The past decade has witnessed growing interest in genetic predisposition to common diseases, and along with rapid advancements in high-throughput genotyping technology, has resulted in a tremendous amount of published epidemiological evidence on gene-disease associations. Reported genetic associations with common diseases have become numerous and are mostly of small magnitude [1]. With this growth in evidence has come an increasing need to collate and summarize the evidence in order to identify true genetic associations among the large volume of false positives [2]. Convincing evidence of true association therefore requires careful control over potential biases and chance effects. Control over bias is important both in study design [3] and in considering the selective availability of data on associations that have been examined [4]. Because most genetic associations are small, large sample sizes are necessary for their detection, especially when many genetic variants are investigated simultaneously, as in genome-wide association studies. Furthermore, replication of findings in independent data sets is now widely regarded as a prerequisite for convincing evidence of association. Thus, multiple studies from several independent groups exist. Systematic reviews and meta-analyses have become a common approach to summarize gene-disease associations [5]. However, as apparently inconsistent results and interpretations of syntheses of evidence on the same genetic association have been reported [6], it is useful to lay out issues in the methods and conduct of these types of research studies. The Human Genome Epidemiology Network (HuGENet) has developed “HuGE reviews” (typically systematic reviews) as the cornerstone of an online resource containing the cumulative and changing information on epidemiologic aspects of human genes [7]. Many of these contain meta-analyses, or statistical syntheses of the findings from multiple studies. In this article, we describe some key components of the methodology for undertaking systematic reviews and meta-analyses of genetic association studies. Detailed discussions of these issues and further recommendations can be found in the HuGE Review Handbook [8] (see also Box 1). Additional detailed guidance on conducting systematic reviews is available in the Cochrane Handbook for Systematic Reviews of Interventions [9] and The Handbook of Research Synthesis [10].

Why Systematic Reviews?

Traditional (narrative) reviews are subjective, and as such have a number of disadvantages that lead them to being more prone to bias and error [11]. Narrative reviews rarely state how studies were selected for inclusion, how (or whether)
their quality was assessed, or how the findings from multiple studies were synthesized in order to draw conclusions. This can lead to a review that supports and reinforces the author’s view, which can be misleading. In contrast, systematic reviews are designed as rigorous research studies. They allow a more objective appraisal of the evidence than a narrative review by aiming to identify, critique, and synthesize evidence from all relevant existing studies on the topic in question using predefined methods. Figure 1 outlines the general processes involved in a systematic review.

Data for systematic reviews are typically collated primarily from publications, often supplemented by data from correspondence with authors of the original studies. In the field of human genome epidemiology, however, the ability to look at increasing numbers of markers prevents the publication of summary data for each genetic variant in a traditional journal format. A variety of sources may therefore be used, including the published literature, online supplementary materials, online databases, and study authors. An increasingly popular technique is the meta-analysis of individual participant data through the collaboration of a consortium of investigators. This is likely to minimize bias compared with a literature-based systematic review, but possibly at the expense of precision, since a comprehensive systematic review may include evidence from studies that are not involved in the consortium. The two approaches are complementary to one another, and some of their potential advantages and disadvantages are summarized in Table 1.

**Preparation of a Protocol**

Like other research studies, systematic reviews should ideally be carefully planned with a detailed protocol prepared in advance. Producing a protocol with criteria predefined for study selection will minimize any selection bias based on study results. Such a protocol should therefore clearly formulate the review question, explain the rationale for conducting such a review, define a priori eligibility criteria for study inclusion, describe the methods for conducting a comprehensive search for studies, indicate the methods for assessing study validity and relevance, and state whether a meta-analysis is planned and describe the methodology to be used for conducting such an analysis.

**Objectives and Eligibility Criteria**

The objectives of a systematic review of genetic association studies will typically be (i) to identify all epidemiological
investigations of the associations of interest; (ii) to assess the validity of the evidence; (iii) to determine whether an association exists; (iv) to assess whether associations are consistent across studies in magnitude and direction; and (v) to quantify the likely magnitude of an association if it exists. Associations of interest have in the past usually been identified through biological plausibility, although candidate associations are now frequently identified from genome-wide association studies.

