ELSEVIER

Journal of Clinical Epidemiology 56 (2003) 297–303

Journal of Clinical Epidemiology

Meta-analyses of molecular association studies: Methodologic lessons for genetic epidemiology

John Attia[a,*], Ammarin Thakkinstian[b], Catherine D'Este[a]

[a]*Centre for Clinical Epidemiology and Biostatistics, Faculty of Health, University of Newcastle, Royal Newcastle Hospital, David Maddison Building, Newcastle NSW 2300, Australia*

[b]*Clinical Epidemiology Unit, Faculty of Medicine, Mahidol University, Ramathibodi Hospital, Rama 6 Road, Bangkok 10400, Thailand*

Received 15 April 2002; received in revised form 19 September 2002; accepted 21 November 2002

Abstract

Meta-analyses of population-based molecular association studies have become increasingly common over the last 10 years, but little attention has been paid to methodology. In addition to the traditional considerations pertinent to any meta-analysis, there are genetic issues particular to molecular association studies: checking Hardy-Weinberg equilibrium, handling data from more than two groups while avoiding multiple comparisons, and pooling data in a way that is sensitive to genetic models. We systematically reviewed all meta-analyses of molecular association studies identified via MEDLINE. Of a total of 37 studies, eight (22%) described the search terms. Nineteen (51%) did not state inclusion or exclusion criteria. Heterogeneity was assessed in 28 (76%), but only 7 of 37 (19%) studies checked for publication bias. Nine (24%) studies assessed the goodness-of-fit of Hardy-Weinberg equilibrium, and eight (22%) gave any biological rationale to justify the choice of genetic model used for pooling. There is a need for greater communication between epidemiologists and geneticists to develop methods appropriate to this area.

Keywords: Systematic review; Meta-analysis; Molecular association studies; Hardy-Weinberg equilibrium; Multiple comparisons; Genetic models

1. Introduction

Meta-analysis has been increasingly used for summarizing data in many fields of medical research. It is a powerful method for pooling or aggregating previous research when individual studies have insufficient power to detect an association, for exploring sources of heterogeneity, or for identifying subgroups associated with the outcome of interest.

The reliability of the results from systematic reviews and meta-analyses of observational studies depends not only on the validity of the primary studies included in the analyses, but also on rigorous methodology. These methodologic considerations have been articulated in detail and include such issues as defining a clear research question; explicitly stating search strategies, sources, and inclusion/exclusion criteria; assessing (and, if appropriate, exploring) heterogeneity; pooling data using appropriate statistical methods; and investigating publication bias [1].

With the rise of genetic and molecular epidemiology, population-based molecular association studies have become a popular research design. These are usually observational designs testing the association of a genetic polymorphism with a dichotomous disease outcome (e.g., the association of ACE gene polymorphism with ischemic stroke [2] or dopamine receptor 2 gene with schizophrenia [3]). For continuous outcomes, most designs are cross-sectional or cohort studies that test mean differences in the phenotype of interest among those with different genotypes/gene polymorphisms (e.g., apolipoprotein B gene polymorphism and serum lipids [4] or body mass index and Trp64Arg polymorphism in the beta 3 adrenergic receptor [5]). Definitions of the terms used in these studies are shown in Table 1 (based on definitions in [6]).

We undertook this review to summarize the effect of the vitamin D receptor gene polymorphisms on bone mineral density and risk of fractures. Systematic reviews and meta-analyses of population-based molecular association studies are subject to the traditional epidemiologic methodologic considerations, but there are additional considerations stemming from the genetic issues particular to these studies. These include checking Hardy-Weinberg equilibrium (HWE), handling data from more than two groups while avoiding multiple comparisons, and pooling data in a way that is sensitive to genetic models. These considerations are not novel but have not been explicitly stated in the literature. In response to these considerations, we embarked on a systematic review of all previous meta-analyses of

* Corresponding author. Tel.: +61-2492-36152; fax: +61-2492-36148.
E-mail address: jattia@cceb.newcastle.edu.au (J. Attia).

