How to Use an Article About Genetic Association: B: Are the Results of the Study Valid?

John Attia; John P. A. Ioannidis; Ammarin Thakkinstian; et al.

JAMA. 2009;301(2):191-197 (doi:10.1001/jama.2008.946)

http://jama.ama-assn.org/cgi/content/full/301/2/191

How to Use an Article About Genetic Association

B: Are the Results of the Study Valid?

John Attia, MD, PhD
John P. A. Ioannidis, MD, PhD
Ammarin Thakkinstian, PhD
Mark McEvoy, MMedSc
Rodney J. Scott, PhD
Cosetta Minelli, PhD
John Thompson, PhD
Claire Infante-Rivard, MD, PhD
Gordon Guyatt, MD, MSc

In the first article of this series, we reviewed the basic genetics concepts necessary to understand genetic association studies. In this second article, we enumerate the major issues in judging the validity of these studies, framed as critical appraisal questions. Was the disease phenotype properly defined and accurately recorded by someone blind to the genetic information? Have any potential differences between disease and nondisease groups, particularly ethnicity, been properly addressed? In genetic studies, one potential cause of spurious associations is differences between cases and controls in ethnicity, a situation termed *population stratification*. Was measurement of the genetic variants unbiased and accurate? Methods for determining DNA sequence variation are not perfect and may have some measurement error. Do the genotype proportions observe Hardy-Weinberg equilibrium? This simple mathematic rule about the distribution of genetic groups may be one way to check for errors in reading DNA information. Have the investigators adjusted their inferences for multiple comparisons? Given the thousands of genetic markers tested in genome-wide association studies, the potential for false-positive and false-negative results is much higher than in traditional medical studies, and it is particularly important to look for replication of results.

JAMA. 2009;301(2):191-197 www.jama.com

CLINICAL SCENARIO

A 55-year-old man with a family history of dementia is inquiring about genetic testing for Alzheimer disease, in particular testing for *APOE*. Armed with knowledge of the basic genetic concepts outlined in the introductory article of this 3-part series,[1] you return to your electronic medical reference discussion. Of the studies they cite, you focus on the largest study (n=6852), with longest follow-up (up to 9 years), representing a general, community-based population aged 55 years and older,[2] and using the stronger of the candidate gene study designs (cohort rather than case-control). The authors report a relative risk of 2.1 (95% confidence interval, 1.7-2.7) for dementia in *APOE e4* (e for epsilon) heterozygotes and 7.8 (95% confidence interval, 5.1-11.9) for *APOE e4* homozygotes compared with e3/e3 individuals.

Initial epidemiologic studies addressing a novel association tend to overestimate the magnitude of association,[3] possibly as a result of publication bias (studies addressing previously unreported associations are published only if they show significant results), and this phenomenon is even more frequent in genetic association studies.[4] This highlights the importance of examining the validity of such studies,[5-9] the focus of this article.

THE FRAMEWORK

We adopt the same framework as previous users' guides:

- Are the results of the study valid?
- What are the results?
- Will the results help me in caring for my patients?

This article deals with the first of these questions; the final article in the series will address the latter 2. BOX 1 provides a summary of our guides and BOX 2 provides a glossary of genetic terms.

ARE THE RESULTS OF THE STUDY VALID?

Similar to traditional prognostic or etiologic studies, genetic association may use cohort or case-control designs.[10-13] Cohort studies sample a group of people (eg, older individuals) who vary in their genetic characteristics (eg,

Author Affiliations are listed at the end of this article.
Corresponding Author: John Attia, MD, PhD, University of Newcastle/Royal Newcastle Hospital, Centre for Clinical Epidemiology and Biostatistics, Level 3, David Maddison Bldg, Newcastle 2300, Australia (john.attia@newcastle.edu.au).
Users' Guides to the Medical Literature Section Editor: Drummond Rennie, MD, Deputy Editor, *JAMA*.

Box 1. Critical Appraisal Guide to Genetic Association Studies

A. Are the results of the study valid?

Was the disease phenotype properly defined and accurately recorded by someone blind to the genetic information?

Have any potential differences between disease and nondisease groups, particularly ethnicity, been properly addressed?

Was measurement of the genetic variants unbiased and accurate?

Do the genotype proportions observe Hardy-Weinberg equilibrium?

Have the investigators adjusted their inferences for multiple comparisons?

