Review Article
Diagnostic Genetics

Ann Lab Med 2015;35:283-287
http://dx.doi.org/10.3343/alm.2015.35.3.283
ISSN 2234-3806 • eISSN 2234-3814

ANNALS OF LABORATORY MEDICINE

Meta-Analysis of Genetic Association Studies

Young Ho Lee, M.D.

Division of Rheumatology, Department of Internal Medicine, Korea University College of Medicine, Seoul, Korea

The object of this review is to help readers to understand meta-analysis of genetic association study. Genetic association studies are a powerful approach to identify susceptibility genes for common diseases. However, the results of these studies are not consistently reproducible. In order to overcome the limitations of individual studies, larger sample sizes or meta-analysis is required. Meta-analysis is a statistical tool for combining results of different studies on the same topic, thus increasing statistical strength and precision. Meta-analysis of genetic association studies combines the results from independent studies, explores the sources of heterogeneity, and identifies subgroups associated with the factor of interest. Meta-analysis of genetic association studies is an effective tool for garnering a greater understanding of complex diseases and potentially provides new insights into gene–disease associations.

Key Words: Gene, Polymorphism, Association study, Meta-analysis

Received: July 15, 2014
Revision received: November 2, 2014
Accepted: March 4, 2015

Corresponding author: Young Ho Lee
Division of Rheumatology, Department of Internal Medicine, Korea University Anam Hospital, Korea University College of Medicine, 73 Inchon-ro, Seongbuk-gu, Seoul 136-705, Korea
Tel: +82-2-920-5645
Fax: +82-2-922-5974
E-mail: lyhcgh@korea.ac.kr

INTRODUCTION

Genetic association studies evaluate the association between a disease and susceptible genetic variants in order to identify genetic variants that affect susceptibility to common diseases [1]. Genetic association studies are a powerful approach for identifying susceptibility genes underlying common diseases and provide linkage analyses when investigating complex diseases. The number of genetic association studies is increasing exponentially; however, the results are not consistently reproducible. Genetic association studies of candidate genes have shown that the majority of initial positive associations cannot be reproduced in subsequent studies [2, 3]. These findings suggest that a large number of original findings are false-positive reports (type I errors), or that small genetic effects were undetectable (false-negative, type II errors) in the majority of these studies. However, it is possible that the lack of reproducibility may be due to true variability in the association of different populations [2]. Candidate genes occasionally demonstrate only small effects in genetic association tests (mean odds ratio [OR] of 1.33 in 55 meta-analyses); therefore, genetic association studies require thousands of subjects to have a reasonable chance of discovering an effect [2]. Thus, to detect a small or moderate genetic effect of polymorphisms, large sample sizes or meta-analysis is required [4].

Meta-analysis is a statistical tool for combining results from different studies on the same topic and is becoming a popular method for resolving discrepancies in genetic association studies. Meta-analysis is an objective, quantitative synthesis of research findings, enabling the identification of genuine associations [5]. Meta-analysis increases statistical strength and precision in estimating effects by combining the results of previous studies, thus overcoming the problem of small sample size and the inadequate statistical strength of complex trait genetic studies [6]. Testing for and quantifying inter-study heterogeneity is an important step of meta-analysis [5]. Meta-analysis analyzes the discrepancies in the results of independent studies by addressing inter-study heterogeneity, thus potentially providing new insights into gene–disease associations. Meta-analysis combines the results from independent studies, explores the sources of

heterogeneity, and identifies subgroups associated with the factor of interest. When properly performed, meta-analysis of genetic association studies is considered as a decisive evidence [7], and is an effective tool for gaining a greater understanding of complex diseases. The object of this review is to help readers to understand meta-analysis of genetic association study. We deal with meta-analysis process of genetic association studies.

META-ANALYSIS METHOD

The general process by which a meta-analysis of genetic association studies is performed is outlined in Table 1. Following these steps will facilitate performing and understanding meta-analyses.

1. Genetic models

In a genetic association study on a polymorphism with two alleles (A and a), where one of the alleles may be associated with a disease (let A be the susceptibility allele), subjects are classified into three genotypes: AA, Aa, and aa. Association studies are the following genetic models [8]: (1) the allele contrast model, in which the numbers of allele A are compared with that of allele a (A vs. a); (2) the recessive model, in which the AA genotype is compared with the combined Aa+aa genotype (AA vs. Aa+aa); (3) the dominant model, in which the combined AA+Aa genotype is compared with the aa genotype (AA+Aa vs. aa); (4) the homozygote contrast (additive) model, in which the AA genotype is compared with the aa genotype (AA vs. aa); and (5) the co-dominant model, in which the combined AA+Aa genotype is compared with the aa genotype (AA+aa vs. Aa). Genetic association studies do not adopt a specific model, and thus multiple genetic models need to be examined [9].