0895-4356/03/$ – see front matter
doi: 10.1016/S0895-4356(03)00011-8

Table 1
Definitions of terms in molecular association studies

Term	Definition
Admixture	Describes the varying ethnic origins of one person's genetic make-up. For example, because of intermarriage with caucasians, the genome of American Indians may reflect varying degrees of admixture (i.e., varying mixtures of native and caucasian genes)
Allele	One of several variants of a gene, usually referring to a specific site within the gene
Codominant	Describes any trait that increases proportionately in expression when comparing those with no copy, one copy, or two copies of that allele (i.e., those with one copy of the allele show more of the trait than those without, and those with two copies show more of the trait than those with one copy)
Complete overdominant	Describes the situation where both alleles of a pair are fully expressed in heterozygotes (i.e., both alleles have an effect even when present in one copy each)
Dominant	Describes any trait that is expressed in a heterozygote (i.e., one copy of that allele is sufficient to manifest its effect)
Genotype	The genetic constitution of an individual, either overall or at a specific gene
Heterozygous	An individual is heterozygous at a gene location if (s)he has two different alleles at that location
Homozygous	An individual is homozygous at a gene location if (s)he has two identical alleles at that location
Linkage	The tendency of genes or other DNA sequences at specific loci to be inherited together as a consequence of their physical proximity on a single chromosome
Phenotype	The observable characteristics of a cell or organism, usually the result of the product coded by a gene (genotype)
Polymorphism	The existence of two or more variants in a gene, occurring at significant frequencies in the population
Population stratification	Describes the situation where a population may be composed of multiple subgroups of different ethnicity. Where case and control groups differ in the make up of their subgroups, this may lead to confounding if some subgroups differ with respect to the outcome of interest
Recessive	Describes any trait that is expressed in a homozygote (i.e., two copies of that allele are necessary to manifest its effect)

population-based molecular association studies with two objectives: (1) to review the methodologic rigor of these meta-analyses along traditional epidemiologic criteria and (2) to determine how these meta-analyses have addressed the additional genetic considerations outlined above.

2. Methods

2.1. Database and searching strategy

We searched Medline (January 1966 through August 2000, OVID version 7.8, Millenium Source ID 1.3932.1.156, revision 1.303) using the following strategy:

1. exp genetics (MeSH heading)
2. exp alleles (MeSH heading)
3. exp polymorphisms (MeSH heading)
4. 1 or 2 or 3
5. Limit 4 to human, English language, and meta-analysis (publication type)

The search yielded 162 references. Forty-two studies were identified as potentially relevant from the abstracts, and full-text versions were obtained. Five studies were excluded from review: Two were not meta-analyses, two were not polymorphism studies, and one was an article about statistical methods.

2.2. Data extraction

Two authors (JA and AT) carefully reviewed the remaining 37 meta-studies. Data collection sheets were independently completed for each study. Disagreements in data extraction were resolved by discussion and review of the original article. Because there is no validated quality scoring system for molecular association studies, we selected five items that we felt were most reflective of methodologic rigor from the list of 35 outlined in the proposed guidelines for reporting meta-analysis of observational studies [1]. This does not deny the importance of the items not chosen, although we judged these not to be paramount to quality (e.g., statement of hypothesis, disclosure of funding source, reporting use of hand searching, etc.). Data on the following points were collected:

- Explicit search terms and sources: Did the authors state their search terms and sources in a manner that would allow replication?
- Explicit inclusion/exclusion criteria: Did the authors clearly state their criteria for inclusion and exclusion of studies?
- Heterogeneity test: Did the authors perform any test of heterogeneity? What level of significance did they set? If there was evidence of heterogeneity, did they try to identify the cause?
- Pooling results: Did the authors use accepted statistical methods to pool results from individual studies?
- Publication bias: Did the authors explore the possibility of publication bias? How did they assess it (e.g., funnel plot)?

In addition, data on the two following issues, particular to population-based molecular association studies, were collected:

- The HWE test: Did the authors check HWE for each included study, and if they did so, how did they assess it?
- Choice of genetic model: What genetic model did the authors use to pool data (e.g., allele frequencies, choosing one genetic model, testing multiple models, or

Table 2
Summary of meta-analyses of molecular association studies ($n = 37$)

Characteristics	Number (%)
Explicit statement of databases used	
Yes	24 (65)
No	13 (35)
Explicit search terms	
Yes	9 (24)
No	28 (76)
Explicit inclusion/exclusion criteria	
Yes/yes	4 (11)
Yes/no	9 (24)
No/yes	5 (14)
No/no	19 (51)
Performing HWE test	
Yes	9 (24)
Possibly	1 (3)
No	27 (73)
Type of outcome	
Continuous	5 (14)
Dichotomous	30 (81)
Both	2 (5)
Type of genetic model	
Alleles	9 (24)
Dominant	7 (19)
Recessive	4 (11)
Homozygous	1 (3)
Multiple comparisons	16 (43)
Correct type one error	
Yes	2 (17)
No	14 (83)
Method of pooling effect	
Continuous outcome: seven studies	
Standardized mean difference	3 (43)
ANOVA	1 (14)
Standardized mean difference + regression	1 (14)
Standardized percentage mean difference	2 (29)
Dichotomous outcome: 32 studies	
Pooled odds ratio	21 (66)
Pooled risk ratio	3 (9)
Chi-square	3 (9)
Logistic regression + pooled odds ratio	4 (13)
Pooled frequency	1 (3)
Type of meta-analysis model	
Continuous outcome: seven studies	
Fixed effects model	7 (100)
Dichotomous outcome: 32 studies	
Fixed effects model	23 (72)
Random effects model	5 (16)
Fixed + random effects model	2 (6)
Not mentioned	1 (3)
No model	1 (3)
Heterogeneity assessment	
Yes	28 (76)
No	9 (24)
Publication bias check	
Yes	7 (19)
No	30 (81)

Abbreviations: HWE, Hardy-Weinberg equilibrium.

multiple pairwise comparisons)? Did the authors give any biologic rationale for the model used? If they used multiple genetic models, did they correct for multiple comparisons?