Are the results consistent with those of other studies?

B. What are the results of the study?

How large and precise are the associations?

C. How can I apply the results to patient care?

Does the genetic association improve predictive power beyond easily measured clinical variables?

What are the absolute and relative effects?

Is the risk-associated allele likely to be present in my patient?

Is the patient likely better off knowing the genetic information?

APOE e2/e2, e2/e3, e2/e4) and follow them forward in time to determine who has the outcome of interest (eg, Alzheimer dementia). In case-control studies, investigators choose affected individuals (case patients, eg, those with Alzheimer dementia) and a sample of unaffected individuals from the same underlying population and determine the genetic characteristics of the individuals in each of the 2 groups.

Case-control studies in traditional epidemiology are subject to a number of potential biases, many of which are less of a concern in genetic studies. In contrast to most environmental exposures, the genetic "exposure" does not vary with age or calendar year, there is no recall bias and no choice of exposure made by the participant, and the exposure is not influenced by disease (or treatment). The case-control design also facilitates large sample sizes and therefore power, which is particularly important for detecting potentially small genetic effects. Our discussion will focus on validity issues of particular relevance for genetic studies.

Was the Disease Phenotype Properly Defined and Accurately Recorded by Someone Blind to the Genetic Information?

In the absence of a standardized definition of the disease or trait of interest, investigators may run association analyses with varying definitions and report only the most significant findings, resulting in spurious associations.[14] On the other hand, what appears at first glance to be a single disease entity may in fact consist of many genetically separate but clinically similar diseases, a situation called *genetic heterogeneity*. In this situation, including diseases with different genetic etiologies may dilute or obscure a true association.

Even if the disease definition is well standardized, it is important to ask whether the disease phenotype has been appropriately measured during the study. Misclassification (here, categorizing people as having dementia when they do not or vice versa) may affect the strength of the genetic association. If the misclassification is a result of random error, the association will be diluted. If misclassification errors are influenced by previous knowledge of the genotype of each individual, eg, if *APOE* genotype influences the diagnosis of dementia, then the genetic effect may be overestimated. Thus, individuals conducting the phenotyping should be blind to the genotyping result (and vice versa).

In our clinical scenario, because different etiologies of late dementia are likely to have different genetic determinants, researchers who do not separate individuals with Alzheimer disease from those with vascular dementia (common) and Lewy body dementia (rare) may fail to establish genetic links. Slooter et al[2] separate Alzheimer from vascular dementia and use widely accepted definitions. Moreover, the investigators made meticulous efforts to minimize misclassification caused by measurement error by using a panel of several tests and by blinding appropriately.

Have Any Potential Differences Between Disease and Nondisease Groups, Particularly Ancestry, Been Properly Addressed?

As we have pointed out, some common variables that, in traditional epidemiologic studies, can cause bias as a result of an association with the condition of interest and misdistribution in exposed and unexposed populations (we call such variables *confounders*) are less likely to introduce bias in genetic epidemiology. Genetic studies, however, may yield misleading results if their disease and nondisease populations include a different ethnic/racial mix; this particular form of confounding is referred to as *population stratification*. The problem occurs if the likelihood of developing the condition of interest varies with ancestry. If ancestry groups also happen to differ in allele frequency of genetic polymorphisms unrelated to the condition of interest, the result will be spurious associations.

Most association studies of unrelated individuals try to avoid this problem by using populations that are homogeneous in terms of ancestry. Self-reporting will usually suffice at least for populations of European ancestry,[15-17] although there are rare examples, such as genes that regulate susceptibility to lactose intolerance, in which there is marked variation. To address such possibilities, a number of techniques have been developed to check for differences in the potential mix of ancestries and, if differences are found, to make corrections; these corrections use self-reported ethnicity, family-based controls, or statistical techniques termed *genomic control* to test for pat-

terns in unlinked markers.[18,19] For example, a spurious association between the CYP3A4-V polymorphism and prostate cancer in blacks disappeared when results were adjusted for additional genetic markers associated with ancestry in the population studied.[20]

Ancestry is not the only potential confounder that may compromise the validity of a genetic association study. For example, 2 genome-wide association (GWA) studies showed an association between type 2 diabetes and a single-nucleotide polymorphism (SNP) in the *FTO* (fat mass and obesity associated) gene.[21,22] These studies selected diabetic patients and controls irrespective of their body mass index (BMI); another study that matched diabetic patients and controls on BMI showed no association. Thus, although the study accurately identified the association between diabetes and the particular SNP, the causal association is probably between the candidate allele and BMI regulation/obesity, not type 2 diabetes.

