CHAPTER 18.

Combining molecular and genetic data from different sources

Evangelia E. Ntzani, Muin J. Khoury, and John P. A. Ioannidis

Summary

The rapidly growing number of molecular epidemiology studies is providing an enormous, often multidimensional, body of evidence on the association of various disease outcomes and biomarkers. The testing and validation of statistical hypotheses in genetic and molecular epidemiology presents a major challenge requiring methodological rigor and analytical power. The non-replication of many genetic and other biomarker association studies suggests that there may be an abundance of spurious findings in the field. This chapter will discuss ways of combining evidence from different sources using meta-analysis methods. Research synthesis not only aims at producing a summary effect estimate for a specific biomarker, but also offers a unique opportunity for a meticulous attempt to critically appraise a research field, identify substantial differences between or within studies, and detect sources of bias. Systematic reviews and meta-analyses in human genome epidemiology are specifically discussed, as they comprise the bulk of the available evidence in molecular epidemiology where these methods have been applied to date. Considered here are issues regarding validity and interpretation in genetic association studies, as well as strategies for developing and integrating evidence through international consortia. Finally, there is a brief look at how combining data through meta-analysis may be applied in other areas of molecular epidemiology.

Introduction

The number of molecular epidemiology studies is constantly growing, and this trend is expected to accelerate (1–4), especially with improvements in genotyping technology that allow massive testing of genetic variants in minimal time and at a decreasing cost on a genome-wide association study platform (5–8). The number of potentially identifiable genetic markers, and the multitude of clinical outcomes to which these may be associated, make the testing and validation of statistical hypotheses in genetic and molecular epidemiology a task of unprecedented scale. Currently, more than 6000 original articles on human genome epidemiology findings are published annually, and the numbers are increasing (9,10).

Yet, there has been considerable concern about non-replication in gene-disease association studies (11–19) and other areas of molecular epidemiology. The combination of high-throughput genotyping, selective reporting, and exploratory statistical analyses in studies with limited sample sizes could potentially generate a scientific literature replete with spurious findings and lead to wasted resources, unless mechanisms are put in place to promptly evaluate evidence as it accumulates (20,21). Related concerns also apply to other fields of molecular epidemiology where large amounts of data are produced and it is important to achieve unbiased integration of the evidence.

Combining evidence from different sources is discussed here. The goal of research synthesis is to estimate and explain between-study heterogeneity, arrive at summary effects, and appraise the quality and reliability of the evidence procured by many studies on the same research question. Specifically, systematic reviews and meta-analyses in human genome epidemiology are discussed, as they comprise the bulk of the available evidence in molecular epidemiology where these methods have been applied. Issues regarding validity and interpretation in genetic association studies are considered, as well as strategies for developing and integrating high-quality genomic evidence through international consortia. Finally, means for applying combined data through meta-analysis in other areas of molecular epidemiology are briefly examined.

Systematic reviews and meta-analyses: Definitions

Systematic reviews and meta-analyses provide valuable tools for summarizing genetic effects and for identifying and explaining the underlying differences and observed discrepancies between studies. The term systematic review has been used as a contrast to traditional review. Systematic reviews use a predefined, structured approach to the collection and integration of available evidence, whereas traditional reviews offer a non-structured, non-standardized appraisal of the current literature distorted in varying degrees by the reviewer's personal opinion and experience. The goal of this systematic approach is to guarantee the transparency and completeness of the review process. Meta-analyses use quantitative research synthesis methodology to derive summary estimates of the studied effects and to describe and explain the variability between and within studies (22). Systematic reviews and meta-analyses are well-established approaches to research synthesis in clinical trials, where their strengths and limitations have been widely assessed (23). Increasingly, they have also been applied to observational studies (24); meta-analyses of observational data are currently as common as those of clinical trials. Meta-analyses of gene–disease association studies have been accepted as a key method for establishing the genetic components of complex diseases (14,17). In 1998, The Human Genome Epidemiology Network (HuGENet) was launched as a global collaboration of individuals and organizations interested in accelerating the development of the knowledge base on genetic variation and common diseases. HuGENet has promoted the publication of HuGE reviews as a means of integrating evidence from human genome epidemiologic studies, that is, population-based studies of the impact of human genetic variation on health and disease (25). Initial efforts to apply quantitative methods were cautious, but there is now wide agreement that a meta-analysis of the evidence is almost always indicated and can provide more useful insights than a simple narrative review, provided the caveats of data synthesis are properly recognized. By the end of 2009, approximately 1200 systematic reviews and meta-analyses had been published on human genome epidemiology topics (http://www.cdc.gov/genomics/hugenet/default.htm); most of them, however, tried to integrate information on only one or a few specific gene–disease associations at a time. The unknown extent of unpublished data and the likelihood of biases inherent in single studies threaten the credibility of genetic findings.

While most meta-analyses in the past have been retrospective exercises, there is an increasing interest for prospective collaborative analyses that use the same statistical methods as traditional retrospective meta-analyses. Collaborative meta-analyses may be undertaken by consortia or networks of investigators working on the same disease and/or set of research questions. Participating teams may combine already-collected data, perform projects that use both retrospectively and prospectively collected information, or develop new collaborative projects on a completely prospective basis. With the advent of genome-wide association studies (GWAS), it is common practice to immediately seek replication of proposed discovered associations by other teams of investigators and publish the combined data in the same article. As a more recent alternative, meta-analyses have been implemented by combining

multiple data sets at the discovery stage under a consortium umbrella (26–28). This is a prospective use of meta-analysis methods. Furthermore, for many diseases and research questions, numerous such coalitions of investigators may exist; bringing their data together presents a new field of application for meta-analysis methods.

Reviewing methods: Basic aspects

Recommendations have been developed regarding the conduct of systematic reviews and meta-analyses. In 2006, HuGENet posted online the first edition of a handbook for conducting HuGE reviews (29,30). The reporting of these studies may need further improvement and standardization in the literature and should become more evidence-based with increasing experience. Such standards may follow the examples of similar initiatives for genetic association studies (e.g. STREGA (31)), as well as other designs and disciplines (e.g. CONSORT (32,33), MOOSE (34), PRISMA (35), STARD (36) and TREND (37)).

First, typical retrospective systematic reviews and meta-analyses will be discussed. A typical systematic review includes the following stages: 1) formulation of the research question requiring appraisal of the available evidence, 2) identification of the eligible studies and data extraction, 3) synthesis of the available evidence, 4) assessing and addressing potential biases, and 5) interpreting the results.

Research questions

Formulating the research question is fundamental for systematic reviews, as for any other research endeavour. Decisions must be made upfront about which gene and variants and which disease and outcomes to assess, as well as the eligibility criteria for the study designs and the study and population characteristics. Different eligibility criteria may lead to different data being synthesized and possibly different conclusions.

Data

Identifying the studies eligible for inclusion in a systematic review requires comprehensive, systematic literature searches. One must specify which eligible databases to search, and decide whether to consider data without regard for their prior publication in peer-reviewed literature or the specific language(s) of publication (38). (For more details on issues pertaining to the eligibility and choice of sources of data, see (39).) Data extraction for published information is typically performed by two independent investigators with critical discussion of any discrepancies.

Data synthesis

Synthesizing the available evidence is best done in a quantitative way, producing summary estimates of the assessed effect and estimates of the between-study heterogeneity, as well as measures of the uncertainty thereof. A quantitative synthesis must be strongly encouraged, whenever feasible, as a means of producing a summary estimate, but most importantly for quantification of heterogeneity and identification of potential bias. Some key issues on methods for evaluation of between-data set heterogeneity and for obtaining summary effects will be touched on briefly; a discussion on issues of multivariate models and adjustments will follow.

Heterogeneity

One should distinguish between clinical, biological and statistical heterogeneity. Statistical heterogeneity can be tested in any quantitative synthesis. Its presence may signal genuine biological and clinical heterogeneity or bias and errors. Often it is difficult to pinpoint what the exact reasons are for heterogeneity, and inferences should be cautious. Conversely, the absence of demonstrable statistical heterogeneity cannot be interpreted as proof of clinical and biological homogeneity.

Several heterogeneity tests and metrics are traditionally used in meta-analyses. (For more details and mathematical formulas, see (39).) The Q statistic provides a χ^2-based test and is considered significant for $P < 0.10$, but it is still underpowered in most meta-analyses whenever there are few (roughly < 20) data sets combined (40).