The applicability of a systematic review depends on the selection of appropriate studies to be included in it. Eligibility criteria for the review, defined a priori, should allow relevant studies investigating the gene(s), allelic variant(s), disease(s), clinical subtypes, and any other outcomes of interest (such as molecular biomarkers) to be easily identifiable and included in the review. Appropriate study designs and study populations should also be considered in order to allow the reliable assessment of associations and interactions.

### Searching for Eligible Studies

A key characteristic of a systematic review is a comprehensive search. Limiting a review to studies identified only from MEDLINE (or PubMed) is usually insufficient, for two reasons. First, any review restricted to published literature is prone to publication bias, whereby only a subset comprising the most “publishable” findings are available from the totality of evidence on a particular genetic association. Thus, sources beyond the published literature should be examined. Second, even within the published literature, reporting bias may be present whereby studies with different conclusions are published in different types of journals. MEDLINE is just one of several major sources of bibliographic information. These biases may, to some extent, be reduced by searching comprehensively for studies available through other sources. Other bibliographic databases likely to be useful for genetic association studies include EMBASE and the Science Citation Index. Overlap between these various databases is far from complete; for example, of approximately 4,800 journals indexed in EMBASE, 1,700 are not indexed in MEDLINE (see Chapter 6 in [9]). The implications of this incomplete overlap for genetic association studies are not clear. Furthermore, bibliographic searches may still not retrieve all articles that are in the indexed journals [12].

Further sources that might be searched include reference lists of study reports and review articles, Web sites, theses, books, reports, and online databases (including data repositories for genome-wide and other association studies [13]). Sources other than peer-reviewed journals are often referred to as “grey literature.”

Once the searches have been undertaken, the records identified need to be assessed against the eligibility criteria for the review. A typical example of the process from search to selection is illustrated in Figure 2.

### Minimizing Human Error

All reasonable attempts should be made to prevent the introduction of errors and personal biases. An important attribute of systematic reviews is that criteria for study inclusion are clearly prespecified. This is perhaps the most notable difference between a narrative review and a systematic review. Independent duplication of steps in the review process, such as selection of studies, extraction of data, and critical assessment of methods used in the individual studies, can further reduce biases and minimize errors. Accidental omission of data, or accidental duplication of a study in a meta-analysis, may lead to spurious false-negative or false-positive findings.

### Assessing Validity of Each Study

The validity of a meta-analysis depends on the validity of the studies included within it, so it is important that each component study is appraised before being included. Because effect magnitudes in genetic association studies are generally small (i.e., odds ratio < 1.2), even small biases may be important. The most important sources of bias in genetic association studies are less well understood than those in other study designs such as randomized trials. Extensive discussions of potential biases are available, with the principal candidates being case definition, population stratification (confounding due to sub-populations in the sample that differ both in genotype prevalence and disease risk), and methods in the collection, handling, and processing of DNA and the determination of genotypes (including blinding to case-control status) [3,14]. Little empirical evidence associating study results with study characteristics for genetic association studies exists as yet, however. This evidence will perhaps be most reliably derived from meta-epidemiological studies, in which results of studies with different characteristics are compared within meta-analyses, and findings are synthesized across meta-analyses to enhance power [15,16].

The appraisal of potential biases is difficult in practice, not only because of the uncertainty over which study characteristics are important, but also because of incomplete or variable reporting of the methods used in the studies themselves [17]. Initiatives such as the STREGA (STrengthening the REporting of Genetic Association Studies) statement, which offers guidelines for reporting

| Table 1: Data Sources for Meta-Analysis of Genetic Association Studies |
|---------------------------|----------------|----------------|
| **Approach**              | **Advantages** | **Disadvantages** |
| Literature-based systematic review | Low resource requirements. | Can be prone to serious bias due to selective availability of published data. |
| Literature-based systematic review and liaison with investigators | Relatively low resource requirements. | May be prone to bias due to selective availability of published data. |
| Consortium-based meta-analysis of existing data | Data not subject to selective availability; potential to harmonize data. | Data may be incomplete; inclusion in consortium may be related to findings. |
| Consortium-based meta-analysis with prospective genotyping | Complete data; availability of data unrelated to findings; ability to harmonize methods. | Genotyping costs. |
of individual genetic association studies, may improve the situation in the future.