To address the effect of author affiliation on methodologic rigor, we created an ad-hoc quality score that consisted of the number of appropriately addressed considerations from the list of five described above (excluding the two genetic considerations). For the sake of clarity, we used "meta-analysis" to refer to the process of pooling studies, "meta-study" to refer to a study using the meta-analysis process, and "study/studies" to refer to the individual component studies that are pooled in a meta-study.

3. Results

3.1. General observations

Thirty-seven meta-studies were included in this review [2–5,7–39]. The majority of reviews were in cardiovascular medicine (12 meta-studies, 32%) and psychiatry (10 meta-studies, 27%), followed by oncology (5 meta-studies, 14%) and neurology (4 meta-studies, 11%). Thirty meta-studies addressed a dichotomous outcome (81%) [2,3,7–9,11–14,16–23,25,26,28–31,33–39], five addressed a continuous outcome (14%) [4,5,10,24,32], and two (5%) addressed both types of outcomes within the same meta-study [15,27]. The latter meta-studies were scored for overall methods (i.e., counted as one article) rather than being scored for methods for each outcome (i.e., counted as two articles).

As a simple exploratory analysis, we judged investigator expertise from author affiliations stated in the papers; 20 (54%) meta-studies [3,4,7–10,13,15,18,21–25,27,30,31,34–36] were undertaken by groups with primary genetic or clinical expertise, and 17 (46%) [2,5,11,12,14,16,17,19,20,26,28,29,32,33,37–39] were undertaken by teams that included at least one epidemiologist or statistician. Table 2 summarizes the results.

3.2. Explicit search terms and sources

Only nine (24%) meta-studies explicitly described the search terms [11,13,16,22–24,32,33,39]. Twenty-four meta-studies (65%) stated the databases used to identify studies [2,4,5,7,11,13,15,16,19,21–24,27–35,38,39]. MEDLINE was almost exclusively used. A recent study indicates that MEDLINE may miss a significant number of clinical trials and that other databases (e.g., EMBASE) and hand searching and personal contact should be used [40]. However, the quality of the additional trials found by these methods was not different from those found on MEDLINE alone, and the bias created is likely to be small. This issue has not been similarly explored specifically for molecular association studies.

3.3. Explicit inclusion/exclusion criteria

Nine (24%) meta-studies described only inclusion criteria [2,5,7,11,13,15,17,23,39], five (14%) meta-studies described only exclusion criteria [16,22,25,30,31], four (11%) meta-studies described both [32–35], and 19 (51%) meta-studies

did not describe either [3,4,8–10,12,14,18–21,24,26–29,36–38].

3.4. Assessing heterogeneity

Assessing heterogeneity means answering the question "How different are the results of these studies?" This question may be answered clinically (e.g., by looking at the varying populations, patient characteristics, disease severity, diagnostic criteria, treatments, outcome measures between studies, etc.) or statistically by using various tests (e.g., Breslow-Day test or Q-test). Because these tests have low power to detect heterogeneity [41,42], the threshold P value should be increased to 0.10 or higher [43].

Heterogeneity has a direct bearing on deciding whether and how to pool results. When there is no heterogeneity, some advocate a fixed-effects model [44], whereas others advocate a random-effects model [45,46], and some recommend using both methods and comparing the results [47,48]. Pooling may be done using fixed- or random-effects models. The fixed-effects model assumes the treatment effect to be the same in each study and therefore includes only within-study variance terms. The random-effects model estimates the average treatment effects and variability between studies and therefore incorporates between-study variance terms. The arguments for using one model rather than another are summarized by Petiti [41] and Egger [49]; in practice, most authors use the fixed effects model when there is no heterogeneity [44] and a random effects model when there is heterogeneity [45,46]. Some authors, however, have argued that the focus should be on trying to understand the sources of heterogeneity rather than providing a possibly meaningless summary measure [44,50–53].