Readers should consider whether diseased and nondiseased groups were similar with respect to other important characteristics that are likely to be genetically determined and associated with the outcome of interest. Alternatively, they may determine whether the investigators adjusted for such characteristics.

Returning to the clinical scenario, one might imagine that ancestry and alcoholism are characteristics that are both genetically influenced and that would be associated with Alzheimer dementia. Slooter et al[2] recruited their entire cohort from among the white population of the Netherlands, which is likely a homogeneous group with little genetic variability; this is verified by results from a recent GWA study from the same cohort.[23] They did not, however, consider alcohol history.

Box 2. Glossary

Allele
One of several variants of a gene, usually referring to a specific site within the gene

Genetic heterogeneity
A situation in which a particular phenotype may result from more than one genetic variant

Genetic marker
A specific genetic variant known to be associated with a recognizable trait

Genome-wide association (GWA) study
A study that evaluates association of genetic variation with outcomes or traits of interest by using 100 000 to 1 000 000 or more markers across the genome

Genotype
The genetic constitution of an individual, either overall or at a specific gene

Haplotype
Alleles that tend to occur together on the same chromosome due to single-nucleotide polymorphisms (SNPs) being in proximity and therefore inherited together

Hardy-Weinberg equilibrium (HWE)
A situation in which a defined population displays constant genotype frequencies from generation to generation, and those genotype frequencies can be calculated from the allele frequencies based on the HWE formula

Heterozygous
An individual is heterozygous at a gene location if (s)he has 2 different alleles (one on the maternal chromosome, one on the paternal) at that location

Homozygous
An individual is homozygous at a gene location if (s)he has 2 identical alleles at that location

Phenotype
The observable characteristics of a cell or organism, usually being the result of the product coded by a gene (genotype)

Polymorphism
The existence of 2 or more variants of a gene, occurring in a population

Population stratification
Describes the situation in which a population may be composed of multiple subgroups of different ethnicity; case and control group differences in the mix can confound the comparison and lead to spurious genetic associations

Was Measurement of the Genetic Variants Unbiased and Accurate?

Genotyping error is a threat to the validity of genetic association studies. Genotyping may go wrong if there is a problem with the biological material (the samples) or with the application of the molecular technique that is used to call alleles.

The biological material that provides the source for genotyping may differ between diseased and nondiseased participants in ways that lead to inaccuracies in genotyping. For example, in a GWA study for type 2 diabetes, blood stored in 1958 provided the basis for genotyping nondiseased individuals, whereas blood drawn more recently was used for genotyping diseased individuals. The older blood resulted in genotyping errors[24] that led to some false-positive SNP associations.

Genotyping error may occur even when disease and nondisease samples are drawn and stored in identical ways. Although laboratory-based methods and DNA information may have the cachet of being absolute, these data are subject to error in the same way as traditional epidemiologic information. Genotyping error rates vary widely,

Box 3. Checking Hardy-Weinberg Equilibrium

Readers can check whether the data at a biallelic single-nucleotide polymorphism (SNP) are consistent with Hardy-Weinberg equilibrium (HWE) by inserting the numbers in each genotype group into an online program.[36] For example, an article may report that among 100 controls, there are 80 homozygote wild types, 12 heterozygotes, and 2 homozygous variants. The program calculates the expected distribution among the 3 genotype groups, the χ^2 value, and the corresponding *P* value.

Genotypes	*Observed, No.	Expected, No.
Homozygote reference	80	79.2
Heterozygote	18	19.6
Homozygote variant	2	1.2
Var allele frequency	0.11	

χ^2 = 0.65

χ^2 test *P* value = .42 with 1 degree of freedom.

