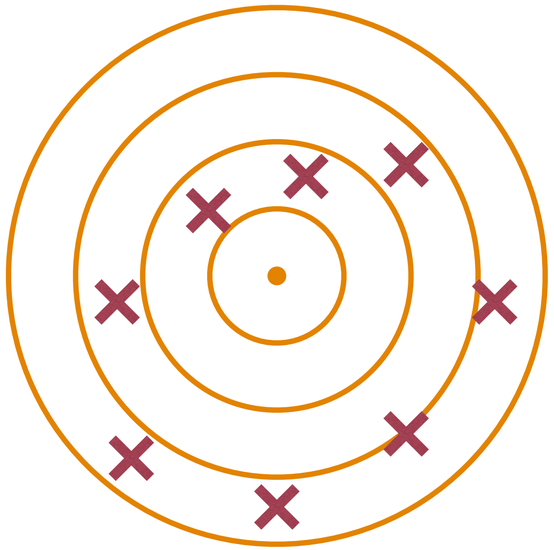
Lecture 4

Meta-Analysis in GWAS: An Overview and Applications

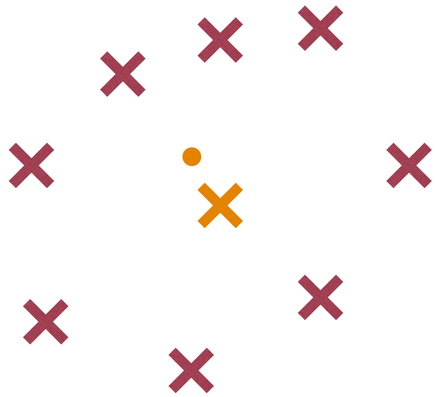
by Dr. Mustafa İsmail Özkaraca

Contents

What is Meta-Analysis?

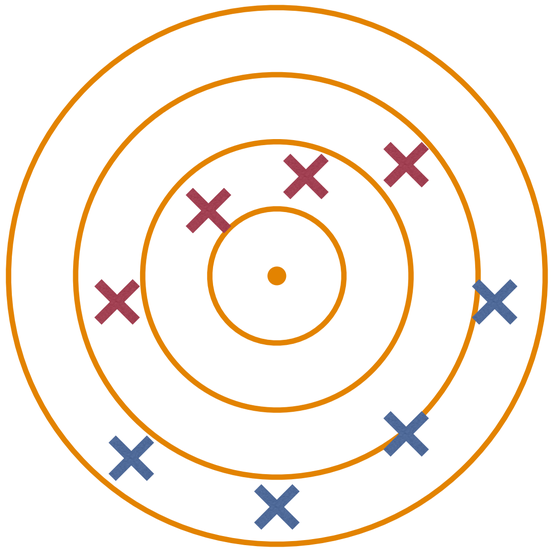


Q: Where is the **centre** of the board?

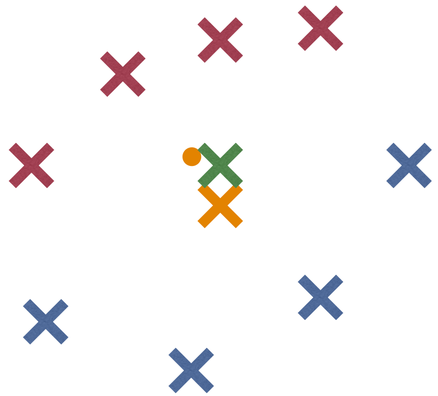


- Single estimates
- Center of 8 estimates (**Meta Estimate**)
- True center

What is Meta-Analysis?



Q: Where is the **centre** of the board?