The between-study variance, τ^2, is not commonly used as a metric of heterogeneity, because its magnitude depends on the respective effect size metric (e.g. standardized mean difference, odds ratio, hazard ratio) and it is not comparable among meta-analyses using different effect metrics (41). However, a useful metric often neglected is the ratio of τ over the effect size. This ratio can provide a measure of the extent of variability (between-study standard deviation) as compared with the effect size. Given that many molecular epidemiology effects are small, the relative magnitude of the uncertainty versus the effect is a useful measure to consider. The most popular metric for conveying between-study heterogeneity is nevertheless the I^2. This metric has the major advantage that it is independent of the number of studies, and

thus can be standardized for use across different meta-analyses with different effect metrics and different numbers of studies (40). I^2 is directly interpreted as the percentage of total variation across studies due to heterogeneity rather than chance, and it takes values between 0 and 100% (42). Values over 50% indicate large heterogeneity. However, I^2 also becomes uncertain when only a few studies are combined, as in the large majority of current meta-analyses (41,43), and therefore presentation of 95% confidence intervals should be considered routinely. This will help avoid spuriously strong inferences regarding heterogeneity or lack thereof.

Summary effects

To date, most meta-analyses have used either fixed or random effects methods for combining the data across eligible studies and data sets. Fixed effects models assume a common effect estimate for all studies and attribute all observed between-study variability to chance. Fixed effects models include inverse-variance weighting, and Mantel-Haenszel and Peto methods (44), and seem inappropriate in the presence of demonstrable or anticipated heterogeneity if used as the single methodology for the effect estimate calculation. In the absence of demonstrable heterogeneity, keep in mind that failure to reject the null hypothesis of homogeneity does not prove homogeneity. Random effects assume that there is a different underlying effect size for each study. There are many different proposed estimators for the between-study variance; the most popular was suggested by DerSimonian and Laird (45). Random effects accommodate between-study heterogeneity and thus should be preferred in the presence or anticipation of heterogeneity. In the absence of any heterogeneity, fixed and random effects give similar results in any case.

Unfortunately, these issues are not yet well understood in the literature, as shown by empirical evaluations of candidate gene meta-analyses and also meta-analyses of GWAS (46,47). Until recently, the choice of model for combining results from candidate gene studies lay on the straightforward concept of underlying heterogeneity. Nevertheless, in a GWAS setting, the presence of heterogeneity may not necessarily correspond to replication failure, but can signal difficulty in extending the probed association in diverse populations (48). In light of the generally limited power to detect moderate signals of effect at the discovery stage, the exclusive use of random effects models, and the more conservative confidence intervals produced when heterogeneity is present would result in forbidding possibly true signals to pass the genome-wide significance threshold and seek further replication however large the discovery data sets might be (power desert phenomenon) (49). Thus, it would be more appropriate to report the results from both models and make critical decisions on the basis of the stage at which meta-analysis is performed.

Besides traditional fixed and random effects models, there is an increasing application of more fully Bayesian methods in meta-analysis. Their discussion is beyond the scope of this chapter, but the interested reader is referred to a reference textbook (50) and the WinBUGS software manual (51).

Adjustments for other variables

Both adjusted and unadjusted effect estimates from single studies may be combined in meta-analyses. Questionnaire-based data are used to some extent to adjust effects estimates, including minimum information, such as age and sex, or more complex data, such as clinical features of the disease under study defining a potentially differentiating risk profile, where genetic or other molecular information could add additional information (52). This is more likely to be the case in large multicenter clinical trials or cohort studies, where a “nested” genetic association study is performed.

An issue with adjusted estimates is to ensure that similar or at least comparable adjustments have been performed across different studies. For retrospective efforts there is usually large variability in the types of adjustments. Moreover, even data on the same variables may have been collected across different studies using different questionnaires or procedures, and standardization may be difficult or even impossible. Finally, caution should be used when differentiating between variables that are independent predictors and others that may be surrogates of the genetic/molecular effect under study.

Assessing and addressing potential biases

There are often considerable and justifiable concerns regarding the quality and validity of molecular epidemiology studies. Critical appraisal of the studies included in a systematic review is of paramount importance for identifying the sources of bias inherent in each study. The types of biases include selection bias, information bias and confounding. Moreover, issues such as multiple testing should be considered, as well as concerns pertinent to specific types of biomarkers and studies (e.g. Hardy–

Weinberg equilibrium violations for genetic association studies). Appraisal of potential biases is often hampered by poor reporting of the primary studies. Poor reporting of observational studies (53) is a common challenge in synthesizing evidence; statements about STrengthening the Reporting of OBservational studies in Epidemiology (STROBE) (54–56), and STrengthening the REporting of Genetic Associations (STREGA) (31), a similar effort in human genome epidemiology sponsored by the Human Genome Epidemiology Network (HuGENet), contribute to the transparency of reporting and the prompt identification of potential sources of study discrepancies and bias.

Detailed discussion of the specific biases that may be encountered in single studies is beyond the scope of this chapter. Some suggested references follow for the interested reader: selection bias (57–65); information bias involving biomarker measurement (e.g. genotyping), capture of environmental factors, or outcome assessment (8,66–69); and confounding which for genetic-association studies in particular manifests primarily through population stratification (70–75). For genome-wide investigations and other massive testing approaches, even minor biases on any of these fronts may create some highly statistically significant spurious signals among the many thousands being probed. Therefore, careful selection of cases and controls, high standards of genotyping and quality control, and routine use of principal component analysis, genomic control, family-based design, or other techniques that more rigorously control for stratification, are indispensable.

Interpreting the results

Interpreting the results of a systematic review and meta-analysis on an assessed biomarker includes consideration of the quantity and quality of the evidence and rigorous scrutiny for publication bias and selective reporting in the field at large. In terms of the quantity of accumulated evidence, it is unclear how much genetic information would be sufficient to validate a genetic association. Empirical evidence has demonstrated that initial research publications often fail to predict the subsequently established genetic effects and may even show substantial discrepancies with later research (14,76).

Publication and selective reporting bias

The tendency to publish studies with positive rather than negative results (preferring studies with large effects or statistically significant results) introduces publication bias (13,34). Publication bias is very difficult to address in a retrospective collection of published evidence. Tests such as funnel plots are notoriously unreliable and subjective, and they should be abandoned. Even formal statistical testing for funnel plot asymmetry cannot fully discriminate between publication bias and other sources of bias or genuine heterogeneity. In addition, the tests are generally underpowered (77–79) and subject to extensive limitations that make them useful only in a few meta-analyses (80). If these tests are employed, a suitable modified regression test should be selected that has appropriate type 1 error properties (81). Such tests would be more correctly called tests for small study effects, since they essentially evaluate whether small studies differ in their results from larger ones.

Another common issue that could have an increasingly important impact in molecular epidemiology is selective reporting of specific analyses and outcomes among the many that may be performed, often in pursuit of nominal statistical significance (82–85). Ideally, straightforward *a priori* hypotheses should be explicitly reported, and study objectives and future analyses should be documented at their outset under a collaborative initiative (20,31). However, this may not be as transparent as it should be, and lack of transparency is compounded by the exploratory nature of much molecular epidemiological research. A meta-analysis diagnostic that can be used to evaluate the presence of "significance-chasing" biases, including publication and other selective reporting biases, has been proposed (86). The test is most useful for application across many meta-analyses (e.g. evaluation of large research fields), while it is expected to be underpowered for meta-analyses with few studies.

Causal inference

An observed association may be spurious or real. Spurious associations may be due to chance, bias within studies, or bias across studies (reporting biases affecting the whole research field of interest). For genetics of common diseases, real associations, not attributable to confounding, may be due to a direct causal variant or to a variant in linkage disequilibrium (LD) with a direct causal variant (13,87,88). They can be a source of the heterogeneity found between studies of gene-disease associations.

Traditional epidemiological criteria for establishing causation include consistency, strength, biological plausibility (including analogy), dose–response, temporality, experimental

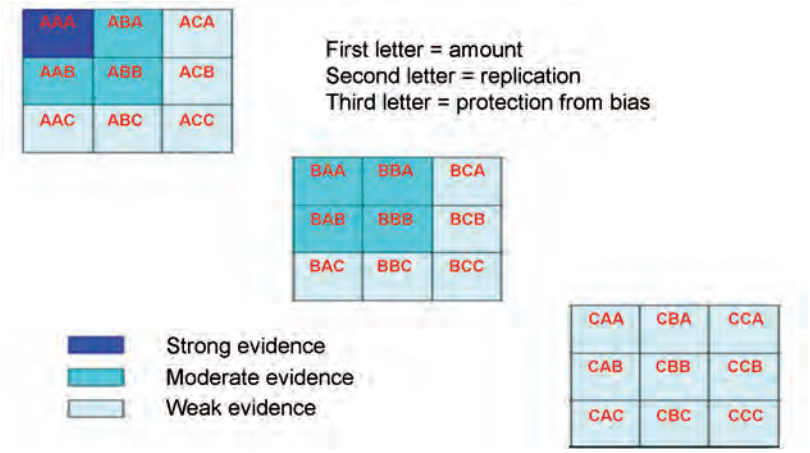
support, and coherence (89,90). Nevertheless, rarely are all of these issues taken into account, and the last three are not really relevant to human genome epidemiology. In genetic epidemiology, replication as an expression of consistency has received the greatest attention (13,14,17). Strength would be difficult to assess, as genetic effects are generally modest with odds ratios below two or even below 1.5 (91). Furthermore, the size of an effect is a characteristic of the genetic association being studied rather than a biologically consistent feature, as it depends on the relative prevalence of other causes (92).