(if <.05 then not consistent with HWE)

There are limitations to the hypothesis testing, whether done by the authors or the online program. Most HWE tests are weak because most sample sizes are small, and thus the likelihood of a false negative because of inadequate power is high. On the other hand, with very large sample sizes, the tests can detect very small deviations from HWE that are of no importance. In the setting of genome-wide association studies, a large number of SNPs are expected to have nominally significant deviations from HWE. For example, with 500 000 tested SNPs, 25 000 of them may have $P<.05$ on HWE testing by chance alone. Therefore, in GWA studies far more strict thresholds are appropriate to identify worrisome HWE deviation.

from less than 1% up to 30%,[25] and rates of up to a few percent are not uncommon in even the best studies.[26-28] Genome-wide association studies should aim to minimize genotyping error rates. Another useful piece of information is the "call rate" of genotyping, ie, the proportion of samples in which the genotyping provides an unambiguous reading. If this proportion is not high, then information is lost. In many studies, investigators decide to avoid analyzing SNPs in which the call rate is less than 90% or even less than 95%. Even high call rates can, however, fail to prevent bias if specific genotypes have lower call rates than others, eg, heterozygotes are more likely to get ambiguous readings or false readings than homozygotes.

These sources of error are most easily detected by the researchers using the raw data; it is impossible for a reader to identify them from the limited data usually reported in an article. A reader may, however, seek a description of how samples were handled, what genotyping method was used, whether any quality checks were implemented, whether any rules were established to say when the genotyping results would be considered valid, and the extent of missing data.

Returning to our clinical scenario, Slooter et al[2] refer to an earlier article from their team for genotyping details[29]; in this article, they state that genotyping was performed independently and in triplicate and without knowledge of the outcome status. They also state that their original cohort had 7983 persons, and they had to exclude 14% of the participants (n=1131) because *APOE* genotype could not be determined. There is no mention about whether this loss may have been related to underlying genotype or to Alzheimer disease, but at face value, it seems unlikely. Although the method was not specified, given the prospective cohort design, one may assume that samples were stored in similar conditions regardless of the subsequent development of dementia.

Do the Genotype Proportions Observe Hardy-Weinberg Equilibrium?

Failure to observe Hardy-Weinberg equilibrium (HWE) is one way of detecting possible genotyping error, although it is nonspecific and may be insensitive.[30-32] Investigators typically conduct statistical tests to check whether the observed genotype frequencies are consistent with HWE; $P<.05$ is the usual threshold for declaring Hardy-Weinberg "disequilibrium."[33] However, with simultaneous testing of a large number of possible associations, as in GWA studies, it is expected that 5% of SNPs will violate HWE simply because of multiple testing. In this setting, investigators may use more stringent *P*-value thresholds. Empirical studies suggest that disequilibrium is common and many articles do not explicitly acknowledge this[34,35]; as discussed in the first article, there are many reasons for disequilibrium (eg, inbreeding) aside from bias or error.

Therefore, readers should look for evidence that the investigators have tested for HWE and raise their level of skepticism about the results if they have not. Given that erroneous reports of HWE occur, they may even check for HWE themselves by using a simple freely available statistical program (**Box 3**). For a cohort study, HWE should be tested in the whole study population, whereas for a case-control study, it should be tested in the controls because these are supposedly representative of the general population.

In our scenario, Slooter et al[2] found that their study population did observe HWE ($P=.45$ in a well-powered study of n=6852). Given that this is a 3-allele system, we are not able to use the online program to check HWE.

Have the Investigators Adjusted for Multiple Comparisons?

One of the main reasons for false-positive results is inadequate attention to the problem of multiple comparisons. The scenario of an experiment testing 100 SNPs for association with a disease outcome in which no real association exists illustrates the magnitude of the problem. If the threshold *P* value of .05 is left unchanged, then the chance of finding an apparent but spurious positive association in this scenario can be calculated as $((1-(1-.05)^{100})\times 100)$, or 99.4%. The easiest method to correct for this problem of multiple comparisons is the Bonferroni method, in which the threshold *P* value is divided by the number of tests. In this example, the *P* value would be set at .05/100, or .0005. This is, however, probably overly conservative and stringent, and authors have suggested many other methods[17,37-40] (BOX 4). This potential for false-positive results also makes genetic association studies particularly susceptible to publication bias, in which initially strongly positive results find their way into publication more easily, whereas studies with negative results take longer to get published.[44] Such bias is not corrected by simply accounting for multiple comparisons.