Single estimates (Worse)

Informed Meta Estimate

Single estimates (Better)

Center of 8 estimates (**Meta Estimate**)

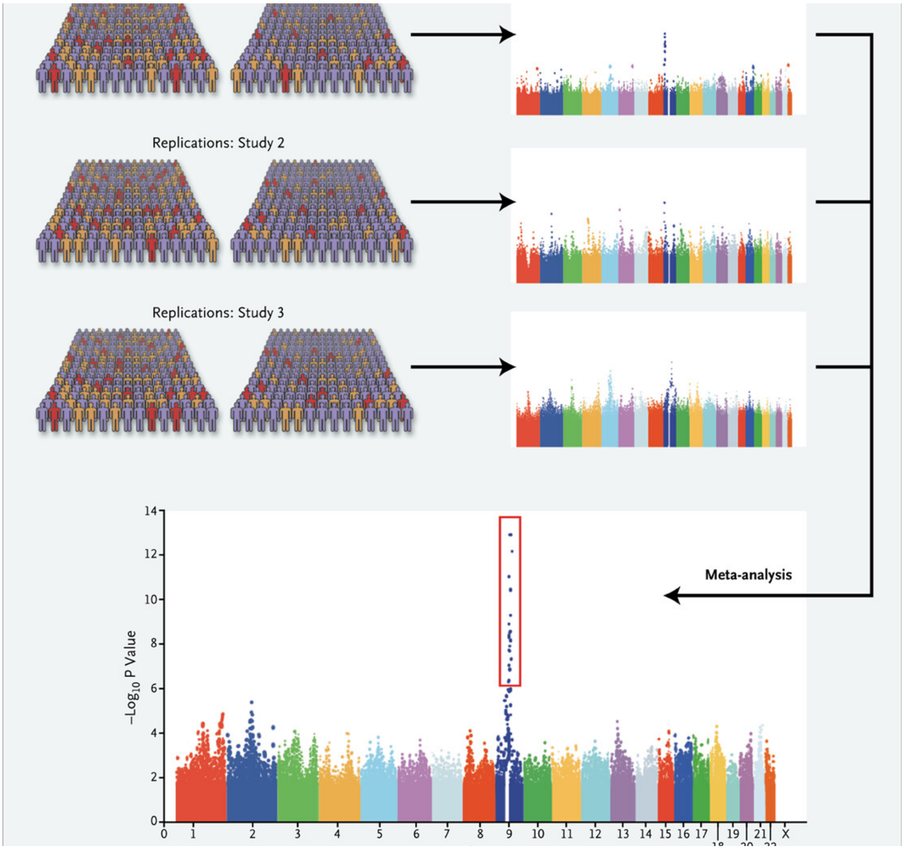
True center

"*Meta-analysis* is the art of **harmonising multiple results** in order to arrive **more accurate** conclusions". (MIO)

Why Meta-Analyse in GWAS?

RECALL: (GWAS Challenges)

- Multiple Testing Burden
- GWAS requires large-scale datasets
- GWAS software capable handling large-scale datasets



[1] Figures From:

Manolio, Teri A.

"Genomewide association studies and assessment of the risk of disease."

New England journal of medicine 363.2 (2010): 166-176.

Types of Meta-Analysis in GWAS

Individual Participant Data Meta-Analysis

Summary Data Meta-Analysis

Types of Meta-Analysis in GWAS

PART 1: pvalue based Methods

Fisher's method:

$$\chi^2 = 2\sum_{i=1}^{K} \log(pval_i)$$

Under the null (i.e. homogeneity): $\chi^2 \sim \chi^2_{2K}$

Z-scores method:

$$z = 2\frac{\sum_{i=1}^{K} z_i\sqrt{N_i}}{\sqrt{\sum_{i=1}^{K} N_i}}$$

N_i is sample size,

$$z_i = \Phi^{-1}(1 - p_i/2) * (direction\ of\ study\ i)$$

Disadvantages:

A. No information about the size of the effects,

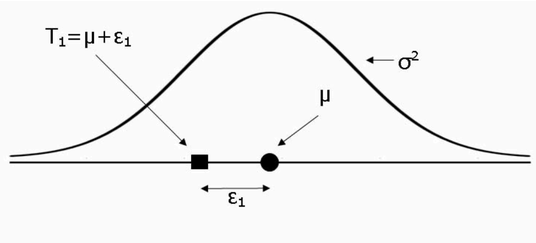
B. Assumes homogeneity,

C. No direction of effect (for Fisher's method)

Types of Meta-Analysis in GWAS

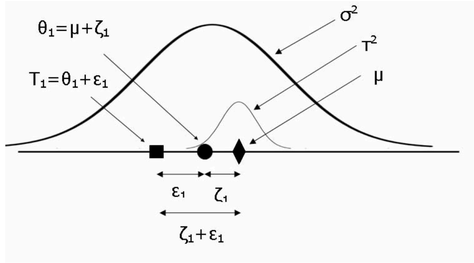
PART 2: Effect size based Methods

Fixed Effect Model [1]



Fixed effect model. The observed effects are sampled from a distribution with true effect μ, and variance σ^2. The observed effect T_1 is equal to $\mu+\varepsilon_i$.

Random Effects Model [1]



Random effects model. The observed effect T_1 (box) is sampled from a distribution with true effect θ_1, and variance σ^2. This true effect θ_1, in turn, is sampled from a distribution with mean μ and variance τ^2.

[1] Figures are from:

https://meta-analysis.com/download/Meta-analysis_fixed_effect_vs_random_effects 072607.pdf

Borenstein, Michael, Larry Hedges, and Hannah Rothstein.
"Meta-analysis: Fixed effect vs. random effects."
Meta-analysis.com (2007): 1-162.

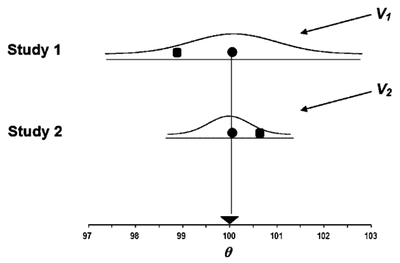


Figure 3. Schematic of the fixed-effect model.

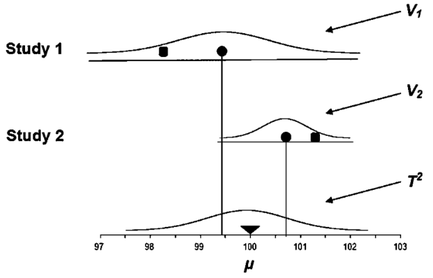


Figure 4. Schematic of the random-effects model.

Figure 3 and Figure 4 are from: Borenstein, Michael, et al. ”A basic introduction to fixed-effect and random-effects models for meta-analysis.” (2010).

Types of Meta-Analysis in GWAS

PART 2: Effect size based Methods

Fixed Effect (Inverse Variance Method - IVM [2]):

$$\widehat{\beta_{meta}} = \sum_{i=1}^{K} \omega_i \hat{\beta}_i$$

where $\omega_i = \frac{V_i^{-1}}{\sum_{i=1}^{K} V_i^{-1}}$,

V_i is the variance for study i.

Random Effects (DerSimonian and Laird Estimator [3]):

$$\widehat{\beta_{meta}} = \sum_{i=1}^{K} \omega_i \hat{\beta}_i$$

where $\omega_i = \frac{(V_i + T^2)^{-1}}{\sum_{i=1}^{K} (V_i + T^2)^{-1}}$,

V_i is the variance for study i,

T^2 is the between-study variance.

[2] Borenstein, Michael, et al. ”A basic introduction to fixed-effect and random-effects models for meta-analysis.” (2010).

[3] DerSimonian, Rebecca, and Nan Laird. ”Meta-analysis in clinical trials.” (1986).

Types of Meta-Analysis in GWAS

PART 2: Effect size based Methods (When to use Random Effects Model?)