We found that heterogeneity was assessed statistically in 28 (76%) meta-studies. All studies used standard methods (e.g., Q-test). Thirteen (46%) meta-studies found heterogeneity [7–9,14–16,22,27,29,32,34,37,38], but only eight of these explored the reasons for the heterogeneity [7,8,14,15,27,29,32,34]. Five of these eight meta-studies [7,8,14,15,27] performed a sensitivity analysis (e.g., pooling results with and without the studies suspected to be the source of heterogeneity). Subgroup analyses according to suspected sources of heterogeneity were performed in the other three meta-studies (29, 32, 34). All of these 13 meta-studies pooled results: Four used the random effects model [7,16,22,34], seven used the fixed effects model [8,14,15,27,29,32,38], one used both [37], and one did not specify [9].

Of the 15 meta-studies that were homogeneous, 13 used the fixed effects model to pool results [5,11–13,19–21,23,28,30,33,36,39], one used the random effects model [35], and one used both [24]. Only one meta-study stated the cut-off level of significance used for declaring heterogeneity ($P < 0.10$) [16].

3.5. Publication bias

Publication bias refers to the possibility that negative studies are not published as readily as positive studies; a meta-analysis that depends solely on published studies may therefore give a falsely skewed positive result. A graphic method of detecting this bias has been suggested [54,55], plotting effect size on the vertical axis versus some measure of precision (e.g., sample size or standard error) on the horizontal axis. Smaller studies have greater scatter, whereas larger studies have smaller scatter; if there is no bias, this leads to a triangle shape (inverted funnel) to the plot; bias leads to asymmetry. This asymmetry may be detected by visual inspection or by formal statistical tests (e.g., Begg's or Egger's test), although the use of these is debated [42,55,56].

We found that only 7 of 37 (19%) meta-studies checked for publication bias [5,12,15,21,27,28,38]. Visual inspection of a funnel plot was the sole method; no study used the formal statistical tests available for assessing funnel plots.

3.6. HWE

The HWE law states that if two alleles, A and a, with frequencies p and q, respectively, are in equilibrium in a population, the proportions of those with genotypes AA, Aa, and aa will be p^2, 2pq, and q^2, respectively. Deviation from these proportions has traditionally been taken as an indication that the alleles are not segregating independently, that there is nonrandom mating, or that the alleles reflect recent mutations that have not reached equilibrium. However, three other possibilities need to be considered: genotyping error in the lab, the presence of bias in the selection of control subjects, or the existence of population stratification. Therefore, goodness-of-fit of HWE [57,58] should be checked before pooling data. However, for pooling case-control studies, subjects in the case group are more likely to represent a skewed population. We therefore agree with Schaid et al. [58] that HWE should be assessed only in control subjects. There is no consensus on whether to include studies that are not in HWE. We suggest that sensitivity analysis should be performed, pooling with and without studies not in HWE, to test the robustness of the results. In our review, we noted that only nine (24%) meta-studies assessed the goodness-of-fit of HWE for each included study [3,14,21,24,25,31,33,35,38]. Among them, six meta-studies tested only the control group [21,24,25,31,35,38], two meta-studies tested case and control groups separately [14,33], and one meta-study did not describe the method [3]. Only two meta-studies stated that they excluded studies that were not in equilibrium [25,35].

3.7. Multiple groups, multiple comparisons, and genetic models

Usual meta-analysis pools studies in which there are two groups (treatment/control, exposure/non-exposure); in these cases, pooling is simply a statistical process. By contrast, genetic association studies have a minimum of three (genetic risk) groups, and these must be pooled in a way that reflects the biology of the genetic model. For example, if the alleles of the gene of interest are A and a, and A is the "increasing"

allele (i.e., the one causing an effect), the three genotype groups are AA, Aa, and aa. Dichotomizing these three groups can be done in many ways: by comparing allele frequency (A and a), by assuming a genetic model (dominant: AA + Aa versus aa; recessive: AA versus Aa + aa; complete overdominant or homozygous: AA + aa versus Aa; or codominant: AA versus Aa versus aa, see Table 1), by multiple pairwise comparisons (e.g., AA versus aa, AA versus Aa, or Aa versus aa), or by assessing many of the above models at the cost of multiple comparisons. In this last case, the *P* value must be adjusted to account for the multiple comparisons. In our review, we noted that nine (24%) meta-studies determined association solely by comparing allele frequency between cases and control subjects [13,14,17,18,26,30,33,37,38].

Twelve (33%) meta-studies determined the effect by using only one genetic model: seven (19%) meta-studies used the dominant model [7–9,11,23,24,32], one (3%) meta-study used a homozygous model [21], and four (11%) meta-studies used the recessive model [16,22,34,39]. Of these 12, eight meta-studies gave implicit or explicit biological reasons for choosing one particular genetic model [7,8,11,16,22,24,34,39], one meta-study pooled using the model suggested by previous studies [21], and three meta-studies did not give any biological reasoning [9,23,32]. Sixteen (43%) meta-studies determined the effect using multiple comparisons [2–5,10, 12,15,19,20,25,27–29,31,35,36]. Among these, only two adjusted the *P* value for the inflated probability of type one error. One meta-study [36] corrected using the Bonferroni method, and another meta-study [25] did not describe the method of correction.