In GWA studies, in which more than 500 000 SNPs are tested simultaneously, the multiple comparison problem takes on a magnitude never imagined in traditional epidemiology. To avoid false-positive results, a consensus seems to be forming that for such large-scale studies, a *P* value in the range of 5×10^{-8} (as opposed to the usual 5×10^{-2}) should be considered the threshold for claiming what is called "genome-wide" significance.[45,46] Increasingly, full results from GWA studies are publicly available, providing further insurance against publication bias.

In our scenario, Slooter et al[2] have not adjusted their results for multiple comparisons. They test only the *APOE* polymorphism (although they address 3 outcomes, myocardial infarction, stroke, and Alzheimer disease). They reasonably consider theirs a hypothesis-testing rather than hypothesis-generating study.

Are the Results Consistent With Those of Other Studies?

Any users' guide—whether for diagnosis, therapy, prognosis, or harm—could include a validity criterion demanding replication. Although we have not included this criterion in considering other sorts of individual studies, the multiple comparison problem and the forces that lead to differential publication of positive results suggest that, here, it is particularly important. Until results are replicated in similar populations, one should interpret them with caution.[8,47]

Most of the genetic associations between SNPs and complex diseases are small (much smaller than the odds ratios >2.0 observed for apoE e2/e3/e4),[48] and therefore even sizeable studies may fail to detect underlying associations.[49] Therefore, given that most individual studies are not large enough to detect these small effect sizes, typically, GWA studies pick the SNPs that have the lowest *P* values and test them in additional replication samples (either other GWA studies or focused studies targeting only the specific SNPs) to increase sample size and power until the cumulative results pass genome-wide significance or similar thresholds. Even more teams may then continue to try to replicate these associations, and all these data become essential in judging the credibility of these associations.

Therefore, just as we suggest that clinicians interested in issues of therapy, diagnosis, prognosis, and harm first seek a systematic review, so also do we suggest that they do the same for genetic associations.[50,51] The Human Genome Epidemiology Network (HuGE Net) group is emerging as the Cochrane equivalent for genetic association studies. The HuGE Net Web site lists many of the meta-analyses performed to date[52,53] and also hosts the HuGE Navigator, where one can determine what single studies, GWA studies, meta-analyses, and synopses are available.[54,55] Another possible aid in searching for previous genetic association studies is the genetic association database maintained by the National Institutes of Health.[56]

A MEDLINE search using *apoE* and *dementia* as search terms and restricted to English and meta-analysis, or a search on the HuGE Navigator, leads to 2 meta-analyses in the general population[57,58] and a Web site collating all the Alzheimer genetic association studies as an all-encompassing synopsis.[59] The meta-analyses demonstrate that results for the *APOE* e2/e3/e4 polymorphism are largely consistent across studies. This is probably the exception even among well-replicated genetic associations, and it reflects the fact that the apoE-dementia association is much stronger than almost any other associations recorded to date.

Box 4. Some Options for Adjustment for Multiple Comparisons

The Bonferroni correction is overly conservative and stringent, and there have been many suggestions for other methods. Two of the more popular ones include the following.

False-discovery rate calculations estimate the proportion of associations that are seemingly "discovered" (pass some required threshold of evidence) but are nevertheless expected to be false positives. The Benjamini-Hochberg method is used when loci (or single-nucleotide polymorphisms) are independent,[41] whereas the Benjamini-Lui method is applied when there is correlation or linkage disequilibrium between loci.[42] Both methods work on ranking the *P* values of the associations within one study and adjusting that *P* value by its position in the ranking list.

The *false-report probability rate* similarly states how likely an association is to be false if it emerges with a given level of statistical significance, given the power of the study and the perceived prior odds of an association being true.[17] The developers of this method have constructed a user-friendly spreadsheet to allow easy calculations.[43]

RETURN TO THE CLINICAL SCENARIO

Slooter et al[2] meet the crucial validity criteria:

- The authors defined a homogeneous group of dementia patients, separating Alzheimer from vascular dementia and using proper definitions and meticulous measurement schemes to determine outcomes.
- They chose a homogeneous ethnic group and provided a table showing similar characteristics in diseased and nondiseased groups, although alcohol is a significant confounder that is not included.
- They did not report sufficient information to ensure that genotyping error has been eliminated, but the population observes HWE and the association is too strong to be accounted for by genotyping error.
- They did not adjust for multiple comparisons in their study, but they studied only 1 polymorphism chosen according to previous work suggesting an association.
- Most important, the specific *APOE* association with Alzheimer dementia has been reproduced many times and meta-analyses of the results show consistent results across studies.