2. Heterogeneity

Meta-analysis analyzes the variance in the results of independent studies and examines the existence of heterogeneity in primary studies. The heterogeneity test examines the following null hypothesis: no differences exist between the findings of primary studies. Cochran's Q test is used to determine whether differences between primary studies exist or whether variations are due to chance [10]. Cochran's Q value is calculated by summing the squared deviations of each study's estimate from the overall estimate, and then comparing it with the chi-squared distribution with κ-1 degrees of freedom (df) (where κ is the number of studies) [10]. A heterogeneity P value <0.10 (not 0.05) indicates the presence of heterogeneity because of the low statistical strength of Cochran's Q test [11]. However, the Q test can be unreliable when a small number of studies are included in the meta-analysis.

Another commonly used method for testing heterogeneity is the I^2 value, which quantifies the effect of heterogeneity and does not depend on the number of studies or the type of outcome data. I^2 values range between 0% and 100% and represent the proportion of inter-study variability that can be attributed to heterogeneity rather than chance ($I^2 = 100\% \times [Q - df]/Q$) [12]. I^2 values of 25%, 50%, and 75% are assigned as low, moderate, and high estimates, respectively [9, 12]. We evaluate Cochran's Q and I^2 values by using meta-analysis software such as the Review Manager (Cochrane Collaboration, London, UK) or Comprehensive Meta-Analysis program (Biosta, Englewood, NJ, USA) by inputting the number of subjects and genotype data in the cases and controls.

Table 1. Steps in performing a meta-analysis for genetic association studies

1.	Check study quality: Hardy-Weinberg test
2.	Check inter-study heterogeneity: Cochran Q test, I^2
3.	Meta-analysis: Fixed or random effects model, Forrest plot
4.	Perform subgroup analysis: Ethnicity-specific analysis
5.	Check publication bias: Funnel plot, Egger's regression test
6.	Present meta-analysis result

3. Fixed vs. random effects models

Meta-analysis combines the effect sizes of the included studies by weighting the data according to the amount of information in each study. The weights are calculated by using the sample size and variability within each study. The fixed effects model assumes that genetic factors have similar effects on disease susceptibility in all of the studies, and that the observed variations between studies are caused by chance alone [13]. The random effects model assumes that different studies exhibit substantial diversity and assesses both intra-study sampling errors and inter-study variances [14]. The choice of meta-analysis model depends on the presence or absence of heterogeneity. In the absence of heterogeneity, a fixed effects model is used for meta-analysis. When a significant Q value (P<0.10) is calculated, indicating the existence of heterogeneity in studies, a random effects model is used for meta-analysis [15]. When study groups are homogeneous, both models would offer similar results; however, in case of heterogeneity, the random effects model usually provides wider confidence intervals (CIs) than the fixed effects model [16]. We perform meta-analysis of fixed or random effects

http://dx.doi.org/10.3343/alm.2015.35.3.283

models by using the Review Manager or Comprehensive Meta-Analysis program after inputting the number of subjects and genotype data into the cases and controls.

4. Evaluating causes of heterogeneity

It is important to assess whether heterogeneity exists in the studies included in meta-analysis as well as to determine possible causes of heterogeneity, because heterogeneity can lead to bias in meta-analysis results—referred to as "mixing apples and oranges" [17]. Subgroup analysis may also be used to assess the impact of heterogeneity. In subgroup analysis, meta-analysis is performed on the basis of factors such as ethnicity, number of studies, the Hardy-Weinberg equilibrium (HWE), or clinical features, to assess the impact of a potential source of heterogeneity. The HWE refers to a situation, in which the frequencies of genotypes are predicted on the basis of the frequencies of two alleles according to the simple Mendelian inheritance model [18]. Departures from HWE can arise from genotyping errors, population stratification, and selection bias in the recruitment of controls [19]. Whether the genotype frequencies of the controls are in HWE must be determined, because genotyping errors are a significant cause of deviation from HWE. The distribution of genotypes in the control group is tested for deviation from HWE by performing an exact test. Subgroup analysis is recommended, except for studies, in which the controls are not in HWE [11]. We can check HWE of genotypes in control groups by analyzing the genotype data by using an HWE web tool (http://ihg.gsf.de/cgi-bin/hw/hwa1.pl).