In theory, biological plausibility should be an important criterion for causation, bringing under the same denominator epidemiologic evidence and diverse forms of biological evidence (93–99). Biological data on gene function, and on the tissue(s) in which a gene is expressed, could contribute to making a causal inference about gene-disease associations. On the other hand, there is concern that a biological argument can be constructed for virtually any associated allele because of the "...relative paucity of current understanding of the mechanisms of action of complex trait loci." (11). Thus, some form of mechanistic evidence might be identified and (mis)used selectively to reinforce an assertion of causality. Empirical evidence suggests very low agreement between biological and epidemiological evidence for common genetic variants and complex diseases (100). While candidate gene studies are often based on some biological knowledge of the candidate gene, genome-wide linkage and association studies initially identify variants without consideration of their biological function. Yet, the absence of mechanistic evidence or evidence of high quality would not exclude inferring that an association is causal if other guidelines for causation are satisfied. As knowledge of the genome is incomplete, biological plausibility may not always be apparent (97,101-103).

Criteria for assessing cumulative evidence

For genetic associations, a consensus approach recently developed interim guidelines for grading of the cumulative epidemiological evidence (104). The grading considers three aspects (known as the Venice criteria): amount of evidence, consistency of replication, and protection from bias (Table 18.1, Figure 18.1). Particularly for retrospective meta-analyses, protection from bias cannot be assumed if: the effect size is small (odds ratio deviating less than 1.15 from the null), the summary results lose their formal statistical significance when the first study that proposed the association is removed or when Hardy–Weinberg equilibrium-violating studies are removed, there are strong signals of small-study effects (e.g. a significant modified regression test) or significance-chasing bias (as discussed above), or if there are other demonstrable major biases in any aspect. Additional, yet weaker, signals for potential bias would be unclear/misclassified phenotypes with possible differential misclassification against genotyping or vice versa, major concerns for population stratification, or any other reason (case-by-case basis) that would jeopardise the validity of the proposed association. For prospective consortium-endorsed meta-analyses, all of the above parameters must be taken into consideration with the exception of small effect size, small study effects, and significance-chasing bias based on the basic assumption that selective reporting bias is not operating in the field.

Figure 18.1. Categories for the credibility of cumulative epidemiological evidence. The three letters correspond (in order) to amount of evidence, replication and protection from bias. Evidence is categorized as strong when there is A for all three items, and is categorized as weak when there is a C for any of the three items. All other combinations are categorized as moderate



Source: (104). Reproduced with permission of Oxford University Press.

Table 18.1. Considerations for epidemiologic credibility in the assessment of cumulative evidence on genetic associations

Criteria	Categories	Proposed operationalization
Amount of evidence	A Large-scale evidence B Moderate amount of evidence C Little evidence	Thresholds may be defined based on sample size, power, or false discovery rate considerations. The frequency of the genetic variant of interest should be accounted for. As a simple rule, it is suggested that category A contains over 1000 subjects (total number of cases and controls assuming 1:1 ratio) evaluated in the least common genetic group of interest, B corresponds to 100–1000 subjects evaluated in this group, and C corresponds to <100 subjects evaluated in this group.*
Replication	A Extensive replication including at least one well-conducted meta-analysis with little between-study inconsistency B Well-conducted meta-analysis with some methodological limitations or moderate between-study inconsistency C No association, no independent replication, failed replication, scattered studies, flawed meta-analysis, or large inconsistency	Between-study inconsistency entails statistical considerations (e.g. defined by metrics such as I^2, where values of ≥50% are considered large, and values of 25–50% are considered moderate inconsistency) and also epidemiological considerations for the similarity/standardization, or at least harmonization, of phenotyping, genotyping, and analytical models across studies.
Protection from bias	A Bias, if at all present, could affect the magnitude but probably not the presence of the association B No obvious bias that may affect the presence of the association, but there is considerable missing information on the generation of evidence C Considerable potential for or demonstrable bias that can affect even the presence or not of the association	A prerequisite for A, is that the bias due to phenotype measurement, genotype measurement, confounding (population stratification), and selective reporting (for meta-analyses) can be appraised as not being high, plus there is no other demonstrable bias in any other aspect of the design, analysis, or accumulation of the evidence that could invalidate the presence of the proposed association. In category B, although no strong biases are visible, there is no such assurance that major sources of bias have been minimized or accounted for, as information is missing on how phenotyping, genotyping and confounding have been handled. Given that occult bias can never be ruled out completely, note that even in category A the qualifier "probably" is used.

* For example, if the association pertains to the presence of homozygosity for a common variant and if the frequency of homozygosity is 3%, then category A amount of evidence requires over 30 000 subjects, and category B between 3000 and 30 000.
Adapted from (104)

Networks in human genome epidemiology

Although meta-analyses of published data provide a mechanism for combining evidence from different sources, they cannot overcome methodological flaws originating from the primary studies. An alternative approach that may also help improve the quality of the primary data is a meta-analysis of individual participant data (MIPD), which involves collecting and analysing detailed data on individual subjects and, ideally, prospective meta-analysis of data collected from consortia of investigators (105).

Meta-analysis of individual participant data (MIPD)

The MIPD may offer some advantages over the meta-analysis of published data. In theory these advantages include: standardization of definitions of cases, molecular markers and other variables of interest, enhanced ability to contact meta-analysis of time-to-event outcomes, testing of the assumptions of time-to-event models, better control of confounding, standardized multivariable and adjusted analyses, consistent treatment of subpopulations, and assessment of sampling bias. Not every one of these advantages may be relevant in all MIPD applications and some may be impossible. For example, when studies have already been established with specific case definitions, it may not be possible to go back and achieve perfect standardization of definitions across all studies, or some adjusting variables may have been collected only in some of the studies but not

others. Furthermore, an MIPD is far more labour-intensive and time-consuming than a meta-analysis of published data and may remain a retrospective effort (106).

Consortia and prospective collaborative efforts

An increasing number of consortia of investigators have been operating in molecular epidemiology. The value of such collaborative multicentre studies has long been recognized by epidemiologists for tackling important questions that are beyond the scope of a study at a single institution (107). Collaboration is of even greater significance in human genome epidemiology, due to the intrinsic characteristics of the field that can be better addressed through collaborative efforts (108): small sample sizes, weak expected genetic effects, genotype frequency variation in populations of different ethnic origin, and publication/selective reporting bias. Networks of scientists from multiple institutions can cooperate in research efforts involving, but not limited to, the conduct, analysis and synthesis of information from multiple population studies (3,20).

HuGENet has launched a global network of consortia working on human genome epidemiology, aimed at coordinating different research teams working on the same theme (109,110). The goal of the HuGENet Network of Investigator Networks is to create a resource to share information, offer methodological support, generate inclusive synopses of studies conducted in specific fields, and to facilitate rapid confirmation of findings. As of this writing, consortia in the Network of Investigator Networks comprise between five and more than 1000 teams, with accumulated sample sizes ranging from 3000 to over a half-million participants. Many other new consortia are continuously being developed. In particular for GWAS, it has become standard practice to try to replicate the derived associations across several other replicating teams as part of the first article to be published on a new proposed association (48). The replicating teams may already belong to an established consortium. Alternatively, their assembly may occur on an opportunity basis, but this may also form a nucleus for further collaborations. Besides choosing research targets based on agnostic massive testing approaches, other targets selected for study by consortia may be chosen based on *a priori* biological plausibility, supporting linkage evidence from genome-wide data, a perception of potentially high population risk (e.g. a common polymorphism), the number and consistency of published reports for a specific molecular marker, or a high-profile controversy in the literature (111,112). Also, consortia are increasingly being used to replicate associations derived from genome-wide association approaches independently from the first article that describes and partially replicates the associations (113).