Given that we are satisfied with the validity of the study, we continue our critical appraisal. In the next article, we will discuss how to interpret results of genetic association studies and how to apply this information in the context of patient care.

Author Affiliations: Centre for Clinical Epidemiology and Biostatistics, University of Newcastle, Hunter Medical Research Institute, and Department of General Medicine, John Hunter Hospital, Newcastle, Australia (Dr Attia); Department of Hygiene and Epidemiology, University of Ioannina, School of Medicine, Ioannina, Greece, and Center for Genetic Epidemiology and Modeling, Tufts Medical Center, Department of Medicine, Tufts University School of Medicine, Boston, Massachusetts (Dr Ioannidis); Clinical Epidemiology Unit, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand (Dr Thakkinstian); Centre for Clinical Epidemiology and Biostatistics, University of Newcastle, Newcastle, Australia (Mr McEvoy); Division of Genetics, Hunter Area Pathology Service, John Hunter Hospital, New Lambton, Australia, and Centre for Information Based Medicine, Faculty of Health, University of Newcastle, Hunter Medical Research Institute, Newcastle, Australia (Dr Scott); Respiratory Epidemiology and Public Health, National Heart and Lung Institute, Imperial College, London, England (Dr Minelli); Department of Health Sciences, University of Leicester, Leicester, England (Dr Thompson); Department of Epidemiology, Biostatistics and Occupational Health, Faculty of Medicine, McGill University, Montreal, Canada (Dr Infante-Rivard); and the Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Canada (Dr Guyatt).

Author Contributions: Dr Attia had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Financial Disclosures: Dr Guyatt reports that his institution receives royalties from publication of the *Users' Guides to the Medical Literature* book. No other authors reported disclosures.
Additional Contributions: We wish to thank Julian Higgins, PhD (Cambridge University), and John Danesh, MBBS, DPhil (Cambridge University), for helpful comments on early drafts of this series.

