Genetic Meta-Analysis

Josée Dupuis
Professor and Chair,
Epidemiology, Biostatistics and Occupational Health

Paul Hanson
Computer Science and Biology

1

1

Outline

- Meta-Analysis Set-up
- P-value Methods
- Effect Size Methods
- Issues specific to meta-analysis of genetic studies
- Using METAL & GWAMA software to perform meta-analyses of genome-wide association scans

2

2

Meta-analysis set-up

- Have a collection of k primary studies (i = 1 to k)
  - Same null hypothesis, H_0: β=0
  - Each study i has estimate $\hat{\beta}_i$ and a p-value p_i for this hypothesis
- Want to test H_0 using all k studies at the α level of significance
- Want to pooled β estimate with a confidence interval
- Key assumption – independence of study results
  - There should be no subjects who are included in more than one study result

3

3

Meta-analysis set-up: Genetic study

- SNP of interest has two alleles, A and G
- Want to estimate the effect of the G allele on a particular phenotype
  - Fasting glucose (β coefficient from linear regression)
  - Diabetes risk (odds ratio from logistic model or β coefficient from logistic regression)
- Use additive regression model to get an estimate of β or odds ratio in each study
- Use meta-analysis approaches to get a pooled estimate of effect (with confidence interval) using estimates from all studies

4

4

Meta-analysis set-up

- First step in a meta-analysis is to identify a consistent measure from each study in the collection to be summarized
  - Often the OR and its SE or regression coefficient and its SE
  - Could be the p-value from a test of association
- Next step is to use these study estimates to arrive at a summary value

5

5

Combining P-values

- Suppose the k studies all have a p-value for association of genotype with trait of interest
- Under the null hypothesis H_0
  - The p-values p_i follow a uniform distribution from 0 to 1
- This fact underlies the p-value methods
  - They all look to see if the set of p_i are consistent with k random observations from a uniform
  - If not, this indicates H_0 may be incorrect

6

6

Fisher's method

- Approach to combining p-values due to Fisher
- It is based on the connection between the uniform and chi-square distributions
  - If variable V follows a uniform distribution then $-2\times\ln(V)$ follows a chi-square distribution with 2 degrees of freedom
  - ln is the natural logarithm

7

7

Fisher's method

- The test statistic is the sum of the $-2 \times \ln(\text{p-value})$ from all the studies
- Reject H_0 if $$\sum_{i=1}^{k} -2\times\ln(p_i) \geq C_{1-\alpha,2k}$$

  $C_{1-\alpha,2k}$ is the upper tail α critical value for a chi-square with 2k degrees of freedom
- This method is not widely used in genetic meta-analysis because it does not take the direction of effect into consideration

8

8

Weighted Z-score approach

- Weighted (signed) Z-score approach takes direction of association and sample size into consideration
- For each study with sample size n_i, the p-value for association is converted to a standard normal deviate with the sign of the normal deviate corresponding to the direction of the association
  - positive = effect allele increases risk or average trait value
    - $Z_i = +\Phi^{-1}(p_i)$, Φ is the inverse normal distribution
  - negative = effect allele decreases risk or average trait value
    - $Z_i = -\Phi^{-1}(p_i)$

$$Z_{meta} = \frac{\sum \sqrt{n_i} Z_i}{\sqrt{\sum n_i}}$$

9

9

Weighted Z-score approach

- Z_{meta} is compared to standard normal distribution to determine significance
- No pooled effect size estimate available when using this approach
- Weighted Z-score approach is often used to meta-analyze studies with different study designs or approaches
  - Effect estimate from each study may not be on the same scale or comparable
    - Different assay used to measure trait
    - Different technologies used in each study
    - Different set of covariates used in model
    - Different transformation applied to trait

10

10

Combining effect sizes

- In medicine, using effect sizes, rather than p-values, has been the main form of meta-analysis
- Combining effect sizes
  - Fixed effects methods
  - Random effects methods
- For case-control studies, the odds ratio (OR) is the usual effect size
- For continuous traits, β estimates from regression is the usual effect size
  - Anything with a confidence interval can be used as an effect size

11

11

Fixed effects

- Combine effects from primary studies
  - Sampling error gives *within* study error (due to sample size)
  - Is there also variability in effect *between* studies?
    - If not, fixed effects analysis
    - If so, random effects analysis
  - Both analyses compute weighted averages of the study effects, but the weights are different