Standardization issues

Members of consortia may share both prospective and retrospective features in the study design and accumulation of information. Standardization is one of the more significant benefits of consortia initiatives. Coordinating centres receive the incoming data, including both genotype and phenotype information, and guarantee adequate quality and transparency. Data standardization is best implemented at the beginning of a *de novo* collaborative study, while developing tools for data collection and definition of data items, and should achieve agreement on common data definitions to which all data layers must conform (114). Nevertheless, it may be difficult to achieve complete standardization if some data are already available. In this situation, consortia should still aim to maximize harmonization of data obtained from different sources.

Standardization or harmonization is crucial in order for a network to perform better than single studies. These processes increase the credibility of the derived evidence even when non-genetic measurements are difficult to standardise across teams. One criterion for the influence and success of a network may be its ability to adopt standards for phenotypes and covariates to establish the use of consistent definitions in subsequent studies. Standardization of genotypes, on the other hand, is usually achieved through central genotyping of all samples (115). Quality control of genotype results is typically straightforward, but additional checks are required in a multiteam collaboration. In the absence of central quality control, consortia may depend on post-hoc analyses, such as deviation from Hardy–Weinberg equilibrium (116) in the controls, to identify possible genotyping (or other) errors. Although large between-study heterogeneity in the final analyses may reflect measurement errors, these methods may still miss sizeable errors and their sensitivity and specificity are uncertain.

Meta-analyses of genome-wide association studies

As mentioned above, for many diseases several GWAS are performed and each may be

accompanied by replication efforts by several other teams (117,118). These studies may have used different platforms, but it is still possible to combine data for markers that are in perfect or almost-perfect LD with a correlation coefficient r^2 approximating 1.00 (119,120). Examples of meta-analyses of several GWAS are available in the early literature (121,122). Meta-analysis is currently considered standard practice for a GWAS setting (123–126) (Figure 18.2). Apart from using meta-analysis in a sequential, multiple-stage manner to continuously update, refute or replicate association signals, it can also be implemented early on at the discovery stage by combining multiple data sets under a consortium umbrella, thus augmenting power to detect signals for subsequent replication (15,127).

Heterogeneity in the genome-wide association setting, where massive testing of agnostic (rather than candidate) markers takes place, has some special features. As previously mentioned, besides bias and errors, the possibility of genuine heterogeneity must be seriously considered, due to differential LD for the culprit gene variant and heterogeneity due to association with correlated phenotypes across the populations enrolled in different studies being combined (47).

Figure 18.2. Typical work flow for conducting a meta-analysis of GWA data sets

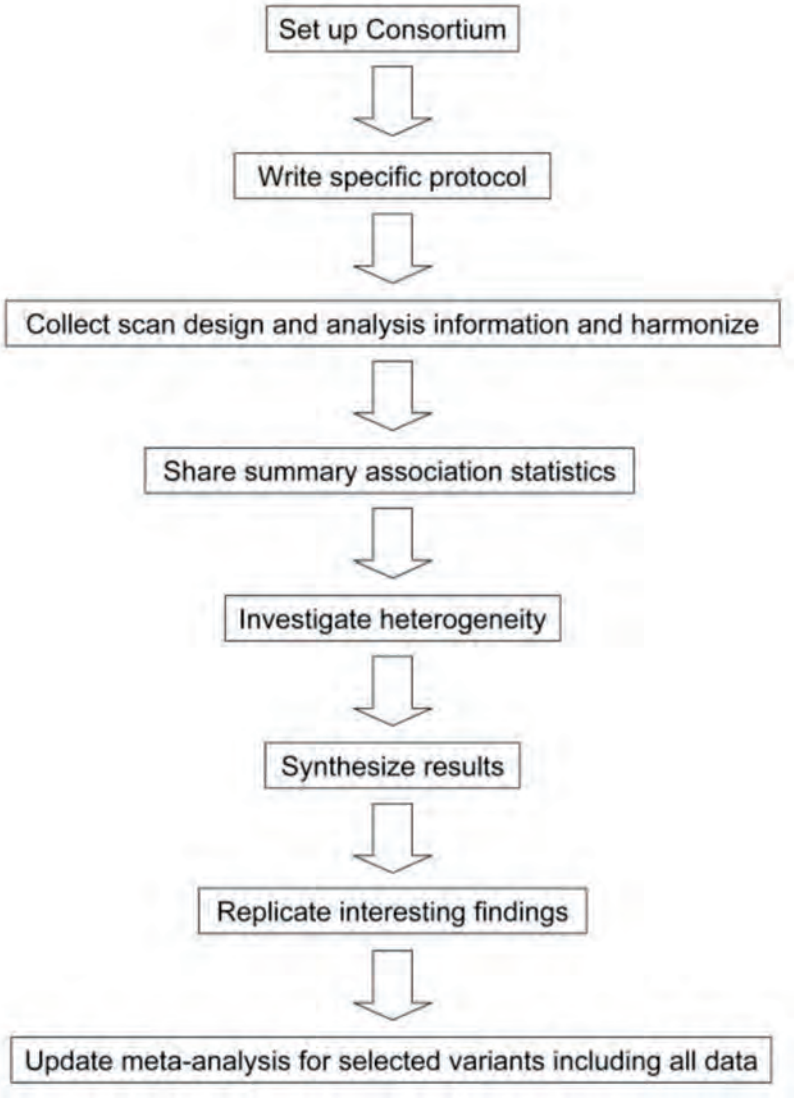


Figure compiled from (104).

Other applications of meta-analysis

Many fields of molecular epidemiology are characterized by large data sets that can be generated easily, due to the availability of sophisticated, low-cost technology. These data sets, derived from linkage scans, microarray-based gene expression profiling, mass spectra-based proteomics and many other massive testing platforms, usually capture information on hundreds of thousands of biological variables from a sometimes limited number of samples. To maximize the power to detect genuine signals requires combining data sets across different studies. However, combining data poses a further challenge, since the available data sets may have been obtained with different experimental conditions, platforms, analysis techniques or even sample types (e.g. different tissue, treatment conditions, or species). Meta-analysis could provide an appropriate framework for large data set synthesis. A few of these meta-analysis applications are mentioned here, but these are only indicative and the list is continuously expanding. Also discussed briefly are some issues that arise in the combination of information on other non-genetic biomarkers.

Meta-analysis of linkage signals in genome scans

Many teams of investigators have performed genome scans evaluating linkage between specific chromosomal loci and specific complex diseases (128–130). However, low linkage signals (131,132) and discrepancies in the findings of different teams often make the available evidence on a quantitative trait extremely difficult to summarize. Genome scan meta-analysis (GSMA) has been used as a method for summation of data from diverse genome scans through meta-analysis (131), and for formally testing whether the heterogeneity for specific chromosomal loci across genome scans (heterogeneity-based genome search meta-analysis (HEGESMA)) is large or small (133–134).

Microarrays and other multidimensional biology platforms

For various diseases, microarray platforms allow assessing differential expression of a large subset of genes. Research groups have approached the issue of synthesis across different platforms from different methodological perspectives (135–139). Significant computational power, multiple testing assumptions, and appropriate incorporation of heterogeneity estimates are only a few of the more challenging methodological issues. Given the small sample sizes of most microarray experiments and the complexity of the signals from single biological factors, meta-analysis may prove to be a very useful approach. Some non-parametric meta-analysis methods may allow synthesizing data from diverse platforms and different types of multidimensional data (140–142).

Meta-analyses of non-genetic prognostic markers

Besides the very large literature on genetic markers, there is also a burgeoning literature on non-genetic biomarkers. Single prognostic molecular markers, or combinations thereof, are still often considered in prognostic and predictive analyses for various clinical outcomes, such as mortality or other disease outcomes. Estrogen and other hormones, nutritional and related biochemical markers, and lipid or DNA adduct biomarkers are some of the commonly encountered examples in the literature (143–146).

Pertinent research synthesis methodology includes meta-analysis models as described above for genetic risk factors. Some of these predictors may be continuous variables, but the meta-analysis methods for combining information are very similar to the methods for combining data from binary markers (for details see (44)). Adjustment for covariates is more common in this literature, and may present problems related to the standardization of multivariate models and adjustments across the studies to be combined. Lack of standardization of biomarker measurements tends to be a more prominent problem than for genetic biomarkers, and error rates are expected to be larger and more variable across studies. Otherwise, heterogeneity testing and bias detection follow largely the same principles as described above for genetic markers.

Empirical evidence has shown that readily accessible published data can be misleading, producing a view of the literature that is distorted in a positive direction. An empirical evaluation has shown that almost all published prognostic marker studies on cancer report statistically significant results (147). Another empirical evaluation has shown that after standardising the definitions for the prognostic marker and the outcome under study, and, more importantly, after retrieving additional information that is unpublished or mentioned in only a cursory fashion in published articles, the statistical significance and predictive effect of a postulated prognostic/predictive factor may be abrogated (148).