REFERENCES

1. Attia J, Ioannidis JPA, Thakkinstian A, et al. How to use an article about genetic association: A: background concepts. *JAMA*. 2009;301(1):74-81.
2. Slooter AJ, Cruts M, Hofman A, et al. The impact of APOE on myocardial infarction, stroke, and dementia: the Rotterdam Study. *Neurology*. 2004; 62(7):1196-1198.
3. Ioannidis JP. Contradicted and initially stronger effects in highly cited clinical research. *JAMA*. 2005; 294(2):218-228.
4. Ioannidis JP, Trikalinos TA. Early extreme contradictory estimates may appear in published research: the Proteus phenomenon in molecular genetics research and randomized trials. *J Clin Epidemiol*. 2005; 58(6):543-549.
5. Colhoun HM, McKeigue PM, Davey Smith G. Problems of reporting genetic associations with complex outcomes. *Lancet*. 2003;361(9360):865-872.
6. Cordell HJ, Clayton DG. Genetic association studies. *Lancet*. 2005;366(9491):1121-1131.
7. Hattersley AT, McCarthy MI. What makes a good genetic association study? *Lancet*. 2005;366(9493): 1315-1323.
8. Ioannidis JP, Boffetta P, Little J, et al. Assessment of cumulative evidence on genetic associations: interim guidelines. *Int J Epidemiol*. 2008;37(1):120-132.
9. Little J, Bradley L, Bray MS, et al. Reporting, appraising, and integrating data on genotype prevalence and gene-disease associations. *Am J Epidemiol*. 2002;156(4):300-310.
10. Laupacis A, Wells G, Richardson WS, Tugwell P; Evidence-Based Medicine Working Group. Users' guides to the medical literature, V: how to use an article about prognosis. *JAMA*. 1994;272(3):234-237.
11. Randolph A, Cook DJ, Guyatt G. Prognosis. In: Guyatt G, Rennie D, Meade M, Cook D, eds. *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice*. New York, NY: McGraw-Hill; 2008.
12. Levine M, Walter S, Lee H, Haines T, Holbrook A, Moyer V; Evidence-Based Medicine Working Group. Users' guides to the medical literature, IV: how to use an article about harm. *JAMA*. 1994;271(20):1615-1619.
13. Levine M, Ioannidis J, Haines T, Guyatt G. Harm (observational studies). In: Guyatt G, Rennie D, Meade M, Cook D, eds. *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice*. New York, NY: McGraw-Hill; 2008.
14. Contopoulos-Ioannidis DG, Alexiou GA, Gouvias TC, Ioannidis JP. An empirical evaluation of multifarious outcomes in pharmacogenetics: beta-2 adrenoceptor gene polymorphisms in asthma treatment. *Pharmacogenet Genomics*. 2006;16(10):705-711.
15. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007;447(7145):661-678.
16. Evangelou E, Trikalinos TA, Salanti G, Ioannidis JP. Family-based versus unrelated case-control designs for genetic associations. *PLoS Genet*. 2006; 2(8):e123.
17. Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. *J Natl Cancer Inst*. 2004; 96(6):434-442.
18. Barnholtz-Sloan JS, McEvoy B, Shriver MD, Rebbeck TR. Ancestry estimation and correction for population stratification in molecular epidemiologic association studies. *Cancer Epidemiol Biomarkers Prev*. 2008;17(3):471-477.
19. Pritchard JK, Rosenberg NA. Use of unlinked genetic markers to detect population stratification in association studies. *Am J Hum Genet*. 1999;65(1): 220-228.
20. Kittles RA, Chen W, Panguluri RK, et al. CYP3A4-V and prostate cancer in African Americans: causal or confounding association because of population stratification? *Hum Genet*. 2002;110(6):553-560.
21. Frayling TM, Timpson NJ, Weedon MN, et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science*. 2007;316(5826):889-894.
22. Zeggini E, Weedon MN, Lindgren CM, et al; Wellcome Trust Case Control Consortium (WTCCC). Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science*. 2007;316(5829):1336-1341.
23. Richards JB, Rivadeneira F, Inouye M, et al. Bone mineral density, osteoporosis, and osteoporotic fractures: a genome-wide association study. *Lancet*. 2008; 371(9623):1505-1512.
24. Clayton DG, Walker NM, Smyth DJ, et al. Population structure, differential bias and genomic control in a large-scale, case-control association study. *Nat Genet*. 2005;37(11):1243-1246.
25. Akey JM, Zhang K, Xiong M, Doris P, Jin L. The effect that genotyping errors have on the robustness of common linkage-disequilibrium measures. *Am J Hum Genet*. 2001;68(6):1447-1456.
26. Bogardus ST Jr, Concato J, Feinstein AR. Clinical epidemiological quality in molecular genetic research: the need for methodological standards. *JAMA*. 1999;281(20):1919-1926.
27. Mein CA, Barratt BJ, Dunn MG, et al. Evaluation of single nucleotide polymorphism typing with invader on PCR amplicons and its automation. *Genome Res*. 2000;10(3):330-343.
28. Pompanon F, Bonin A, Bellemain E, Taberlet P. Genotyping errors: causes, consequences and solutions. *Nat Rev Genet*. 2005;6(11):847-859.
29. Slooter AJ, Cruts M, Kalmijn S, et al. Risk estimates of dementia by apolipoprotein E genotypes from a population-based incidence study: the Rotterdam Study. *Arch Neurol*. 1998;55(7):964-968.
30. Cox DG, Kraft P. Quantification of the power of Hardy-Weinberg equilibrium testing to detect genotyping error. *Hum Hered*. 2006;61(1):10-14.
31. Hosking L, Lumsden S, Lewis K, et al. Detection of genotyping errors by Hardy-Weinberg equilibrium testing. *Eur J Hum Genet*. 2004;12(5):395-399.
32. Leal SM. Detection of genotyping errors and pseudo-SNPs via deviations from Hardy-Weinberg equilibrium. *Genet Epidemiol*. 2005;29(3):204-214.
33. Montoya-Delgado LE, Irony TZ, de B Pereira CA, Whittle MR. An unconditional exact test for the Hardy-Weinberg equilibrium law: sample-space ordering using the Bayes factor. *Genetics*. 2001;158(2): 875-883.