3.8. Study quality by author affiliation

We examined whether epidemiologists or geneticists met the methodologic criteria for their meta-studies differentially. We derived a quality score based on five traditional considerations (explicit search terms and sources, explicit inclusion/exclusion criteria, assessing heterogeneity, assessing publication bias, and pooling appropriately, with a total score of 5). The following analysis is a rough exploration of the data given the possible misclassification resulting from judging author expertise by affiliation (as stated in the papers) and the possible error in judging quality using scores [59].

The 17 meta-studies [2,5,11,12,14,16,17,19,20,26,28,29, 32,33,37–39] that included at least one author with an epidemiologic or statistical affiliation met a median score of two traditional epidemiologic criteria (range 0–4), as did the 20 meta-studies [3,4,7–10,13,15,18,21–25,27,30,31,34–36] with primarily genetics or clinical expertise (range 0–3; Mann-Whitney test, $P = .520$). Geneticists and clinicians did a little better with the genetic considerations: 7 of 20 meta-studies (35%) checked HWE, compared with only 3 of 17 (18%) with epidemiologic or statistical input. Groups with epidemiologic input expressed results using allele frequencies more frequently than genetics or clinical groups (6 of 17 [35%] versus 3 of 20 [15%]), which is a paradoxical finding given that this method assumes HWE. Geneticists chose a single genetic model more frequently than epidemiologists (8 of 20 [40%] versus 4 of 17 [23%]), but both groups gave biologic rationale for their choice as often as each other (5 of 8 [63%] versus 3 of 4 [75%]). The most common method in both groups was to test multiple models (9 of 20 [45%] for geneticists versus 7 of 17 [41%] for epidemiologists).

4. Discussion

We have critically reviewed the methods used in the meta-analysis of molecular association studies. Although we did not search other databases (e.g., EMBASE) and therefore may have missed some meta-studies, we do not believe that this would significantly affect our results. Unlike systematic reviews of an effect size, where missing studies may affect their pooled magnitude, in our review we focused on an appraisal of methods. The reviews we included presumably reflect those published in the highest-impact journals and should reflect the highest-quality papers. This makes the methodologic failings we discuss even more pertinent.

We encourage meta-analysts of population-based molecular association studies to adhere to guidelines derived for traditional meta-analyses [1]. In our review, almost 80% did not describe search terms, and 35% did not state their sources. Half did not state any inclusion/exclusion criteria. Most meta-studies (80%) did not check for publication bias. One quarter of meta-studies pooled results without checking heterogeneity, and of those that found heterogeneity, all went on to pool results, most (70%) using a fixed effects model. We agree with Poole et al. [60] that, in some of these cases, it would be more helpful not to pool or to pool only subsets of data based on biological considerations (in particular ethnicity).

Exploring heterogeneity is just as important in population-based molecular association studies as in traditional epidemiologic studies. Although genetic studies are less susceptible to some sources of bias (e.g., misclassification or measurement error), there are additional sources of heterogeneity unique to genetic studies. Population stratification is a much-discussed example; this is a phenomenon whereby allele frequency and disease outcome vary between subpopulations (e.g., ethnic group), and the varying proportions of these subpopulations in the case and control groups confound the results [61,62]. Other unique potential sources of heterogeneity include population admixture (whereby the genotype of the individuals reflects the mixture of various ethnic backgrounds) and allelic heterogeneity. As in traditional meta-analysis, pooling despite the presence of heterogeneity may give misleading results. As stated by Poole and Greenland, "Meta-analysts rarely entertain the possibility of concluding that study-specific results are too heterogeneous to aggregate" [60].

We also highlight the need to more fully integrate genetic concepts in the meta-analytic process. Basic concepts, such as HWE, were not addressed in most studies.

The most surprising result of our study was the lack of coherence or transparency in choosing a genetic model to pool results. The major differences between meta-analyses of traditional epidemiologic studies and molecular association are the presence of at least three effect groups (for a diallelic locus) and the need to pool results in a way that reflects the underlying biology of the genetic effect. A number of meta-studies compared allele frequencies in case and control subjects and tested for the difference in proportions. We suggest that this method be discouraged because it yields a test of significance but does not provide any estimate of the magnitude of the risk associated with a particular genotype. Strictly speaking, odds ratios derived in this way represent the odds of having that allele given that one is a case or a control subject; this does not easily translate into a risk of disease given a specific genotype. We suggest that potential methods for pooling should express genotype risk in a clinically useful manner, analogous to a genotype-relative risk. Some meta-studies determined the effect of genotype by specifying a single mode of inheritance (e.g., dominant, recessive, complete overdominant, or codominant), but half of the meta-studies tested multiple modes of inheritance; most did not correct for these multiple comparisons, which increases the risk of false positives. In addition, only a quarter of meta-studies based their choice of a genetic model on any biologic rationale of gene effect.