In all, readily available information on prognostic factors may be the tip of the iceberg, and thus superficial perusal of the literature can lead to erroneous conclusions. This is yet another instance where selective reporting may spuriously inflate the importance of postulated prognostic factors unless retrieval of information and standardization of definitions in the literature are optimized. Meta-analyses of prognostic factors are likely to benefit from efforts to improve the conduct and reporting of primary studies, as exemplified by the REporting recommendations for tumour MARKer prognostic studies (REMARK) statement for tumour prognostic markers (149).

References

1. McPherson JD, Marra M, Hillier L *et al.*; International Human Genome Mapping Consortium (2001). A physical map of the human genome. *Nature*, 409:934–941.doi:10.1038/35057157 PMID:11237014

2. Guttmacher AE, Collins FS (2003). Welcome to the genomic era. *N Engl J Med,* 349:996–998.doi:10.1056/NEJMe038132 PMID:12954750

3. Little J, Khoury MJ, Bradley L *et al.* (2003). The human genome project is complete. How do we develop a handle for the pump? *Am J Epidemiol,* 157:667–673.doi:10.1093/aje/kwg048 PMID:12697570

4. Bell J (2004). Predicting disease using genomics. *Nature*, 429:453–456.doi:10.1038/nature02624 PMID:15164070

5. Lawrence RW, Evans DM, Cardon LR (2005). Prospects and pitfalls in whole genome association studies. *Philos Trans R Soc Lond B Biol Sci,* 360:1589–1595.doi:10.1098/rstb.2005.1689 PMID:16096108

6. Manolio TA, Collins FS, Cox NJ *et al.* (2009). Finding the missing heritability of complex diseases. *Nature*, 461:747–753.doi:10.1038/nature08494 PMID:19812666

7. Marchini J, Donnelly P, Cardon LR (2005). Genome-wide strategies for detecting multiple loci that influence complex diseases. *Nat Genet,* 37:413–417.doi:10.1038/ng1537 PMID:15793588

8. Thomas DC, Haile RW, Duggan D (2005). Recent developments in genomewide association scans: a workshop summary and review. *Am J Hum Genet,* 77:337–345.doi:10.1086/432962 PMID:16080110

9. Lin BK, Clyne M, Walsh M *et al.* (2006). Tracking the epidemiology of human genes in the literature: the HuGE Published Literature database. *Am J Epidemiol,* 164:1–4.doi:10.1093/aje/kwj175 PMID:16641305

10. Yu W, Wulf A, Yesupriya A *et al.* (2008). HuGE Watch: tracking trends and patterns of published studies of genetic association and human genome epidemiology in near-real time. *Eur J Hum Genet*, 16:1155–1158.doi:10.1038/ejhg.2008.95 PMID:18478035

11. Cardon LR, Bell JI (2001). Association study designs for complex diseases. *Nat Rev Genet,* 2:91–99.doi:10.1038/35052543 PMID:11253062

12. Gambaro G, Anglani F, D'Angelo A (2000). Association studies of genetic polymorphisms and complex disease. *Lancet*, 355:308–311.doi:10.1016/S0140-6736(99)07202-5 PMID:10675088

13. Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K (2002). A comprehensive review of genetic association studies. *Genet Med,* 4:45–61.doi:10.1097/00125817-200203000-00002 PMID:11882781

14. Ioannidis JPA, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG (2001). Replication validity of genetic association studies. *Nat Genet*, 29:306–309.doi:10.1038/ng749 PMID:11600885

15. Ioannidis JP (2009). Population-wide generalizability of genome-wide discovered associations. *J Natl Cancer Inst,* 101:1297–1299.doi:10.1093/jnci/djp298 PMID:19726754

16. No authors listed. The Lancet (2003). In search of genetic precision. *Lancet*, 361:357.doi:10.1016/S0140-6736(03)12433-6 PMID:12573365

17. Lohmueller KE, Pearce CL, Pike M *et al.* (2003). Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet,* 33:177–182.doi:10.1038/ng1071 PMID:12524541

18. Editorial (1999). Freely associating. *Nat Genet,* 22:1–2.doi:10.1038/8702 PMID:10319845

19. Tabor HK, Risch NJ, Myers RM (2002). Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nat Rev Genet,* 3:391–397.doi:10.1038/nrg796 PMID:11988764

20. Ioannidis JP, Bernstein J, Boffetta P *et al.* (2005). A network of investigator networks in human genome epidemiology. *Am J Epidemiol,* 162:302–304.doi:10.1093/aje/kwi201 PMID:16014777

21. Ioannidis JP (2007). Non-replication and inconsistency in the genome-wide association setting. *Hum Hered,* 64:203–213.doi:10.1159/000103512 PMID:17551261

22. Lau J, Ioannidis JP, Schmid CH (1997). Quantitative synthesis in systematic reviews. Ann Intern Med, 127:820–826. PMID:9382404

23. Lau J, Ioannidis JP, Schmid CH (1998). Summing up evidence: one answer is not always enough. *Lancet*, 351:123–127.doi:10.1016/S0140-6736(97)08468-7 PMID:9439507

24. Stroup DF, Berlin JA, Morton SC *et al.* for the Meta-analysis of Observational Studies in Epidemiology (MOOSE) Group (2000). Meta-analysis of observational studies in epidemiology: a proposal for reporting. *JAMA*, 283:2008–2012.doi:10.1001/jama.283.15.2008 PMID:10789670

25. Khoury MJ, Dorman JS (1998). The Human Genome Epidemiology Network. *Am J Epidemiol,* 148:1–3. PMID:9663396

26. Easton DF, Pooley KA, Dunning AM *et al.*; SEARCH collaborators; kConFab; AOCS Management Group (2007). Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature*, 447:1087–1093.doi:10.1038/nature05887 PMID:17529967

27. Ioannidis JP, Thomas G, Daly MJ (2009). Validating, augmenting and refining genome-wide association signals. *Nat Rev Genet,* 10:318–329.doi:10.1038/nrg2544 PMID:19373277

28. Khoury MJ, Bertram L, Boffetta P *et al.* (2009). Genome-wide association studies, field synopses, and the development of the knowledge base on genetic variation and human diseases. *Am J Epidemiol,* 170:269–279.doi:10.1093/aje/kwp119 PMID:19498075

29. Little J, Higgins J, editors. The HuGENet™ HuGE review handbook, version 1.0; 2006. Available from URL: http://www.medicine.uottawa.ca/public-health-genomics/web/assets/documents/HuGE_Review_Handbook_V1_0.pdf.

30. Higgins JP, Little J, Ioannidis JP *et al.* (2007). Turning the pump handle: evolving methods for integrating the evidence on gene-disease association. *Am J Epidemiol,* 166:863–866.doi:10.1093/aje/kwm248 PMID:17804859

31. Little J, Higgins JP, Ioannidis JP *et al.*; STrengthening the REporting of Genetic Association Studies (2009). STrengthening the REporting of Genetic Association Studies (STREGA): an extension of the STROBE statement. *PLoS Med,* 6:e22.doi:10.1371/journal.pmed.1000022 PMID:19192942

32. Altman DG, Schulz KF, Moher D *et al.*; CONSORT GROUP (Consolidated Standards of Reporting Trials) (2001). The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med,* 134:663–694. PMID:11304107

33. Moher D, Schulz KF, Altman D; CONSORT Group (Consolidated Standards of Reporting Trials) (2001). The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA*, 285:1987–1991.doi:10.1001/jama.285.15.1987 PMID:11308435

34. Stroup DF, Thacker SB. Meta-analysis in epidemiology. In: Gail MH, Benichou J, editors. Encyclopedia of epidemiologic methods. Chichester (UK): Wiley & Sons Publishers; 2000. p. 557–570

35. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ*, 339 jul21 1;b2535.doi:10.1136/bmj.b2535 PMID:19622551

36. Bossuyt PM, Reitsma JB, Bruns DE *et al.*; Standards for Reporting of Diagnostic Accuracy (2003). Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. *Ann Intern Med,* 138:40–44. PMID:12513043

37. Des Jarlais DC, Lyles C, Crepaz N; TREND Group (2004). Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. *Am J Public Health*, 94:361–366. doi:10.2105/AJPH.94.3.361 PMID:14998794

38. Pan Z, Trikalinos TA, Kavvoura FK *et al.* (2005). Local literature bias in genetic epidemiology: an empirical evaluation of the Chinese literature. *PLoS Med,* 2:e334.doi:10.1371/journal.pmed.0020334 PMID:16285839