12

12

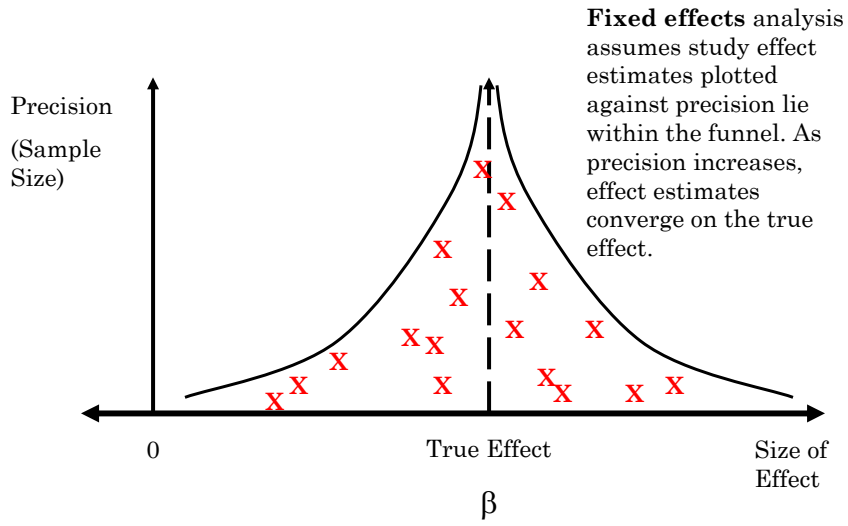
Fixed effects

- Fixed effects analysis *assumes* that all studies are estimating the same underlying population effect parameter β
  - ➢ The expected value for each of the effect sizes in the set of primary studies is β
  - ➢ The only reason studies have different effect estimates is random sampling error
    - ♦ If any primary study had an infinitely large sample size, then the effect size estimate would exactly equal β

13

13

Fixed effects analyses



Fixed effects analysis assumes study effect estimates plotted against precision lie within the funnel. As precision increases, effect estimates converge on the true effect.

14

14

Inverse variance weighting

- Each study i provides an estimate $\hat{\beta}_i$ of the population value β with a standard error estimate $SE(\hat{\beta}_i)$
- For the inverse variance weighting method, we assume that the $\hat{\beta}_i$'s roughly follow a normal distribution
- Since all studies estimate β, it makes sense to estimate β by averaging the study estimates

$$\hat{\beta} = w_1\hat{\beta}_1 + w_2\hat{\beta}_2 + \cdots + w_k\hat{\beta}_k$$

- w_i is the weight for study i
  - $0 \leq w_i$ for each i
  - $w_1 + \ldots + w_k = 1$

15

15

Inverse variance weighting

- Minimizes the standard error of $\hat{\beta}$
  - Primary studies should be weighted in proportion to their precision (reciprocal of SE squared)

$$w_i \propto \frac{1}{SE(\hat{\beta}_i)^2}$$

  - More precise studies get more weight
  - Less precise studies get less weight

16

16

Inverse variance weighting

- The best weight to use is

$$w_i = \frac{\frac{1}{SE\left(\hat{\beta}_i\right)^2}}{\frac{1}{SE\left(\hat{\beta}_1\right)^2} + \frac{1}{SE\left(\hat{\beta}_2\right)^2} + \cdots + \frac{1}{SE\left(\hat{\beta}_k\right)^2}}$$

  - The bottom of the fraction is the same for each study
  - Ensures that the weights sum to 1

17

17

Inverse variance weighting

- The standard error of the combined estimate

$$SE\left(\hat{\beta}\right) = \sqrt{\frac{1}{\sum_{i=1}^{k} \frac{1}{SE\left(\hat{\beta}_i\right)^2}}}$$