34. Salanti G, Amountza G, Ntzani EE, Ioannidis JP. Hardy-Weinberg equilibrium in genetic association studies: an empirical evaluation of reporting, deviations, and power. *Eur J Hum Genet*. 2005;13(7):840-848.
35. Xu J, Turner A, Little J, Bleecker ER, Meyers DA. Positive results in association studies are associated with departure from Hardy-Weinberg equilibrium: hint for genotyping error? *Hum Genet*. 2002;111(6):573-574.
36. Tufts University Comparative and Molecular Pharmacogenomics Laboratory. A simple calculator to determine whether observed genotype frequencies are consistent with Hardy-Weinberg equilibrium. http://www.tufts.edu/~mcourt01/Documents/Court%20lab%20-%20HW%20calculator.xls. Accessed October 25, 2008.
37. Freimer N, Sabatti C. The use of pedigree, sib-pair and association studies of common diseases for genetic mapping and epidemiology. *Nat Genet*. 2004;36(10):1045-1051.
38. Ioannidis JP. Calibration of credibility of agnostic genome-wide associations. *Am J Med Genet B Neuropsychiatr Genet*. 2008. doi:10.1002/ajmg.b.30721.
39. Province MA. Sequential methods of analysis for genome scans. *Adv Genet*. 2001;42:499-514.
40. Sabatti C. Avoiding false discoveries in association studies. *Methods Mol Biol*. 2007;376:195-211.
41. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B*. 1995;57:289-300.
42. Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. *Ann Stat*. 2001;29:1165-1188.
43. Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. http://jnci.oxfordjournals.org/cgi/content/full/96/6/434/DC1?maxtoshow=&HITS=10&hits=10&RESULTFORMAT=1&author1=wacholder&andorexacttitle=and&andorexacttitleabs=and&andorexactfulltext=and&searchid=1&FIRSTINDEX=0&sortspec=relevance&resourcetype=HWCIT. Accessed October 25, 2008.
44. Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG. Replication validity of genetic association studies. *Nat Genet*. 2001;29(3):306-309.
45. Hoggart CJ, Clark TG, De Iorio M, Whittaker JC, Balding DJ. Genome-wide significance for dense SNP and resequencing data. *Genet Epidemiol*. 2008;32(2):179-185.
46. McCarthy MI, Abecasis GR, Cardon LR, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet*. 2008;9(5):356-369.
47. Chanock SJ, Manolio T, Boehnke M, et al; NCI-NHGRI Working Group on Replication in Association Studies. Replicating genotype-phenotype associations. *Nature*. 2007;447(7145):655-660.
48. Ioannidis JP, Trikalinos TA, Khoury MJ. Implications of small effect sizes of individual genetic variants on the design and interpretation of genetic association studies of complex diseases. *Am J Epidemiol*. 2006;164(7):609-614.
49. Moonesinghe R, Khoury MJ, Liu T, Ioannidis JP. Required sample size and nonreplicability thresholds for heterogeneous genetic associations. *Proc Natl Acad Sci U S A*. 2008;105(2):617-622.
50. Munafo MR, Flint J. Meta-analysis of genetic association studies. *Trends Genet*. 2004;20(9):439-444.
51. Salanti G, Sanderson S, Higgins JP. Obstacles and opportunities in meta-analysis of genetic association studies. *Genet Med*. 2005;7(1):13-20.
52. Khoury MJ, Dorman JS. The Human Genome Epidemiology Network. *Am J Epidemiol*. 1998;148(1):1-3.
53. Little J, Higgins J. *The HuGENet HuGE Review Handbook*. Ottawa, Ontario, Canada: HuGENet Canada; 2008. http://www.hugenet.ca. Accessed October 25, 2008.
54. Human Genome Epidemiology Network. A navigator for human genome epidemiology. http://www.hugenavigator.net/. Accessed October 25, 2008.
55. Yu W, Gwinn M, Clyne M, Yesupriya A, Khoury MJ. A navigator for human genome epidemiology. *Nat Genet*. 2008;40(2):124-125.
56. Becker KG, Barnes KC, Bright TJ, Wang SA. The genetic association database. *Nat Genet*. 2004;36(5):431-432.
57. Farrer LA, Cupples LA, Haines JL, et al; for the APOE and Alzheimer Disease Meta Analysis Consortium. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease: a meta-analysis. *JAMA*. 1997;278(16):1349-1356.
58. Rubinsztein DC, Easton DF. Apolipoprotein E genetic variation and Alzheimer's disease: a meta-analysis. *Dement Geriatr Cogn Disord*. 1999;10(3):199-209.
59. Bertram L, McQueen MB, Mullin K, Blacker D, Tanzi RE. Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. *Nat Genet*. 2007;39(1):17-23.