1. Cochran's Q [4]

$$Q = \sum_{i=1}^{K} \omega_i(\widehat{\beta_i} - \widehat{\beta_{IVM}})^2, \text{ under the null } Q \sim \chi^2_{K-1}.$$

2. I^2 [4]

$$I^2 = 100 * \left(1 - \frac{K-1}{Q}\right)$$

3. T^2 [4]

$$T^2 = \frac{Q - K + 1}{A - \frac{B}{A}} \text{ where } A = \sum_{i=1}^{K} \omega_i \text{ and } B = \sum_{i=1}^{K} \omega_i^2$$

"**The strategy of starting with a fixed-effect model and then moving to a random-effects model** if the test for heterogeneity is significant relies on a flawed logic and **should be strongly discouraged**." [5]

"Fixed-effects calculations avoided power deserts and maximized discovery of association signals at the expense of much higher false-positive rates. Therefore, random- and fixed-effects models are preferable for different purposes (**fixed effects for initial screenings, random effect is for generalizability applications**)." [6]

"**Random effects models are not used in discovery efforts** owing to far more limited power than fixed effects models; however, random effects models are more appropriate than fixed effects models when the aim is to consider the generalizability of the observed association and estimate the average effect size of the associated variant and its uncertainty across different populations: for example, for predictive purposes[28,33]." [7]

"The conventional wisdom in the statistical literature is that **when heterogeneity is present** or even likely, **the random effects model is more appropriate than the fixed effects model**. **We suggest that this might not be the right approach for GWAS**, because (i) the number of studies being combined is often not very large (leading to an imprecise heterogeneity estimate) and (ii) the form of the heterogeneity typically does not fit a Gaussian random effects model." [8]

[4] Borenstein, Michael, et al. *Introduction to meta-analysis*. John Wiley & Sons, 2009.
[5] Borenstein, Michael, et al. "A basic introduction to fixed-effect and random-effects models for meta-analysis." (2010).
[6] Pereira, Tiago V., et al. "Discovery properties of genome-wide association signals from cumulatively combined data sets." (2009).
[7] Evangelou, Evangelos, and John PA Ioannidis. "Meta-analysis methods for genome-wide association studies and beyond." (2013).
[8] Begum, Ferdouse, et al. "Comprehensive literature review and statistical considerations for GWAS meta-analysis." (2012)

Types of Meta-Analysis in GWAS

PART 2: Effect size based Methods (When to use Random Effects Model?)

"A popular way out of this dilemma is **to start with a fixed effects meta-analysis but to report the random effects meta-analysis** when heterogeneity is found in a top hit." [9]

Example Study: GWAS on Height [10]

A.5.4 million individuals, 5 ancestry groups.

B.Meta-Analysis of Meta-analyses.

C.IVM applied within each ancestry groups. Then IVM applied between ancestry groups.

Heterogeneity in meta analysis of GWAS:

"Using data from a meta-analysis of seven GWA studies on obesity, we concluded that moderate or larger heterogeneity was common in meta-analysis of GWA studies." [11]

[9] Thompson, John R., John Attia, and Cosetta Minelli. "The meta-analysis of genome-wide association studies." (2011).
[10] Yengo, L., Vedantam, S., Marouli, E. et al. A saturated map of common genetic variants associated with human height. Nature 610, 704–712 (2022).
[11] Pei, Yu-Fang, et al. "Exploring the major sources and extent of heterogeneity in a genome-wide association meta-analysis." (2016).