18

18

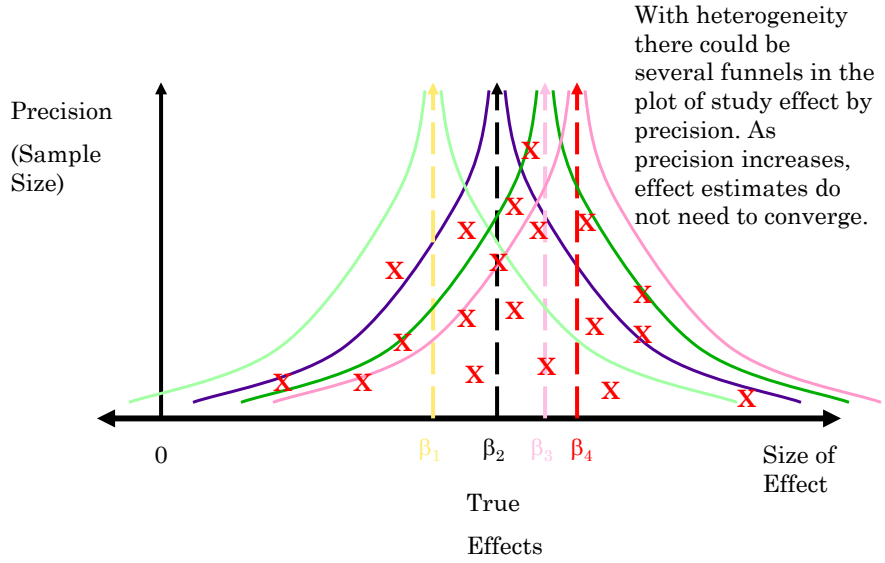
Heterogeneity

- With heterogeneity there is no single population effect size β
  - Each study effect size $\hat{\beta}_i$ is estimates a study-specific population effect β_i
  - The effect size for study i approaches β_i as the sample size increases
- It is not the case that effect sizes for all of the primary studies would approach the same value as their sample sizes increase
- Some of the β_i might be the same, but not all of them

19

19

Heterogeneity



20

20

Testing for heterogeneity

- Testing for heterogeneity is somewhat problematic due to low power, but
- The test can still be useful to confirm the presence of heterogeneity
  - ➢ It may not be significant every time there is heterogeneity
  - ➢ It tends be significant too often if there are many studies (about 100 or more).
  - ➢ But, when it is significant with a small number of studies it is a strong indication of heterogeneity

21

21

Testing for heterogeneity

- The test is called Cochran's Q-test or just the Q-test
- The form of the Q-test statistic is similar to the computation of variance
  - ➢ Differences between the study effect size and the fixed effects combined effect size are squared
  - ➢ The squared differences are weighted by the inverse of the effect size standard error squared and summed over all the studies

22

22

Testing for heterogeneity

- The formula for the Q-test statistic is

$$Q = \frac{1}{SE(\hat{\beta}_1)^2} \times (\hat{\beta}_1 - \hat{\beta})^2 + \frac{1}{SE(\hat{\beta}_2)^2} \times (\hat{\beta}_2 - \hat{\beta})^2 + \cdots + \frac{1}{SE(\hat{\beta}_k)^2} \times (\hat{\beta}_k - \hat{\beta})^2$$

- $\hat{\beta}$ is the fixed effect combined effect size
- $\hat{\beta}_i$ is the effect size for primary study i
- $SE(\hat{\beta}_i)$ is the standard error of the estimate $\hat{\beta}_i$ from study i
- With homogeneous studies (H_0),
  - Q has a chi-square distribution with k-1 degrees of freedom
  - Q is compared to the upper 0.05 cutoff for a chi-square with k-1 df to determine if the heterogeneity is significant

23

23

Measuring Heterogeneity

- To measure heterogeneity, Higgins and Thompson proposed several statistics
  - Higgins JP, Thompson SG. (2002) Quantifying heterogeneity in a meta-analysis. Stat Med. 21:1539-58.
- $H^2 = Q / (k - 1)$
- Q is Cochran's heterogeneity statistic and k is the number of studies
- H describes the relative excess in Q over its degrees of freedom
  - Remember, a χ^2 statistics with k-1 degrees of freedom has an expected value of k-1

24

24

Measuring Heterogeneity

- Alternate measure of heterogeneity: I^2
  - More widely reported in meta-analysis
- I^2 describes the percentage of total variation across studies that is due to heterogeneity rather than chance
- $I^2 = (H^2 - 1)/H^2$ or alternatively $I^2 = [Q - (k-1)]/Q$
  - Q is Cochran's heterogeneity statistic and k = number of studies
- Negative values of I^2 are set to to zero so that I^2 lies between 0 and 1

25

25

Random effects model

- In the presence of heterogeneity, the fixed effect model is not appropriate
- As an alternative, in a random effects model, the β_i are assumed to follow a normal distribution with
  - mean = β
  - variance = τ^2
- These β_i are called random effects
- This is intended to incorporate a moderate amount of heterogeneity in the meta-analysis model