Assessing Bias in the Review as a Whole

In addition to potential biases in the individual studies, attention should be paid to bias in the collection of studies as a whole. Two particular considerations are reporting biases and bias in the selection of genetic variants to study.

Reporting biases include both publication bias, which typically refers to suppression of evidence concerning an entire study on the basis of its findings, and selective reporting, whereby only the most exciting findings are reported. Selective reporting of only the most promising variants is a natural consequence of an attempt to summarize an association study of many genetic markers within the constraints of a traditional paper journal article. To overcome these constraints, association studies are increasingly exploiting Web-based publication for complete findings, although in practice there are often obstacles to prevent this occurring.

The possibility of reporting biases can only be discarded if it is known that all eligible data from all eligible studies are available (or if a truly unbiased subset of these data is available). It is difficult to imagine situations in which reviewers are confident that all existing studies are either known or reported without bias. The most reliable way to overcome these biases is therefore in the prospective generation of data, for example by using a consortium-based approach.

It is also helpful to keep in mind that the very act of investigating a gene-disease association may be dependent on earlier results. Even in a world of flawless studies that are completely reported, it is likely that the most exciting findings are those that are targets for replication. Initial exciting findings may well be chance effects, and so situations in which a meta-analysis is contemplated (those in which multiple studies exist) comprise a biased subset of possibly important associations. There is good evidence of the exaggerated effect that is often seen in the earliest report of an association compared with subsequent attempts to replicate the finding [18].

Meta-Analysis

Meta-analysis is the statistical synthesis of results from multiple studies [19]: an example is provided in Figure 3 [20]. When implemented and interpreted appropriately, and applied to unbiased and correctly analyzed studies, it provides a powerful tool to understand both similarities and differences in results from multiple studies. By exploiting the totality of evidence, meta-analyses offer enhanced power to detect associations and increased precision in the estimation of their magnitude. Meta-analyses are encouraged

---

do:10.1371/journal.pmed.1000028.g002

Figure 2: A Typical Process for Searching and Assessing Eligibility of Studies
in systematic reviews. However, attention should always be paid to the possibility of reporting biases in reviews based on published literature. Furthermore, it is important that each entry in a meta-analysis represents an independent sample of data. Thus for example, multiple reports of the same study need to be merged to obtain a single “best” answer for that study prior to inclusion in the meta-analysis.

Most meta-analyses are undertaken by calculating weighted averages of the estimates from multiple studies. A metric is chosen for the analysis that ensures comparable quantities with reasonable statistical properties. For example, log odds ratios are typically combined for case-control studies (and then presented as odds ratios). If quantitative traits are measured using different methods across studies, they may be standardized before pooling, usually by expressing the results in terms of standard deviations.

Consistency of results across studies should always be evaluated, and can be achieved using a statistical test of homogeneity or by quantifying the between-study variance (or quantities derived from these such as \( I^2 \), which describes the proportion variation in point estimates that is due to heterogeneity rather than within-study error [21]). Potential reasons for variation in findings across studies should also be investigated. To minimize unreliable consequences from multiple testing, potential sources of variation in study results should be limited in number and prespecified whenever possible. Consideration should be given to using methods to temper the statistical significance of findings based on a small number of studies [22].

Sensitivity analyses are to be encouraged, to ensure that findings are robust. For example, undue influence of the initial study, of smaller studies, or of studies with potential biases may be responsible for inappropriate conclusions. Furthermore, the choice of statistical model (or method) or the inclusion of studies with uncertain eligibility for the systematic review may affect the results, and robustness to these should also be evaluated.

Meta-analysis methods of gene-disease association studies closely follow well-developed methodology for randomized trials [4,6,23]. Due to so-called “Mendelian randomization” in the transmission of genetic material from parents to children, confounding is generally thought to be of minimal concern (although not ignorable, since it may arise through population stratification). Thus, unadjusted analyses are common in meta-analyses, even if matched studies have adjusted for matching factors such as age and sex.