39. Kavvoura FK, Ioannidis JP (2008). Methods for meta-analysis in genetic association studies: a review of their potential and pitfalls. *Hum Genet,* 123:1–14.doi:10.1007/s00439-007-0445-9 PMID:18026754

40. Higgins JP, Thompson SG (2002). Quantifying heterogeneity in a meta-analysis. *Stat Med,* 21:1539–1558.doi:10.1002/sim.1186 PMID:12111919

41. Huedo-Medina TB, Sánchez-Meca J, Marín-Martínez F, Botella J (2006). Assessing heterogeneity in meta-analysis: Q statistic or I2 index? *Psychol Methods*, 11:193–206.doi:10.1037/1082-989X.11.2.193 PMID:16784338

42. Higgins JP, Thompson SG, Deeks JJ, Altman DG (2003). Measuring inconsistency in meta-analyses. *BMJ*, 327:557–560. doi:10.1136/bmj.327.7414.557 PMID:12958120

43. Ioannidis JP, Patsopoulos NA, Evangelou E (2007). Uncertainty in estimates of heterogeneity in meta-analysis. *BMJ*, 335:914–916. doi:10.1136/bmj.39343.408449.80 PMID:17974687

44. Sutton AJ, Abrams KR, Jones DR *et al.*, editors. Methods for meta-analysis in medical research. Chichester (UK): John Wiley & Sons; 2000.

45. DerSimonian R, Laird N (1986). Meta-analysis in clinical trials. *Control Clin Trials,* 7:177–188.doi:10.1016/0197-2456(86)90046-2 PMID:3802833

46. Attia J, Thakkinstian A, D'Este C (2003). Meta-analyses of molecular association studies: methodologic lessons for genetic epidemiology. *J Clin Epidemiol,* 56:297–303.doi:10.1016/S0895-4356(03)00011-8 PMID:12767405

47. Ioannidis JP, Patsopoulos NA, Evangelou E (2007). Heterogeneity in meta-analyses of genome-wide association investigations. *PLoS One*, 2:e841.doi:10.1371/journal.pone.0000841 PMID:17786212

48. Zeggini E, Ioannidis JP (2009). Meta-analysis in genome-wide association studies. *Pharmacogenomics*, 10:191–201.doi:10.2217/14622416.10.2.191 PMID:19207020

49. Pereira TV, Patsopoulos NA, Salanti G, Ioannidis JP (2009). Discovery properties of genome-wide association signals from cumulatively combined data sets. *Am J Epidemiol*,170:1197–1206.doi:10.1093/aje/kwp262 PMID:19808636

50. Spiegelhalter DJ, Abrams KR, Myles PJ, editors. Evidence synthesis. Chichester (UK): John Wiley & Sons Ltd; 2004.

51. Spiegelhalter DJ, Thomas A, Best NG, Lunn D. WinBUGS version 1.4 users manual; 2003.

52. Davis RL, Khoury MJ (2007). The emergence of biobanks: practical design considerations for large population-based studies of gene-environment interactions. *Community Genet,* 10:181–185.doi:10.1159/000101760 PMID:17575463

53. Pocock SJ, Collier TJ, Dandreo KJ *et al.* (2004). Issues in the reporting of epidemiological studies: a survey of recent practice. *BMJ*, 329:883.doi:10.1136/bmj.38250.571088.55 PMID:15469946

54. von Elm E, Altman DG, Egger M *et al.*; STROBE Initiative (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet*, 370:1453–1457.doi:10.1016/S0140-6736(07) 61602-X PMID:18064739

55. STROBE Group. Strengthening the reporting of observational studies in epidemiology; 2005. Available from URL: http://www.strobe-statement.org/.

56. von Elm E, Egger M (2004). The scandal of poor epidemiological research. *BMJ*, 329:868–869.doi:10.1136/bmj.329.7471.868 PMID:15485939

57. Hill HA, Kleinbaum DG. Bias in observational studies. In: Gail MH, Benichou J, editors. Encyclopedia of epidemiologic methods. Chichester (UK): John Wiley & Sons, Ltd; 2000. p. 94–100.

58. Davey Smith G, Ebrahim S (2003). 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol,* 32:1–22.doi:10.1093/ije/dyg070 PMID:12689998

59. Little J, Khoury MJ (2003). Mendelian randomisation: a new spin or real progress? *Lancet*, 362:930–931.doi:10.1016/S0140-6736 (03)14396-6 PMID:14511923

60. Botto LD, Yang Q (2000). 5,10-Methylenetetrahydrofolate reductase gene variants and congenital anomalies: a HuGE review. *Am J Epidemiol,* 151:862–877. PMID:10791559

61. Brockton N, Little J, Sharp L, Cotton SC (2000). N-acetyltransferase polymorphisms and colorectal cancer: a HuGE review. *Am J Epidemiol,* 151:846–861. PMID:10791558

62. Cotton SC, Sharp L, Little J, Brockton N (2000). Glutathione S-transferase polymorphisms and colorectal cancer: a HuGE review. *Am J Epidemiol,* 151:7–32. PMID:10625170

63. Dorman JS, Bunker CH (2000). HLA-DQ and Type 1 Diabetes. *Epidemiol Rev,* 22:218–227. PMID:11218373

64. Madigan MP, Troisi R, Potischman N *et al.* (2000). Characteristics of respondents and non-respondents from a case-control study of breast cancer in younger women. *Int J Epidemiol,* 29:793–798.doi:10.1093/ije/29.5.793 PMID:11034958

65. Morton LM, Cahill J, Hartge P (2006). Reporting participation in epidemiologic studies: a survey of practice. *Am J Epidemiol,* 163:197–203.doi:10.1093/aje/kwj036 PMID:16339049

66. García-Closas M, Wacholder S, Caporaso NE, Rothman N. Inference issues in cohort and case-control studies of genetic effects and gene-environment interactions. In: Khoury MJ, Little J, Burke W, editors. Human genome epidemiology: a scientific foundation for using genetic information to improve health and prevent disease. New York (NY): Oxford University Press; 2004. p. 127–143.

67. Rothman N, Stewart WF, Caporaso NE, Hayes RB (1993). Misclassification of genetic susceptibility biomarkers: implications for case-control studies and cross-population comparisons. *Cancer Epidemiol Biomarkers Prev,* 2:299–303. PMID:8348052

68. Little J, Bradley L, Bray MS *et al.* (2002). Reporting, appraising, and integrating data on genotype prevalence and gene-disease associations. *Am J Epidemiol,* 156:300–310. PMID:12181099

69. Steinberg K, Gallagher M. Assessing genotypes in human genome epidemiology studies. In: Khoury MJ, Little J, Burke W, editors. Human genome epidemiology: a scientific foundation for using genetic information to improve health and prevent disease. New York (NY): Oxford University Press; 2004. P. 79–91.

70. Cardon LR, Palmer LJ (2003). Population stratification and spurious allelic association. *Lancet*, 361:598–604.doi:10.1016/S0140-6736 (03)12520-2 PMID:12598158

71. Thomas DC, Witte JS (2002). Point: population stratification: a problem for case-control studies of candidate-gene associations? *Cancer Epidemiol Biomarkers Prev,* 11:505–512. PMID:12050090

72. Wacholder S, Rothman N, Caporaso NE (2000). Population stratification in epidemiologic studies of common genetic variants and cancer: quantification of bias. *J Natl Cancer Inst,* 92:1151–1158.doi:10.1093/jnci/92.14.1151 PMID:10904088

73. Freedman ML, Reich D, Penney KL *et al.* (2004). Assessing the impact of population stratification on genetic association studies. *Nat Genet,* 36:388–393.doi:10.1038/ng1333 PMID:15052270

74. Ioannidis JP, Ntzani EE, Trikalinos TA (2004). 'Racial' differences in genetic effects for complex diseases. *Nat Genet,* 36:1312–1318.doi:10.1038/ng1474 PMID:15543147

75. Khlat M, Cazes MH, Génin E, Guiguet M (2004). Robustness of case-control studies of genetic factors to population stratification: magnitude of bias and type I error. *Cancer Epidemiol Biomarkers Prev,* 13:1660–1664. PMID:15466984

76. Trikalinos TA, Ntzani EE, Contopoulos-Ioannidis DG, Ioannidis JPA (2004). Establishment of genetic associations for complex diseases is independent of early study findings. *Eur J Hum Genet,* 12:762–769. doi:10.1038/sj.ejhg.5201227 PMID:15213707

77. Egger M, Davey Smith G, Schneider M, Minder C (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315: 629–634. PMID:9310563