26

26

Random effects model

- The variance of an effect size estimate is more complicated than in the fixed effects model
- The variability of an effect size estimate $\hat{\beta}_i$ has two components in this model
  - Variability of $\hat{\beta}_i$ as an estimator of β_i – the *within* study variability (s_i^2)
    - Note that $SE(\hat{\beta}_i)$ provides an estimate of s_i
  - Variability of β_i around β – the *between* study variability (τ^2)

27

27

Random effects model

- With an estimate of τ^2
  - Make the study weights reflect both within and between study variation
  - For study i, the random effects weight is proportional to the reciprocal of the sum of the study standard error squared and the between study variability

$$w_i^* \propto \frac{1}{\left(SE\left(\hat{\beta}_i\right)^2 + \tau^2\right)}$$

28

28

Random effects model

- Sum these values for all of the primary studies and divide each by the total

$$w_i^* = \frac{\frac{1}{\left(SE\left(\hat{\beta}_i\right)^2 + \tau^2\right)}}{\frac{1}{\left(SE\left(\hat{\beta}_1\right)^2 + \tau^2\right)} + \frac{1}{\left(SE\left(\hat{\beta}_2\right)^2 + \tau^2\right)} + \cdots + \frac{1}{\left(SE\left(\hat{\beta}_k\right)^2 + \tau^2\right)}}$$

- I use w* to denote the random effects weights to keep them separate from the fixed effects weights

29

29

How to estimate τ^2

- One approach to estimate τ^2 is from DerSimonian and Laird
  - Method of moments (MOM) estimate
  - Based on the Q-test statistic

$$\hat{\tau}^2 = \frac{Q - (k-1)}{\sum_{i=1}^{k} \frac{1}{SE\left(\hat{\beta}_i\right)^2} - \frac{\sum_{i=1}^{k} \frac{1}{SE\left(\hat{\beta}_i\right)^4}}{\sum_{i=1}^{k} \frac{1}{SE\left(\hat{\beta}_i\right)^2}}}$$

30

30

Meta-analysis issues specific to genetic studies

- Strand issue
- Choice of effect allele
- Combining results from genome-wide association studies using different genotyping platforms
  - Different SNPs typed in each study
  - Number of overlapping SNPs may be small
- Genomic control adjustment

31

31

Meta-analysis of genetic studies: Strand issue

- SNPs can be genotyped on the "forward" strand or on the reverse "strand" depending on primers used
  - Allele A is complementary to T, C is complementary to G
    - A forward strand ⇒ T reverse strand
    - C forward strand ⇒ G reverse strand
    - G forward strand ⇒ C reverse strand
    - T forward strand ⇒ A reverse strand
- For a SNP with alleles A/C, A/G, C/T or G/T, it is easy to "reverse" the strand
  - Substitute T for A, G for C, C for G and A for T
  - Genotype AA becomes TT, AC becomes GT, CC becomes GG, etc.

32

32

Meta-analysis of genetic studies: Strand issue

- However, for SNPs with alleles A/T or C/G, cannot make the correction without information about the strand that the SNPs was typed
  - AA would become TT and TT would become AA, so cannot determine which strand a SNP is genotyped
  - May be able to make use of allele frequency
  - However, for SNPs with frequency ~ 50%, may not be able to resolve the strand issue without further information
- For genome wide association study, some studies report association on the forward strand to facilitate meta-analysis with other studies
- For imputed genotypes, not an issue if all studies used the same imputation approach/reference panel

33

33

Meta-analysis of genetic studies: Choice of effect allele

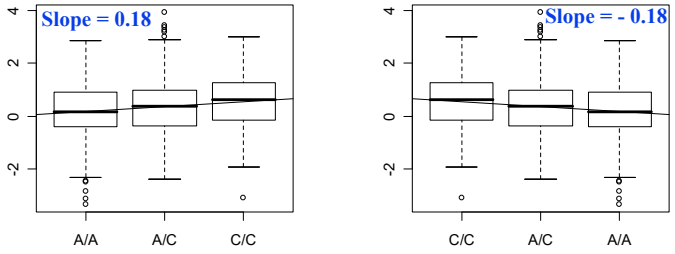
- Most common genetic model used for meta-analysis: additive model
- Each study report the association of trait with genotype coded as 0, 1 or 2 copies (or dosage) of the effect allele
- When studying genetic association, the choice of an effect allele may differ between studies. Some options:
  - Effect allele = minor allele
    - Minor allele may not be the same for SNPs with allele freq. ~ 50%
  - Effect allele = second alphabetical allele
    - Default for some software, such as R