Our review highlights the need to develop new methodology that allows pooling according to a genetic model when biologic data indicate one and that avoids multiple comparisons when such evidence is lacking. The recent development of "model-free" methods for genetic linkage needs to carry over into this area as well. Whereas traditional linkage analysis for Mendelian diseases has used family pedigrees and methods of analysis in which the mode of inheritance needs to be specified, the search for genes involved in complex diseases has given rise to different designs. These designs focus mainly on concordant or discordant sib-pairs or parent-child trios and use the assumption that the more similar the phenotype, the more genes are shared in common, with no need to specify a mode of inheritance [63]. Meta-analytical methods for pooling results from these model-free, family-based methods have been suggested [64,65], but no similar methods have been proposed for population-based molecular association studies. In a submitted manuscript, we propose such a method that pools data across studies without assuming any mode of inheritance and avoids the problem of multiple comparisons.

In conclusion, performing meta-analysis of population-based molecular association studies is relatively new (the first such study we identified was published in 1991). It seems to be an "orphan" field, falling somewhere between genetics and epidemiology, undertaken by geneticists and epidemiologists in roughly even numbers but benefiting from the experience of neither. Our review found that basic lessons and guidelines from meta-analyses in traditional epidemiology were not carried over to meta-analyses of population-based molecular association studies and that there was little regard for genetic considerations. Methodologic advances in this area require true cooperation, communication, and a trans-disciplinary approach from geneticists and epidemiologists.

Acknowledgments

We thank Patrick McElduff for helpful discussions and feedback.

References

[1] Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. JAMA 2000;283:2008–11.

[2] Sharma P. Meta-analysis of the ACE gene in ischemic stroke. J Neurol Neurosurg Psychiatry 1998;64:227–30.

[3] Breen G, Brown J, Maude S, et al. –141 C Del/Ins polymorphism of the dopamine receptor 2 gene is associated with schizophrenia in a British population. Am J Med Genet 1999;88:407–10.

[4] Rantala M, Rantala TT, Savolainen MJ, et al. Apolipoprotein B gene polymorphism and serum lipids: meta-analysis of role of genetic variation in responsiveness to diet. Am J Clin Nutr 2000:713–24.

[5] Allison DB, Heo M, Faith MS, et al. Meta-analysis of association of the Trp64Arg polymorphism in the beta 3 adrenergic receptor with body mass index. Int J Obes 1998;22:559–66.

[6] Strachan T, Read A. Human molecular genetics. 2nd edition. Oxford: Bios Scientific Publishers; 1999.

[7] Carlquist JF, Menlove RL, Murray MB, et al. HLA class II (DR and DQ) antigen associations in idiopathic dilated cardiomyopathy: validation study and meta-analysis of published HLA association studies. Circulation 1991;83:515–22.

[8] Jenkins D, Penny MA, Fletcher JA, et al. HLA class II gene polymorphism contributes little to Hashimoto's thyroiditis. Clin Endocrinol 1992;37:141–5.

[9] Pato CN, Macciardi F, Pato MT, et al. Review of putative association of dopamine D2 receptor and alcoholism: a meta analysis. Am J Med Genet 1993;48:78–82.

[10] Ordovas JM, Lopez-Miranda J, Perez-Jimenez F, et al. Effect of apolipoprotein E and A-IV phenotypes on the low density lipoprotein response to HMG CoA reductase inhibitor therapy. Atherosclerosis 1995;113:157–66.

[11] McWilliams JE, Sanderson BJS, Harris EL, et al. Glutathione S-Transferase M1 (GSTM1) deficiency and lung cancer risk. Cancer Epidemiol Biomark Prev 1995;4:589–94.

[12] Shaikh S, Collier DA, Sham PC, et al. Allelic association between a Ser-9-Gly polymorphism in the dopamine D3 receptor gene and schizophrenia. Hum Genet 1996;97:714–9.

[13] Kunz R, Kreutz R, Beige J, et al. Association between the angiotensinogen 235T-variant and essential hypertension in Whites: a systematic review and methodological appraisal. Hypertension 1997;30: 1331–7.

[14] Weston A, Godbold JH. Polymorphisms of H-ras-1 and p53 in breast cancer and lung cancer: a meta-analysis. Environ Health Perspect 1997;105(Suppl 4):919–26.

