Endocrinol Metab.* 2004;89(9):4234–4237.
- Spiegelman BM. PPAR-gamma: adipogenic regulator and thiazolidinedione receptor. *Diabetes.* 1998;47(4):507–514.
- Yen CJ, Beamer BA, Negri C, et al. Molecular scanning of the human peroxisome proliferator activated receptor gamma (hPPAR gamma) gene in diabetic Caucasians: identification of a Pro12Ala PPAR gamma 2 missense mutation. *Biochem Biophys Res Commun.* 1997;241(2):270–274.
- Tönjes A, Scholz M, Loeffler M, et al. Association of Pro12Ala polymorphism in peroxisome proliferator-activated receptor gamma with pre-diabetic phenotypes: meta-analysis of 57 studies on nondiabetic individuals. *Diabetes Care.* 2006;29(11):2489–2497.
- Gouda HN, Sagoo GS, Harding AH, et al. The association between the peroxisome proliferator-activated receptor- γ 2 (PPARG2) Pro12Ala gene variant and type 2 diabetes mellitus: a HuGE review and meta-analysis. *Am J Epidemiol.* 2010;171(6):645–655.
- Franks PW, Mesa JL, Harding AH, et al. Gene-lifestyle interaction on risk of type 2 diabetes. *Nutr Metab Cardiovasc Dis.* 2007;17(2):104–124.
- Raimondi S, Botteri E, Iodice S, et al. Gene-smoking interaction on colorectal adenoma and cancer risk: review and meta-analysis. *Mutat Res.* 2009;670(1-2):6–14.
- Zeiger JS, Beaty TH, Liang KY. Oral clefts, maternal smoking, and TGFA: a meta-analysis of gene-environment interaction. *Cleft Palate Craniofac J.* 2005;42(1):58–63.
- Khoury MJ, Wacholder S. Invited commentary: from genome-wide association studies to gene-environment-wide interaction studies—challenges and opportunities. *Am J Epidemiol.* 2009;169(2):227–230.
- Clayton D, McKeigue PM. Epidemiological methods for studying genes and environmental factors in complex diseases. *Lancet.* 2001;358(9290):1358–1360.
- Khoury MJ, Beaty TH, Cohen BH. *Fundamentals of Genetic Epidemiology*. New York, NY: Oxford University Press; 1993.
- Yang Q, Khoury MJ. Evolving methods in genetic epidemiology. III. Gene-environment interaction in epidemiologic research. *Epidemiol Rev.* 1997;19(1):33–43.
- Yu W, Gwinn M, Clyne M, et al. A navigator for human genome epidemiology. *Nat Genet.* 2008;40(2):124–125.
- Knol MJ, Egger M, Scott P, et al. When one depends on the other: reporting of interaction in case-control and cohort studies. *Epidemiology.* 2009;20(2):161–166.
- Risch N, Herrell R, Lehner T, et al. Interaction between the serotonin transporter gene (5-HTTLPR), stressful life events, and risk of depression: a meta-analysis. *JAMA.* 2009;301(23):2462–2471.
- Marcus PM, Hayes RB, Vineis P, et al. Cigarette smoking, N-acetyltransferase 2 acetylation status, and bladder cancer risk: a case-series meta-analysis of a gene-environment interaction. *Cancer Epidemiol Biomarkers Prev.* 2000;9(5):461–467.
- Knowler WC, Williams RC, Pettitt DJ, et al. Gm3:5,13,14 and type 2 diabetes mellitus: an association in American Indians with genetic admixture. *Am J Hum Genet.* 1988;43(4):520–526.
- Garcia-Closas M, Rothman N, Lubin J. Misclassification in case-control studies of gene-environment interactions: assessment of bias and sample size. *Cancer Epidemiol Biomarkers Prev.* 1999;8(12):1043–1050.
- Morimoto LM, White E, Newcomb PA. Selection bias in the assessment of gene-environment interaction in case-control studies. *Am J Epidemiol.* 2003;158(3):259–263.
- Egger M, Schneider M, Davey Smith G. Spurious precision? Meta-analysis of observational studies. *BMJ.* 1998;316(7125):140–144.
- Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA.* 2000;283(15):2008–2012.
- Vineis P. Methodological approaches to gene-environment interactions in occupational epidemiology [electronic article]. *Occup Environ Med.* 2007;64:e3.

30. Sanderson S, Tatt ID, Higgins JP. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *Int J Epidemiol.* 2007;36(3):666–676.
31. Minelli C, Thompson JR, Abrams KR, et al. The quality of meta-analyses of genetic association studies: a review with recommendations. *Am J Epidemiol.* 2009;170(11):1333–1343.
32. Egger M, Davey Smith G, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. *BMJ.* 1997;315(7109):629–634.
33. Sanderson S, Salanti G, Higgins J. Joint effects of the *N*-acetyltransferase 1 and 2 (*NAT1* and *NAT2*) genes and smoking on bladder carcinogenesis: a literature-based systematic HuGE review and evidence synthesis. *Am J Epidemiol.* 2007;166(7):741–751.
34. Salanti G, Higgins JP, White IR. Bayesian synthesis of epidemiological evidence with different combinations of exposure groups: application to a gene-gene-environment interaction. *Stat Med.* 2006;25(24):4147–4163.
35. von Elm E, Moher D, Little J. Reporting genetic association studies: the STREGA statement. *Lancet.* 2009;374(9684):98–100.
36. Little J, Higgins JP, Ioannidis JP, et al. STrengthening the REporting of Genetic Association Studies (STREGA): an extension of the STROBE statement [electronic article]. *PLoS Med.* 2009;6(2):e22.
37. Knoppers BM, Fortier I, Legault D, et al. The Public Population Project in Genomics (P3G): a proof of concept? *Eur J Hum Genet.* 2008;16(6):664–665.
38. Seminara D, Khoury MJ, O'Brien TR, et al. The emergence of networks in human genome epidemiology: challenges and opportunities. *Epidemiology.* 2007;18(1):1–8.
39. Kim-Cohen J, Caspi A, Taylor A, et al. MAOA, maltreatment, and gene-environment interaction predicting children's mental health: new evidence and a meta-analysis. *Mol Psychiatry.* 2006;11(10):903–913.