Figure 3: Example of a Forest Plot
The plot shows the relation between N-acetyltransferase (NAT1) acetylation status (slow or rapid) and bladder cancer risk (data from Sanderson et al. [20]). The classification of acetylation phenotype is made either directly, by measuring the metabolism of a probe drug, or indirectly, by measuring the genotype associated with that phenotype.
The era of genome-wide association studies has implications for the meta-analysis of genetic associations. Although, in principle, summary data for the variant(s) of interest in the review should be sought from available genome-wide association studies in the same way as they are sought from candidate gene studies, in practice many genome-wide association studies are designed primarily for detecting associations rather than quantifying them, and the potential for bias must be assessed. Furthermore, the variant(s) of interest might not have been included in a genome-wide scan. While it is possible that some unknown genotypes could be imputed, proposed methods typically require detailed data [24], and it is unclear whether a more naive approach, based on minimal summary data and linkage disequilibrium information, would be appropriate.

Other special considerations in the meta-analysis of gene-disease associations are discussed in detail elsewhere and include the choice of inheritance model [25,26], the treatment of Hardy-Weinberg equilibrium [27,28], and the combination of associations for markers known to be in linkage disequilibrium [29]. Developments in the last area are now allowing complex syntheses of data across whole-genome-wide association studies [30].

**Presenting and Interpreting Results**

A systematic summary of all available evidence should allow the strengths and gaps in the evidence base to be identified and allow recommendations to be made in order to stimulate research to address such gaps. The magnitude of the association between the allelic variants and the clinical outcomes studied in terms of relative and attributable risks in different populations should be summarized in a systematic and concise way. Comments should also be provided on the quality and methodology of studies. Tables should summarize information on each gene-disease association study (possibly online as Web-based supplements, depending on size and journal formats). Table 2 lists some of the characteristics most likely to be relevant. If a meta-analysis is conducted, these should usually be presented as forest plots as in Figure 3. Statistics related to heterogeneity (e.g., between-study variance or \( I^2 \) and confidence intervals) and investigations of bias should be presented. Multiple subgroup and sensitivity analyses are convenently presented as “overview” forest plots without including the individual studies.

A discussion of the results should also address issues such as the overall quantity and quality of the evidence base, the consideration of possible publication and selective reporting biases, the likelihood that associations are causal, and any potential public health applications. For example, results from small or few studies should be interpreted with great caution, as it is common to see dissipation of early claimed effects [31], and early studies have no predictive power for the subsequent picture of the evidence [32]. Interim guidance is available to assess the strength of evidence that an observed association is genuine [33]. These criteria assess the three domains of precision (i.e., through sample size), consistency of results across studies (i.e., through meta-analytic measures of heterogeneity), and protection from bias both within and across studies in a meta-analysis.

Additional detailed guidance for the general reporting of systematic reviews and meta-analyses are available from PRISMA (Preferred Reporting Items for Systematic Reviews; formerly known as QUOROM) [34]. Reporting guidance for meta-analyses of observational studies are also available from MOOSE (Meta-analysis Of Observational Studies in Epidemiology) [35].

**Concluding Remarks**

The strengths and limitations of systematic reviews are well established for clinical trials, largely through the efforts of The Cochrane Collaboration [36]. They are increasingly being applied to observational studies, and currently there are as many meta-analyses of observational data conducted as there are of clinical trials. The citation impact of both types of meta-analyses is equally high, the highest among all study designs in the health sciences [37]. The principal value of a well-conducted systematic review of genetic association studies is in establishing reliably the presence and magnitude of individual gene-disease associations. By complementing both consortium-based pooled analyses and larger-scale attempts to collate genetic association evidence across whole fields, they play an important role alongside other research designs in the integration of evidence on genetic association [5]. At their most ambitious, systematic reviews and meta-analyses can collate evidence across all studied genetic variants for a phenotype, with notable examples provided by three databases of genetic association evidence for Alzheimer disease, Parkinson disease, and schizophrenia that include ad hoc meta-analyses [38–40].