[15] Staessen JA, Wang JG, Ginocchio G, et al. The deletion/insertion polymorphism converting enzyme gene and cardiovascular-renal risk. J Hypertens 1997;15:1579–92.

[16] Christensen M, Gotzche PC, Brosen K. The sparteine/debrisoquine (CYP2D6) oxidation polymorphism and the risk of lung cancer: a meta-analysis. J Clin Pharmacol 1997;51:389–93.

[17] Hokanson JE. Lipoprotein lipase gene variants and risk of coronary disease: a quantitative analysis of population-based studies. Int J Clin Lab Res 1997;27:24–34.
[18] Prasher VP, Chowdhury TA, Rowe BR, et al. ApoE genetype and Alzheimer's disease in adults with Down Syndrome: meta-analysis. Am J Ment Retard 1997;102:103–10.
[19] Arranz MJ, Munro J, Sham P, et al. Meta-analysis of studies on genetic variation in 5-HT_{2A} receptors and clozapine response. Schizophr Res 1998;32:93–9.
[20] Furlong RA, Ho L, Walsh C, et al. Analysis and meta-analysis of two serotonin transporter gene polymorphisms in bipolar and unipolar affective disorders. Am J Med Genet 1998;81:58–63.
[21] Williams J, Spurlock G, Holmans P, et al. A meta-analysis and transmission disequilibrium study of association between the dopamine D3 receptor gene and schizophrenia. Mol Psychiatry 1998;3:141–9.
[22] Rostami-Hodjegan A, Lennard MS, Woods HF, et al. Meta-analysis of studies of the CYP2D6 polymorphism in relation to lung cancer and Parkinson's disease. Pharmacogenetics 1998;8:227–38.
[23] Fujisawa T, Ikegami H, Kawaguchi Y, et al. Meta-analysis of the association of Trp64 Arg polymorphism of Beta3-adrenergic receptor gene with body mass index. J Clin Endocrinol Metab 1998;83:2441–4.
[24] Fujisawa T, Ikegami H, Kawaguchi Y, et al. Meta-analysis of association of insertion/deletion polymorphism of angiotensin I-converting enzyme gene with diabetic nephropathy and retinopathy. Diabetologia 1998;41:47–53.
[25] Hani EH, Boutin P, Durand E, et al. Missense mutations in the pancreatic islet beta cell inwardly rectifying K^+ channel gene (KIR6.2/BIR): a meta-analysis suggests a role in polygenic basis of Type II diabetes mellitus in Caucasians. Diabetologia 1998;41:1511–5.
[26] Joost O, Taylor CA, Thomas CA, et al. Absence of effect of seven functional mutations in the CYP2D6 gene in Parkinson's disease. Mov Disord 1999;14:590–5.
[27] Staessen JA, Kuznetsova T, Wang JG, et al. M235T angiotensinogen gene polymorphism and cardiovascular renal risk. J Hypertens 1999;17:9–17.
[28] Iacoviello L, Burzotta F, Castelnuovo AD, et al. The 4G/5G polymorphism of PAI-1 promoter gene and the risk of myocardial infarction: a meta-analysis. Thromb Haemost 1998;80:1029–30.
[29] Dubertret C, Gorwood P, Ades J, et al. Meta-analysis of DRD3 gene and schizophrenia: ethnic heterogeneity and significant association in Caucasians. Am J Med Genet 1998;81:318–22.
[30] Furlong RA, Ho L, Rubinsztein JS, et al. Analysis of monoamine oxidase A (MAOA) gene in bipolar affective disorder by association studies, meta-analysis, and sequencing of the promoter. Am J Med Genet 1999;88:398–406.
[31] Furlong RA, Rubinsztein JS, Ho L, et al. Analysis and metaanalysis of two polymorphisms within the tyrosine hydroxylase gene in bipolar and unipolar affective disorders. Am J Med Genet 1999;88:88–94.
[32] Juo S-HH, Wyszynski DF, Beaty TH, et al. Mild association between the A/G polymorphism in the promoter of the Apolipoprotein A-I gene and apolipoprotein A-I levels: a meta-analysis. Am J Med Genet 1999;999:235–41.
[33] McCarron MO, Delong D, Alberts MJ. APOE genotype as a risk factor for ischemic cerebrovascular disease: a meta-analysis. Neurology 1999;53:1308–11.
[34] Houlston RS. Glutathione S-transferase M1 status and lung cancer risk: a meta-analysis. Cancer Epidemiol Biomarkers Prev 1999;8:675–82.
[35] Houlston RS. CYP1A1 polymorphisms and lung cancer risk: a meta-analysis. Pharmacogenetics 2000;10:105–14.
[36] Preisig M, Bellivier F, Fenton BT, et al. Association between bipolar disorder and monoamine oxidase A gene polymorphisms: results of multicenter study. Am J Psychiatry 2000;157:948–55.