5. Publication bias

Studies reporting positive effects tend to have a higher likelihood of being published than those that do not, and studies showing no significant result tend to remain unpublished [20]. As meta-analysis includes only published studies, the degree of the actual effect might be overestimated [20]. This outcome is termed "publication bias." Meta-analysis of genetic association studies may be subject to publication bias. The funnel plot is a commonly used graphic test to assess publication bias in meta-analyses [21]. This test is a scatter plot of the effect estimate from each study included in the meta-analysis against the measure of its precision (1/SE) [22]. The effect estimates of small-scale studies will scatter more widely at the bottom of the graph, while the spread among larger-scale studies will be narrower. In the absence of publication bias, the funnel plot resembles a symmetrical inverted funnel; asymmetry suggests the existence of publication bias [23]. However, funnel plots usually require a range of studies of varying scales and subjective judgments. Therefore, other methods such as the Egger's linear regression test [21], which measures funnel plot asymmetry on a natural logarithm scale of ORs, is used. If the existence of publication bias is suggested by the funnel plot or other statistical tests, the results of the meta-analysis should be interpreted cautiously, and the possible impact of publication bias should be noted. Funnel plots and Egger's regression tests can be performed by using the Review Manager or Comprehensive Meta-Analysis program.

EXAMPLES OF META-ANALYSES OF GENETIC ASSOCIATION TESTS

We selected studies for the meta-analysis of association between the Fc receptor like-3 (*FCRL3*)-169C/T polymorphism and rheumatoid arthritis (RA) based on inclusion and exclusion criteria [24]. We included studies if: (1) they were case control studies, (2) the data was original (independence among studies), (3) they provided enough data to calculate ORs, and (4) the distribution of the *FCRL3*-169C/T polymorphism in normal controls was in HWE. We excluded the followings: (1) studies that contained overlapping data, (2) studies, in which the number of null and wild genotypes could not be ascertained, and (3) those, in which family members were studied, as these analyses are based on linkage considerations. Meta-analyses were performed by using: 1) allelic contrast, 2) recessive, 3) dominant, 4) homozygote, and 5) heterozygote contrast models to determine the association between the *FCRL3*-169C/T polymorphism and RA [24].

Step 1: We determined if genotypes in the control group deviated from HWE by using the web tool found at http://ihg.gsf.de/cgi-bin/hw/hwa1.pl. No deviation from HWE was found in any of the studies included in the meta-analysis. Step 2: We calculated Cochran's Q and I^2 values by using Comprehensive Meta-Analysis program. Cochran's Q *P* value and I^2 values were calculated as 0.084 and 34.1%, respectively, indicating existence of inter-study heterogeneity [24]. Step 3: We selected one model between the fixed and random effects models on the basis of heterogeneity and performed statistical analyses of the selected model by using Comprehensive Meta-Analysis program. We selected the random effect model because of the significant inter-study heterogeneity (Cochran's Q *P* value=0.084, I^2 value= 34.1%) [24]. By performing a random effects meta-analysis, no association was found between RA and the *FCRL3*-169C allele in the study subjects (OR=1.046, 95% CI=0.997-1.098, *P*=0.068) [24]. Step 4: We performed ethnicity-specific meta-analyses on the European and Asian populations of the 17 studies included

in the original analysis (9 European, 7 Asian, and 1 Native North American populations). The meta-analysis stratified by ethnicity revealed a significant association between the *FCRL3*-169C allele and RA in the Asian populations under allelic contrast, recessive, dominant, and homozygote contrast models (C allele OR=1.101, 95% CI=1.035-1.174, P=0.002), but not in the European populations (OR=1.012, 95% CI= 0.962-1.065, P=0.643) [24]. Step 5: We generated a funnel plot and performed Egger's regression test by using Comprehensive Meta-Analysis program. The funnel plot showed no evidence of asymmetry and Egger's regression test did not yield a significant P value (Egger's regression test P value=0.863), indicating that publication bias did not affect the meta-analysis [24]. Step 6: We presented the meta-analysis results showing that *FCRL3*-169C/T polymorphism may confer susceptibility to seropositive RA in Asian populations [24].

DISCUSSION

Meta-analysis is a statistical method for combining the results of independent studies, providing a precise estimate of the effect size, and increasing statistical strength, which is especially important when the strength of primary study is limited because of a small sample size. Meta-analysis solves a problem associated with genetic association studies by analyzing variation in the results of different studies by identifying inter-study heterogeneity. Thus, meta-analysis is potentially a powerful tool for assessing the effects of candidate genes [25-27].

Meta-analysis is also applicable to a variety of genetics study designs from family-based linkage studies to genome-wide scans and genome-wide association studies, as well as to population-based association studies [5]. Meta-analysis has been used increasingly for combining and integrating data from a number of independent studies. However, meta-analysis is not a suitable replacement for robust genetic association studies. A major criticism of meta-analysis is that it combines different types of studies, mixing apples and oranges. However, meta-analysis can overcome this problem by assessing heterogeneity in studies, and performing subgroup analysis [28]. The "garbage in, garbage out" metaphor refers to the fact that if a meta-analysis includes low quality studies with bias, the results of the meta-analysis will be biased. The results of meta-analysis depend on the quality of primary research. Therefore, meta-analysis should include studies selected on the basis of inclusion criteria. Despite these limitations, a properly executed meta-analysis is an invaluable link between past and future studies that objectively and quantitatively synthesizes evidence while minimizing bias.