Types of Meta-Analysis in GWAS

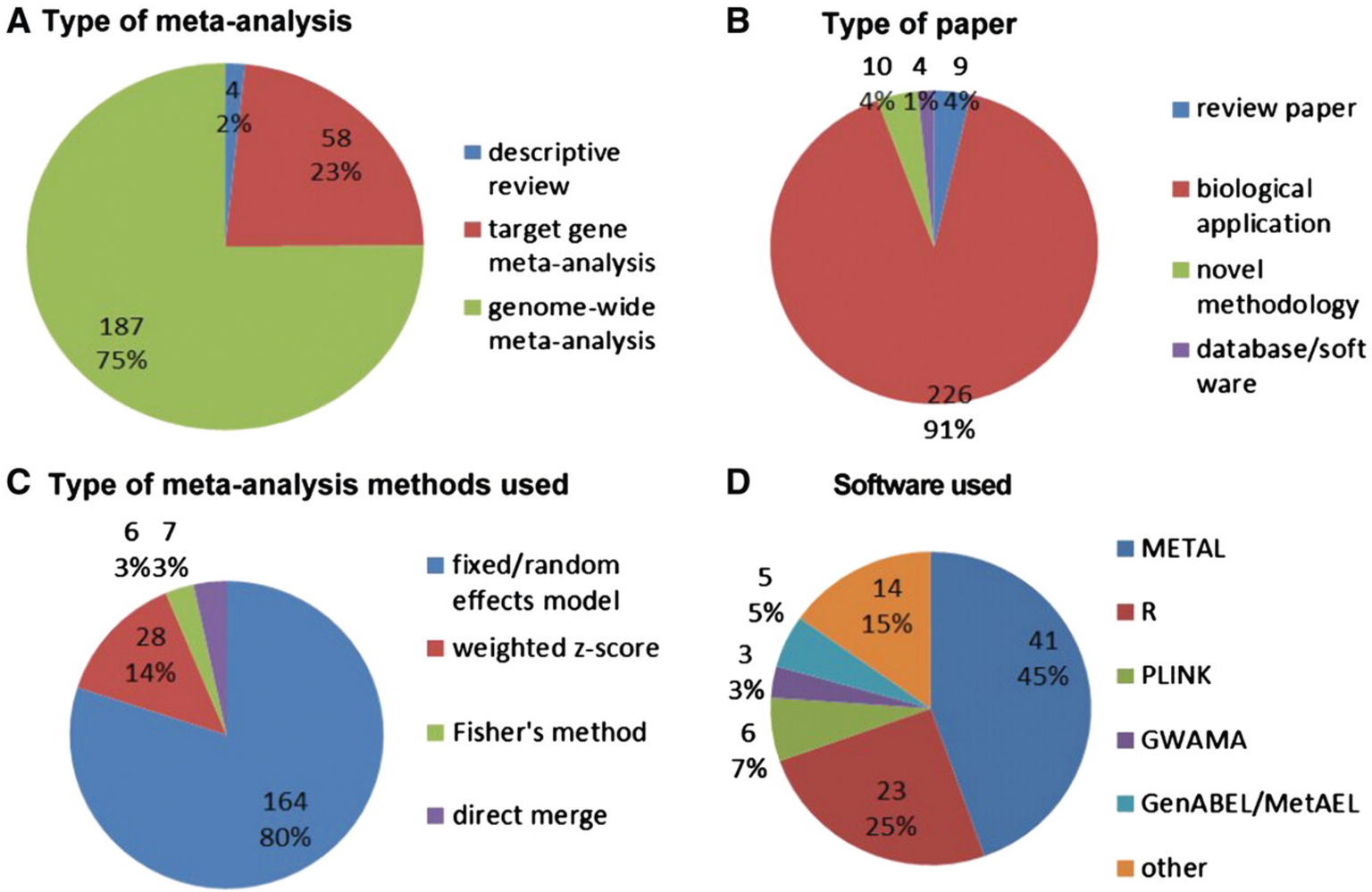


Figure from: Begum, Ferdouse, et al. "Comprehensive literature review and statistical considerations for GWAS meta-analysis." *Nucleic acids research* 40.9 (2012): 3777-3784.

What’s Next

1. Understanding Polygenic Risk Score (PRS)?

2. Calculating PRS

3. Population Bias