78. Lau J, Ioannidis JP, Terrin N *et al.* (2006). The case of the misleading funnel plot. *BMJ*, 333:597–600.doi:10.1136/bmj.333.7568.597 PMID:16974018

79. Terrin N, Schmid CH, Lau J (2005). In an empirical evaluation of the funnel plot, researchers could not visually identify publication bias. *J Clin Epidemiol,* 58:894–901.doi:10.1016/j.jclinepi.2005.01.006 PMID:16085192

80. Ioannidis JP, Trikalinos TA (2007). The appropriateness of asymmetry tests for publication bias in meta-analyses: a large survey. *Can Med Assoc J,* 176:1091–1096. doi:10.1503/cmaj.060410 PMID:17420491

81. Harbord RM, Egger M, Sterne JA (2006). A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. *Stat Med*, 25:3443–3457.doi: 10.1002/sim.2380 PMID:16345038

82. Chan AW, Hróbjartsson A, Haahr MT *et al.* (2004). Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA*, 291:2457–2465.doi:10.1001/jama.291.20.2457 PMID:15161896

83. Chan AW, Krleza-Jerić K, Schmid I, Altman DG (2004). Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. *Can Med Assoc J*, 171:735–740.doi:10.1503/cmaj.1041086 PMID:15451835

84. Chan AW, Altman DG (2005). Identifying outcome reporting bias in randomised trials on PubMed: review of publications and survey of authors. *BMJ*, 330:753.doi:10.1136/bmj.38356.424606.8F PMID:15681569

85. Contopoulos-Ioannidis DG, Alexiou GA, Gouvias TC, Ioannidis JP (2006). An empirical evaluation of multifarious outcomes in pharmacogenetics: beta-2 adrenoceptor gene polymorphisms in asthma treatment. *Pharmacogenet Genomics*, 16:705–711. doi:10.1097/01.fpc.0000236332.11304.8f PMID:17001289

86. Ioannidis JP, Trikalinos TA (2007). An exploratory test for an excess of significant findings. *Clin Trials*, 4:245–253.doi:10.1177/1740774507079441 PMID:17715249

87. Ardlie KG, Kruglyak L, Seielstad M (2002). Patterns of linkage disequilibrium in the human genome. *Nat Rev Genet,* 3:299–309. doi:10.1038/nrg777 PMID:11967554

88. Salanti G, Amountza G, Ntzani EE, Ioannidis JP (2005). Hardy-Weinberg equilibrium in genetic association studies: an empirical evaluation of reporting, deviations, and power. *Eur J Hum Genet,* 13:840–848. doi:10.1038/sj.ejhg.5201410 PMID:15827565

89. Hill AB (1965). The environment and disease: association or causation? *Proc R Soc Med,* 58:295–300. PMID:14283879

90. Surgeon General (Advisory Committee). Smoking and health. U.S. Department of Health, Education and Welfare. Washington (DC): U.S. Government Printing Office; 1964.

91. Ioannidis JP, Trikalinos TA, Khoury MJ (2006). Implications of small effect sizes of individual genetic variants on the design and interpretation of genetic association studies of complex diseases. *Am J Epidemiol,* 164:609–614.doi:10.1093/aje/kwj259 PMID:16893921

92. Rothman KJ, Greenland S, Lash TL, editors. Modern epidemiology. 3rd ed. Philadelphia (PA): Lippincott Williams & Wilkins; 2008.

93. Cloninger CR (2002). The discovery of susceptibility genes for mental disorders. *Proc Natl Acad Sci USA*, 99:13365–13367. doi:10.1073/pnas.222532599 PMID:12374853

94. Glazier AM, Nadeau JH, Aitman TJ (2002). Finding genes that underlie complex traits. *Science*, 298:2345–2349.doi:10.1126/science.1076641 PMID:12493905

95. Harrison PJ, Owen MJ (2003). Genes for schizophrenia? Recent findings and their pathophysiological implications. *Lancet*, 361:417–419.doi:10.1016/S0140-6736(03)12379-3 PMID:12573388

96. Editorial (2001). Challenges for the 21st century. *Nat Genet,* 29:353–354.doi:10.1038/ng1201-353 PMID:11726913

97. Page GP, George V, Go RC *et al.* (2003). "Are we there yet?": Deciding when one has demonstrated specific genetic causation in complex diseases and quantitative traits. *Am J Hum Genet,* 73:711–719.doi:10.1086/378900 PMID:13680525

98. Rebbeck TR, Martínez ME, Sellers TA *et al.* (2004). Genetic variation and cancer: improving the environment for publication of association studies. *Cancer Epidemiol Biomarkers Prev,* 13:1985–1986. PMID:15598750

99. Weiss KM, Terwilliger JD (2000). How many diseases does it take to map a gene with SNPs? *Nat Genet,* 26:151–157. doi:10.1038/79866 PMID:11017069

100. Ioannidis JP, Kavvoura FK (2006). Concordance of functional in vitro data and epidemiological associations in complex disease genetics. *Genet Med,* 8:583–593. doi:10.1097/01.gim.0000237775.93658.0c PMID:16980815

101. Begg CB (2005). Reflections on publication criteria for genetic association studies. *Cancer Epidemiol Biomarkers Prev,* 14:1364–1365.doi:10.1158/1055-9965.EPI-05-0407 PMID:15941941

102. Pharoah PD, Dunning AM, Ponder BA, Easton DF (2005). The reliable identification of disease-gene associations. *Cancer Epidemiol Biomarkers Prev,* 14:1362.doi:10.1158/1055-9965.EPI-05-0405 PMID:15941939

103. Ioannidis JP (2006). Common genetic variants for breast cancer: 32 largely refuted candidates and larger prospects. *J Natl Cancer Inst,* 98:1350–1353.doi:10.1093/jnci/djj392 PMID:17018776

104. Ioannidis JP, Boffetta P, Little J *et al.* (2008). Assessment of cumulative evidence on genetic associations: interim guidelines. *Int J Epidemiol,* 37:120–132.doi:10.1093/ije/dym159 PMID:17898028

105. Ioannidis JP, Rosenberg PS, Goedert JJ, O'Brien TR; International Meta-analysis of HIV Host Genetics (2002). Commentary: meta-analysis of individual participants' data in genetic epidemiology. *Am J Epidemiol,* 156:204–210.doi:10.1093/aje/kwf031 PMID:12142254

106. Steinberg KK, Smith SJ, Stroup DF *et al.* (1997). Comparison of effect estimates from a meta-analysis of summary data from published studies and from a meta-analysis using individual patient data for ovarian cancer studies. *Am J Epidemiol,* 145:917–925. PMID:9149663

107. Seminara D, Obrams GI (1994). Genetic epidemiology of cancer: a multidisciplinary approach. *Genet Epidemiol,* 11:235–254. doi:10.1002/gepi.1370110303 PMID:8088505

108. Kreeger K (2003). Consortia, 'big science' part of a paradigm shift for genetic epidemiology. *J Natl Cancer Inst*, 95:640–641. doi:10.1093/jnci/95.9.640 PMID:12734309

109. Ioannidis JP, Gwinn M, Little J *et al.*; Human Genome Epidemiology Network and the Network of Investigator Networks (2006). A road map for efficient and reliable human genome epidemiology. *Nat Genet,* 38:3–5. doi:10.1038/ng0106-3 PMID:16468121

110. Seminara D, Khoury MJ, O'Brien TR *et al.*; Human Genome Epidemiology Network; Network of Investigator Networks (2007). The emergence of networks in human genome epidemiology: challenges and opportunities. *Epidemiology*, 18:1–8.doi:10.1097/01.ede.0000249540.17855.b7 PMID:17179752

111. Colhoun HM, McKeigue PM, Davey Smith G (2003). Problems of reporting genetic associations with complex outcomes. *Lancet*, 361:865–872.doi:10.1016/S0140-6736(03)12715-8 PMID:12642066

112. Ioannidis JP (2005). Why most published research findings are false. *PLoS Med,* 2:e124.doi:10.1371/journal.pmed.0020124 PMID:16060722

113. Elbaz A, Nelson LM, Payami H *et al.* (2006). Lack of replication of thirteen single-nucleotide polymorphisms implicated in Parkinson's disease: a large-scale international study. *Lancet Neurol,* 5:917–923.doi:10.1016/S1474-4422(06)70579-8 PMID:17052658

114. John EM, Hopper JL, Beck JC *et al.*; Breast Cancer Family Registry (2004). The Breast Cancer Family Registry: an infrastructure for cooperative multinational, interdisciplinary and translational studies of the genetic epidemiology of breast cancer. *Breast Cancer Res,* 6:R375–R389.doi:10.1186/bcr801 PMID:15217505

115. Andrulis IL, Anton-Culver H, Beck J *et al.*; Cooperative Family Registry for Breast Cancer studies (2002). Comparison of DNA- and RNA-based methods for detection of truncating BRCA1 mutations. *Hum Mutat,* 20:65–73.doi:10.1002/humu.10097 PMID:12112659

116. Yonan AL, Palmer AA, Gilliam TC (2006). Hardy-Weinberg disequilibrium identified genotyping error of the serotonin transporter (SLC6A4) promoter polymorphism. *Psychiatr Genet,* 16:31–34.doi:10.1097/01.ypg.0000174393.79883.05 PMID:16395127

117. Hunter DJ, Kraft P (2007). Drinking from the fire hose–statistical issues in genomewide association studies. *N Engl J Med,* 357:436–439.doi:10.1056/NEJMp078120 PMID:17634446

118. Hindorff LA, Junkins HA, Mehta JP, Manolio TA. A catalog of published genome-wide association studies; 2009. Available from URL: http//genome.gov/gwastudies.