In conclusion, meta-analysis of genetic association studies is an effective tool for garnering a greater understanding of complex diseases and potentially provides new insights into gene–disease associations.

Author's Disclosure of Potential Conflicts of Interest

No potential conflicts of interest relevant to this article were reported.

REFERENCES

1. Cardon LR and Bell JI. Association study designs for complex diseases. Nat Rev Genet 2001;2:91-9.
2. Ioannidis JP. Genetic associations: false or true? Trends Mol Med 2003;9:135-8.
3. Colhoun HM, McKeigue PM, Davey Smith G. Problems of reporting genetic associations with complex outcomes. Lancet 2003;361:865-72.
4. Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. Nat Genet 2003;33:177-82.
5. Trikalinos TA, Salanti G, Zintzaras E, Ioannidis JP. Meta-analysis methods. Adv Genet 2008;60:311-34.
6. Gotzsche PC. Why we need a broad perspective on meta-analysis. It may be crucially important for patients. BMJ 2000;321:585-6.
7. Yuan Y and Hunt RH. Systematic reviews: the good, the bad, and the ugly. Am J Gastroenterol 2009;104:1086-92.
8. Minelli C, Thompson JR, Abrams KR, Thakkinstian A, Attia J. The choice of a genetic model in the meta-analysis of molecular association studies. Int J Epidemiol 2005;34:1319-28.
9. Thakkinstian A, McElduff P, D'Este C, Duffy D, Attia J. A method for meta-analysis of molecular association studies. Stat Med 2005;24:1291-306.
10. Whitehead A and Whitehead J. A general parametric approach to the meta-analysis of randomized clinical trials. Stat Med 1991;10:1665-77.
11. Munafò MR and Flint J. Meta-analysis of genetic association studies. Trends Genet 2004;20:439-44.
12. Higgins JP and Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21:1539-58.
13. Davey Smith G and Egger M. Meta-analyses of randomised controlled trials. Lancet 1997;350:1182.
14. DerSimonian R and Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7:177-88.
15. Ried K. Interpreting and understanding meta-analysis graphs--a practical guide. Aust Fam Physician 2006;35:635-8.
16. Zintzaras E and Lau J. Synthesis of genetic association studies for pertinent gene-disease associations requires appropriate methodological and statistical approaches. J Clin Epidemiol 2008;61:634-45.
17. Salanti G, Sanderson S, Higgins JP. Obstacles and opportunities in meta-analysis of genetic association studies. Genet Med 2005;7:13-20.
18. Wittke-Thompson JK, Pluzhnikov A, Cox NJ. Rational inferences about

http://dx.doi.org/10.3343/alm.2015.35.3.283

departures from Hardy-Weinberg equilibrium. Am J Hum Genet 2005; 76:967-86.
19. 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;13:840-8.
20. Dickersin K and Min YI. Publication bias: the problem that won't go away. Ann N Y Acad Sci 1993;703:135-46; discussion 146-8.
21. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315:629-34.
22. Sterne JA and Egger M. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. J Clin Epidemiol 2001;54:1046-55.
23. Lau J, Ioannidis JP, Terrin N, Schmid CH, Olkin I. The case of the misleading funnel plot. BMJ 2006;333:597-600.
24. Song GG, Bae SC, Kim JH, Kim YH, Choi SJ, Ji JD, et al. Association between functional Fc receptor-like 3 (FCRL3) -169 C/T polymorphism and susceptibility to seropositive rheumatoid arthritis in Asians: a meta-analysis. Hum Immunol 2013;74:1206-13.
25. Nath SK, Harley JB, Lee YH. Polymorphisms of complement receptor 1 and interleukin-10 genes and systemic lupus erythematosus: a meta-analysis. Hum Genet 2005;118:225-34.
26. Lee YH, Harley JB, Nath SK. Meta-analysis of TNF-alpha promoter -308 A/G polymorphism and SLE susceptibility. Eur J Hum Genet 2006;14: 364-71.
27. Lee YH, Rho YH, Choi SJ, Ji JD, Song GG. PADI4 polymorphisms and rheumatoid arthritis susceptibility: a meta-analysis. Rheumatol Int 2007;27:827-33.
28. Bailar JC 3rd. The promise and problems of meta-analysis. N Engl J Med 1997;337:559-61.