34

34

Meta-analysis of genetic studies: Choice of effect allele

- P-value not affected by choice of effect allele
- Estimate of effect size has opposite sign depending on choice of effect allele



- Want to take the direction of effect in consideration when combining results from multiple studies

35

35

Meta-analysis of genetic studies: Choice of effect allele

- Each study to be meta-analyzed needs to provide information about the effect and non-effect alleles
- For each SNP, an effect allele is selected
- For studies with the opposite effect allele, the sign of the effect estimate is "flipped"
  - ➢ Positive effect becomes negative effect estimates, and vice versa

36

36

Meta-analysis of genetic studies: Studies with different genotyping platforms

- When studies used the same set of SNPs, meta-analysis can be performed after taking care of the strand and effect allele issues described in the previous slides
- When studies used different sets of SNPs, common SNPs may be analyzed
- However, the set of common SNPs across platforms may be small

37

37

Meta-analysis of genetic studies: Studies with different genotyping platforms

- Imputation software provide a probability for each possible genotype, i.e. three probabilities (summing to 1) for each SNP and each individual
- Typical to use "dosage" as the "genotype" for imputed SNPs
  - Dosage = Prob. of 1 effect allele + 2 × Prob. of 2 effect alleles
  - May also use "most probable" genotypes for each individual, although information about the imputation uncertainty is lost
  - May also use the posterior probabilities in the formulation of the likelihood
- Results from imputed and genotyped SNPs may be meta-analyzed together

38

38

Genomic Control Adjustment

- Genomic control computed under the assumption that most SNPs in a GWAS are unassociated, except for a very small proportion
  - Measure inflation in the test statistics
- To compute genomic control correction (λ_{GC}):
  - Convert test statistic p-values to χ^2 statistic with 1 df
  - Compute median χ^2 statistic from GWAS results
  - Compute ratio of observed to expected median χ^2 statistic
    - Expected median χ^2 statistic with 1 df = 0.4549

$$\lambda_{GC} = \frac{\text{Observed Median } \chi^2 \text{ statistic}}{0.4549}$$

- Median should not be overly affected by a small number of associated SNPs with low p-values

- Devlin B, Roeder K. 1999. Genomic control for association studies. Biometrics 55; 997-1004.
- Reich DE, Goldstein DB. 2001. Detecting association in a case-control study while correcting for population stratification. Genet Epidemiol. 20:4-16.

39

Genomic Control Adjustment

- Corrects for inflation in the test statistics (which may be due to population stratification or cryptic relatedness) by adjusting association statistics at each marker by a uniform overall inflation factor
  - Divide observed χ^2 statistic by λ_{GC} at each marker:

    χ^2 adjusted = observed χ^2 / λ_{GC}
  - Compute adjusted p-value based on observed χ^2 / λ_{GC}
  - Correction is not applied if $\lambda_{GC} < 1$
- Correction may not be sufficient at markers having unusually strong differentiation across ancestral populations
- Correction may be unnecessarily for markers without differentiation across ancestral populations
- Better to investigate why inflated test statistics are observed than to apply a uniform correction

40

Meta-analysis of genetic studies: Software

- Weighted (signed) Z-score approach is default analysis implemented in software METAL
  - ➢http://www.sph.umich.edu/csg/abecasis/Metal/index.html
- Fixed effect inverse variance weighted approach also available in METAL
  - ➢Heterogeneity testing (Q-test) optional
  - ➢Random effect model implemented in a patched version of metal: https://github.com/explodecomputer/random-metal
- METAL will check for strand inconsistencies for SNPs with A/C, A/G, C/T and G/T alleles
  - ➢ Can handle A/T and C/G SNPs when strand information is provided
- METAL will select an effect allele and change the sign of effect estimate for studies with opposite effect allele
- Other meta-analysis software for genetic analysis will perform similar analyses (GWAMA)

41

41

Using METAL & GWAMA software to perform meta-analyses of genome-wide association scans

42

42