[37] Wong NA, Rae F, Simpson KJ, et al. Genetic polymorphisms of cytochrome p4502E1 and susceptibility to alcoholic liver disease and hepatocellular carcinoma in a white population: a study and literature review, including meta-analysis. Mol Pathol 2000;53:88–93.
[38] Kato N, Sugiyama T, Morita H, et al. Angiotensinogen gene and meta-analysis on six reported studies. J Hypertens 1999;17:757–63.
[39] Keavney B, McKenzie C, Parish S, et al. Large-scale test of hypothesised associations between the angiotensin-converting-enzyme insertion/deletion polymorphism and myocardial infarction in about 5000 cases and 6000 controls. Lancet 2000;355:434–42.
[40] Suarez-Almazor ME, Belseck E, Homik J, et al. Identifying clinical trials in the medical literature with electronic databases: MEDLINE alone is not enough. Control Clin Trials 2000;21:476–87.
[41] Petitti DB. Meta-analysis, decision analysis and cost-effectiveness analysis: methods for quantitative synthesis in medicine. New York: Oxford University Press; 1994.
[42] Normand S-LT. Tutorial in biostatistics meta-analysis: formulating, evaluating, combining, and reporting. Statist Med 1999;18:321–59.
[43] Fleiss JL. Statistical methods for rates and proportions. 2nd edition. New York: Wiley; 1981.
[44] Thompson SG, Pocock SJ. Can meta-analyses be trusted? Lancet 1991;338:1127–30.
[45] Fleiss JL, Gross AJ. Meta-analysis in epidemiology, with special reference to studies of the association between exposure to environmental tobacco smoke and lung cancer: a critique. J Clin Epidemiol 1991;44:127–39.
[46] Aitkin M. Meta-analysis by random effects modelling in generalised linear models. Statist Med 1999;18:2343–51.
[47] Olkin I. Diagnostic statistical procedures in medical meta-analysis. Statist Med 1999;18:2331–41.
[48] How to review the evidence: systematic identification and review of the scientific literature. Canberra: NHMRC; 2000.
[49] Egger M, Smith GD, Altman DG. Systematic reviews in health care: meta-analysis in context. 2nd edition. London: BMJ Books; 2001.
[50] Thompson SG. Why sources of heterogeneity in meta-analysis should be investigated. BMJ 1994;309:1351–5.
[51] Thompson SG, Sharp SJ. Explaining heterogeneity in meta-analysis: a comparison of methods. Statist Med 1999;18:2693–708.
[52] Greenland S. Quantitative methods in the review of epidemiologic literature. Epidemiol Rev 1987;9:1–30.
[53] Hardy RJ, Thompson SG. Detecting and describing heterogeneity in meta-analysis. Statist Med 1998;17:841–56.
[54] Steichen TJ, Egger M, Sterne J. Test for publication bias in meta-analysis. Stata Bull 1998;44:3–4.
[55] Begg CB, Mazumdar M. Operating characteristics of rank correlation test for publication bias. Biometrics 1994;50:1088–101.
[56] Egger M, Smith GD, Schneider M, et al. Bias in meta-analysis detected by simple, graphical test. BMJ 1997;315:629–34.
[57] Guo SW, Thompson EA. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 1992;48:361–72.
[58] Schaid DJ, Jacobsen SJ. Biased tests of association: comparison of allele frequencies when departing from the Hardy-Weinberg proportions. Am J Epidemiol 1999;149:706–11.
[59] Juni P, Witschi A, Bloch R, et al. The hazards of scoring the quality of clinical trials for meta-analysis. JAMA 1999;282:1054–60.
[60] Poole C, Greenland S. Random effects meta-analyses are not always conservative. Am J Epidemiol 1999;150:469–75.
[61] Wacholder S, Rothman N, Caporaso N. Counterpoint: bias from population stratification is not a major threat to the validity of conclusions from epidemiological studies of common polymorphisms and cancer. Cancer Epidemiol Biomarkers Prev 2002;11:513–20.
[62] Thomas DC, Witte JS. Point: population stratification: a problem for case-control studies of candidate-gene association? Cancer Epidemiol Biomarkers Prev 2002;11:505–12.
[63] Haines JL, Pericak-Vance MA. Approaches to gene mapping in complex human diseases. New York: Wiley-Liss; 1998.
[64] Gu C, Province MA, Rao DC. Meta-analysis for model-free methods. Adv Genet 2001;42:255–72.
[65] Etzel CJ, Guerra T. Meta-analysis of genetic-linkage analysis of quantitative-trait loci. Am J Hum Genet 2002;71:56–65.