**Table 2: Characteristics Likely To Be of Interest for Each Study Included in a Review**

| Source of Data | Findings | Phenotypes | Genotypes and Other Exposures | Population |
|----------------|----------|------------|-------------------------------|------------|
| ID (e.g., first author) | Genotype data for cases and controls, as presented in the study report | Case definition/diagnosis | Gene |
| Year of publication | Measures of association (e.g., odds ratios and 95% confidence intervals) | Clinical subtypes of interest | Polymorphism (including rs number) |
| Study design | Interactions reported | Ethnicity | Genotyping methods (including blinding of researchers) |
| Assessments of validity | State if data taken directly from publication, from online database, or from correspondence with study authors | Age | Other important participant characteristics |
| State if data taken directly from publication, from online database, or from correspondence with study authors | All references (clearly linking multiple publications from the same study) | Sex |  |

![Table 2: Characteristics Likely To Be of Interest for Each Study Included in a Review](doc10.1371/journal.pmed.1000028.t002)
Within the Summary Points we summarize our key suggestions for methodology of systematic reviews of genetic association studies. These reflect our accumulated experience to date and are not intended to lay down standards for universal application. This is an area of rapid development. In a decade’s time, appropriate methods for the collation and integration of evidence on gene-disease association may be substantially different.

References

1. Khoury MJ, Little J, Gwinn M, Ioannidis PA (2007) On the synthesis and interpretation of consistent but weak gene-disease associations in the era of genome-wide association studies. Int J Epidemiol 36: 439-445.
2. Ioannidis JP (2003) Genetic associations: False or true? Trends Mol Med 9: 135-138.
3. Little J, Bradley L, Bray MS, Clyne M, Dorman J, et al. (2002) Reporting, appraising, and integrating data on genotype prevalence and gene-disease associations. Am J Epidemiol 156: 300-310.
4. Salanti G, Sanderson S, Higgins JPT (2005) Obstacles and opportunities in meta-analysis of genetic association studies. Genet Med 7: 13-20.
5. Ioannidis JPA, Gwinn M, Little J, Higgins JPT, Bernstein JL, et al. (2006) A road map for efficient and reliable human genome epidemiology. Nat Genet 38: 5-5.
6. Kavvoura FK, Ioannidis JPA (2008) Methods for meta-analysis in genetic association studies: A review of their potential and pitfalls. Hum Genet 123: 1-14.
7. Khoury MJ, Dorman JS (1998) The Human Genome Epidemiology Network. Am J Epidemiol 148: 1-5.
8. Little J, Higgins JPT, editors (2007) The HuGENetTM HuGE review handbook. Version 1.0. Available: http://hugenet.ca/. Accessed 29 January 2009.
9. Higgins JPT, Green S, editors (2008) Cochrane handbook for systematic reviews of interventions. Chichester (UK): John Wiley & Sons.
10. Cooper H, Hedges LV, editors (1997) The handbook of research synthesis. New York: The Russell Sage Foundation. 573 p.
11. Khoury MJ, Little J, Gwinn M, Ioannidis PA (2007) On the synthesis and interpretation of consistent but weak gene-disease associations in the era of genome-wide association studies. Int J Epidemiol 36: 439-445.
12. Hopewell S, Clarke M, Lefebvre C, Scherer R (2007) Handsearching versus electronic searching to identify reports of randomized trials. Cochrane Database Syst Rev: MR000001.
13. Froehlich AJ, Higgins JPT (2007) Online genome databases informing human genome epidemiology: A review. BMC Med Res Methodol 7: 31.
14. Clayton DG, Walker NM, Smyth DJ, Pask R, Cooper JD, et al. (2002) Reporting, appraising, and integrating data on genotype prevalence and gene-disease associations. Am J Epidemiol 156: 300-310.
15. Ioannidis JPA, Trikalinos TA, Ntzani EE, Contopoulos-Ioannidis DG (2003) Genetic associations in large versus small studies: An empirical assessment. Lancet 361: 367-371.