119. Voight BF, Scott LJ, Steinthorsdottir V *et al.*; MAGIC investigators; GIANT Consortium (2010). Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet,* 42:579–589.doi:10.1038/ng.609 PMID:20581827

120. Elliott KS, Zeggini E, McCarthy MI *et al.*; Australian Melanoma Family Study Investigators; PanScan Consortium (2010). Evaluation of association of HNF1B variants with diverse cancers: collaborative analysis of data from 19 genome-wide association studies. *PLoS One,* 5:e10858.doi:10.1371/journal.pone.0010858 PMID:20526366

121. Zeggini E, Weedon MN, Lindgren CM *et al.*; Wellcome Trust Case Control Consortium (WTCCC) (2007). Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science*, 316:1336–1341.doi:10.1126/science.1142364 PMID:17463249

122. Evangelou E, Maraganore DM, Ioannidis JP (2007). Meta-analysis in genome-wide association datasets: strategies and application in Parkinson disease. *PLoS One,* 2:e196.doi:10.1371/journal.pone.0000196 PMID:17332845

123. Rivadeneira F, Styrkársdottir U, Estrada K *et al.*; Genetic Factors for Osteoporosis (GEFOS) Consortium (2009). Twenty bone-mineral-density loci identified by large-scale meta-analysis of genome-wide association studies. *Nat Genet,* 41:1199–1206.doi:10.1038/ng.446 PMID:19801982

124. Zeggini E, Scott LJ, Saxena R *et al.*; Wellcome Trust Case Control Consortium (2008). Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet,* 40:638–645.doi:10.1038/ng.120 PMID:18372903

125. McGovern DP, Gardet A, Törkvist L *et al.*; NIDDK IBD Genetics Consortium (2010). Genome-wide association identifies multiple ulcerative colitis susceptibility loci. *Nat Genet,* 42:332–337.doi:10.1038/ng.549 PMID:20228799

126. Houlston RS, Webb E, Broderick P *et al.*; Colorectal Cancer Association Study Consortium; CoRGI Consortium; International Colorectal Cancer Genetic Association Consortium (2008). Meta-analysis of genome-wide association data identifies four new susceptibility loci for colorectal cancer. *Nat Genet,* 40:1426–1435.doi:10.1038/ng.262 PMID:19011631

127. McCarthy MI, Abecasis GR, Cardon LR *et al.* (2008). Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet,* 9:356–369.doi:10.1038/nrg2344 PMID:18398418

128. Risch N (1990). Linkage strategies for genetically complex traits. III. The effect of marker polymorphism on analysis of affected relative pairs. *Am J Hum Genet,* 46:242–253. PMID:2301394

129. Becker KG, Simon RM, Bailey-Wilson JE *et al.* (1998). Clustering of non-major histocompatibility complex susceptibility candidate loci in human autoimmune diseases. *Proc Natl Acad Sci USA,* 95:9979–9984.doi:10.1073/pnas.95.17.9979 PMID:9707586

130. Cornélis F, Fauré S, Martinez M *et al.*; ECRAF (1998). New susceptibility locus for rheumatoid arthritis suggested by a genome-wide linkage study. *Proc Natl Acad Sci USA,* 95:10746–10750.doi:10.1073/pnas.95.18.10746 PMID:9724775

131. Wise LH, Lanchbury JS, Lewis CM (1999). Meta-analysis of genome searches. *Ann Hum Genet,* 63:263–272.doi:10.1046/j.1469-1809.1999.6330263.x PMID:10738538

132. Dempfle A, Loesgen S (2004). Meta-analysis of linkage studies for complex diseases: an overview of methods and a simulation study. *Ann Hum Genet,* 68:69–83.doi:10.1046/j.1529-8817.2003.00061.x PMID:14748832

133. Zintzaras E, Ioannidis JP (2005). HEGESMA: genome search meta-analysis and heterogeneity testing. *Bioinformatics*, 21:3672–3673.doi:10.1093/bioinformatics/bti536 PMID:15955784

134. Terwilliger JD, Ott J. Handbook of human genetic linkage. Baltimore (MD): Johns Hopkins University Press; 1994.

135. Conlon EM, Song JJ, Liu A (2007). Bayesian meta-analysis models for microarray data: a comparative study. *BMC Bioinformatics,* 8:80.doi:10.1186/1471-2105-8-80 PMID:17343745

136. Rhodes DR, Barrette TR, Rubin MA *et al.* (2002). Meta-analysis of microarrays: interstudy validation of gene expression profiles reveals pathway dysregulation in prostate cancer. *Cancer Res*, 62:4427–4433. PMID:12154050

137. Warnat P, Eils R, Brors B (2005). Cross-platform analysis of cancer microarray data improves gene expression based classification of phenotypes. *BMC Bioinformatics,* 6:265.doi:10.1186/1471-2105-6-265 PMID:16271137

138. Smid M, Dorssers LC, Jenster G (2003). Venn Mapping: clustering of heterologous microarray data based on the number of co-occurring differentially expressed genes. *Bioinformatics*, 19:2065–2071.doi:10.1093/bioinformatics/btg282 PMID:14594711

139. Grützmann R, Boriss H, Ammerpohl O *et al.* (2005). Meta-analysis of microarray data on pancreatic cancer defines a set of commonly dysregulated genes. *Oncogene*, 24:5079–5088.doi:10.1038/sj.onc.1208696 PMID:15897887

140. Zintzaras E, Ioannidis JP (2008). Meta-analysis for ranked discovery datasets: theoretical framework and empirical demonstration for microarrays. *Comput Biol Chem,* 32:38–46. PMID:17988949

141. Hsu YH, Zillikens MC, Wilson SG *et al.* (2010). An integration of genome-wide association study and gene expression profiling to prioritize the discovery of novel susceptibility Loci for osteoporosis-related traits. *PLoS Genet,* 6:e1000977.doi:10.1371/journal.pgen.1000977 PMID:20548944

142. Zhong H, Yang X, Kaplan LM *et al.* (2010). Integrating pathway analysis and genetics of gene expression for genome-wide association studies. *Am J Hum Genet,* 86:581–591.doi:10.1016/j.ajhg.2010.02.020 PMID:20346437

143. Friedenreich CM (1994). Methodologic issues for pooling dietary data. *Am J Clin Nutr,* 59 Suppl;251S–252S. PMID:8279435

144. Taioli E, Bonassi S (2002). Methodological issues in pooled analysis of biomarker studies. *Mutat Res,* 512:85–92.doi:10.1016/S1383-5742(02)00027-3 PMID:12220591

145. Veglia F, Matullo G, Vineis P (2003). Bulky DNA adducts and risk of cancer: a meta-analysis. *Cancer Epidemiol Biomarkers Prev,* 12:157–160. PMID:12582026

146. Flores-Mateo G, Navas-Acien A, Pastor-Barriuso R, Guallar E (2006). Selenium and coronary heart disease: a meta-analysis. *Am J Clin Nutr,* 84:762–773. PMID:17023702

147. Kyzas PA, Denaxa-Kyza D, Ioannidis JP (2007). Almost all articles on cancer prognostic markers report statistically significant results. *Eur J Cancer,* 43:2559–2579. PMID:17981458

148. Kyzas PA, Loizou KT, Ioannidis JP (2005). Selective reporting biases in cancer prognostic factor studies. *J Natl Cancer Inst,* 97:1043–1055.doi:10.1093/jnci/dji184 PMID:16030302

149. McShane LM, Altman DG, Sauerbrei W (2005). Identification of clinically useful cancer prognostic factors: what are we missing? *J Natl Cancer Inst,* 97:1023–1025.doi:10.1093/jnci/dji193 PMID:16030294