16. Sterne JAC, Juni P, Schulz KF, Altman DG, Bartlett C, et al. (2002) Statistical methods for assessing the influence of study characteristics on treatment effects in ‘meta-epidemiological’ research. Stat Med 21: 1513-1524.
17. Yusupriya A, Evangelou E, Kavvoura FK, Patsopoulos NA, Clyne M, et al. (2008) Reporting of human genome epidemiology (HuGE) association studies: An empirical assessment. BMC Med Res Methodol 8: 51.
18. Ioannidis JPA, Trikalinos TA (2005) Early extreme contradictory estimates may appear in published research: The Proteus phenomenon in molecular genetics research and randomized trials. J Clin Epidemiol 58: 543-549.
19. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR (2009) Introduction to meta-analysis. Chichester (UK): John Wiley & Sons.
20. Sanderson S, Salanti G, Higgins J (2007) 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 166: 741-751.
21. Higgins JPT, Thompson SG, Deeks JJ, Altman DG (2005) Measuring inconsistency in meta-analyses. BMJ 327: 557-560.
22. Higgins JPT, Thompson SG (2004) Controlling the risk of spurious findings from meta-regression. Stat Med 23: 1605-1622.
23. Munafò MR, Flint J (2004) Meta-analysis of genetic association studies. Trends Genet 20: 439-444.
24. Marchini J, Howie B (2008) Comparing algorithms for genotype imputation. Am J Hum Genet 85: 535-539.
25. Salanti G, Higgins JPT (2008) Meta-analysis of genetic association studies under different inheritance models using data reported as merged genotypes. Stat Med 27: 764-777.
26. Minelli C, Thompson JR, Abrams KR, Thakkinstian A, Attia J (2005) The choice of a genetic model in the meta-analysis of molecular association studies. Int J Epidemiol 34: 1319-1328.
27. Salanti G, Higgins JPT, Trikalinos TA, Ioannidis JPA (2007) Bayesian meta-analysis and meta-regression for gene-disease associations and deviations from Hardy-Weinberg equilibrium. Stat Med 26: 533-567.
28. Minelli C, Thompson JR, Abrams KR, Thakkinstian A, Attia J (2008) How should we use information about HWE in the meta-analyses of genetic association studies? Int J Epidemiol 37: 136-146.
29. Verzilli C, Shah T, Casas JP, Chapman J, Sandhu M, et al. (2008) Bayesian meta-analysis of genetic association studies with different sets of markers. Am J Hum Genet 82: 859-872.
30. Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, et al. (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.
31. Ioannidis JPA, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG (2001) Replication validity of genetic association studies. Nat Genet 29: 306-309.
32. 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.
33. Ioannidis JPA, Boffetta P, Little J, O’Brien TR, Uitterlinden AG, et al. (2008) Assessment of cumulative evidence on genetic associations: Interim guidelines. Int J Epidemiol 37: 120-132.
34. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, et al. (1999) Improving the quality of reports of meta-analyses of randomised controlled trials: The QUOROM statement. Quality of Reporting of Meta-analyses. Lancet 354: 1806-1809.
35. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, et al. (2000) Meta-analysis of observational studies in epidemiology: A proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA 283: 2088-2092.
36. Chalmers I (1993) The Cochrane Collaboration: Preparing, maintaining, and disseminating systematic reviews of the effects of health care. Ann NY Acad Sci 703: 156-163.
37. Patsopoulos NA, Anzalotis AA, Ioannidis JPA (2005) Relative citation impact of various study designs in the health sciences. JAMA 293: 2562-2566.
38. Bertram L, McQueen MB, Mullin K, Blacker D, Tanzi RE (2007) Systematic meta-analyses of Alzheimer disease genetic association studies: The AlzGene database. Nat Genet 39: 17-23.
39. Bagade S, Allen NC, Tanzi RE, Bertram L (2008) The PDGene Database. Alzheimer Research Forum. Available: http://www.pdgene.org/. Accessed 29 January 2009.
40. Allen NC, Bagade S, McQueen MB, Ioannidis JP, Kavvoura FK, et al. (2008) Systematic meta-analyses and field synopsis of genetic association studies in schizophrenia: The SfGene database. Nat Genet 40: 827-834.