Nan. M Laird
Harvey V. Fineberg Professor Emerita of Biostatistics
Harvard University

META-ANALYSIS: HETEROGENEITY CAN BE A GOOD THING

META-ANALYSIS: DEFINITION

The statistical analysis of the findings of a collection of individual studies (Glass, 1976)

Differs from a review, uses a quantitative summary from each study for overall result.

Background

- 1970's used extensively in social sciences
- Big questions---

  Is psychotherapy effective?

  What is the optimal class size for learning?

  Are there teacher expectancy effects?
- The Placebo Effect (Beecher, 1955)

TABLE 2.—*Therapeutic Effectiveness of Placebos in Several Conditions*

Condition	Study	Placebo: Agent	Placebo: Route *		Patients, No.	% Satisfactorily Relieved by a Placebo	
Severe postoperative wound pain	Keats, A. S., and Beecher, H. K.: J. Pharmacol. & Exper. Therap. **100:** 1-13, 1950	Saline	I. V.		118	21	
	Beecher, H. K., and others: U. S. Armed Forces M. J. **2:** 1269-1276, 1951	Saline	S. C.		29	31	
	Keats, A. S., and others: J. A. M. A. **147:** 1761-1763 (Dec. 29) 1951	Saline	I. V.		34	26	
	Beecher and others (1953) [3]	Lactose	P. O.		52	40	33
					36	26	
					44	34	
					40	32	
	Lasagna and others (1954) [4]	Saline	S. C.		14	50	39
					20	37	
					15	53	
					21	40	
					15	40	
					15	15	
Cough	Gravenstein, J. S., and others: J. Appl. Physiol. **7:** 119-139, 1954	Lactose	P. O.		22	36	40
					27	43	
Drug-induced mood changes	Lasagna, L., and others: J. A. M. A. **157:** 1006-1020 (March 19) 1955	Isotonic sodium chloride	S. C.	Normal	20	30	
				"Post-addicts"	30	30	
Pain from angina pectoris	Evans, W., and Hoyle, C.: Quart. J. Med. **2:** 311-338, 1933	Sodium bicarbonate	P. O.		66	38	
	Travell, J., and others: Ann. New York Acad. Sc. **52:** 345-353, 1949	"Placebo"	P. O.		19	26	
	Greiner, T., and others: Am. J. Med. **9:** 143-155, 1950	Lactose	P. O.		27	38	
Headache	Jellinek (1946)	Lactose	P. O.		199	52	
Seasickness	Gay and Carliner (1949)	Lactose	P. O.		33	58	
Anxiety and tension	Wolf and Pinsky (1954) [5]	Lactose	P. O.		31	30	
Experimental cough	Hillis (1952)	Isotonic sodium chloride	S. C.	Many experiments	1	37	
Common cold	Diehl, H. S.: J. A. M. A. **101:** 2042-2049 (Dec. 23) 1933	Lactose	P. O.	Cold acute	110	35	
				Subacute chronic	48	35	
				Total patients	1,082	Average relieved	35.2 ± 2.2%

* I. V., intravenous; S. C., subcutaneous; P. O., oral.

Early Methods for Summarizing

- Largely descriptive reviews:

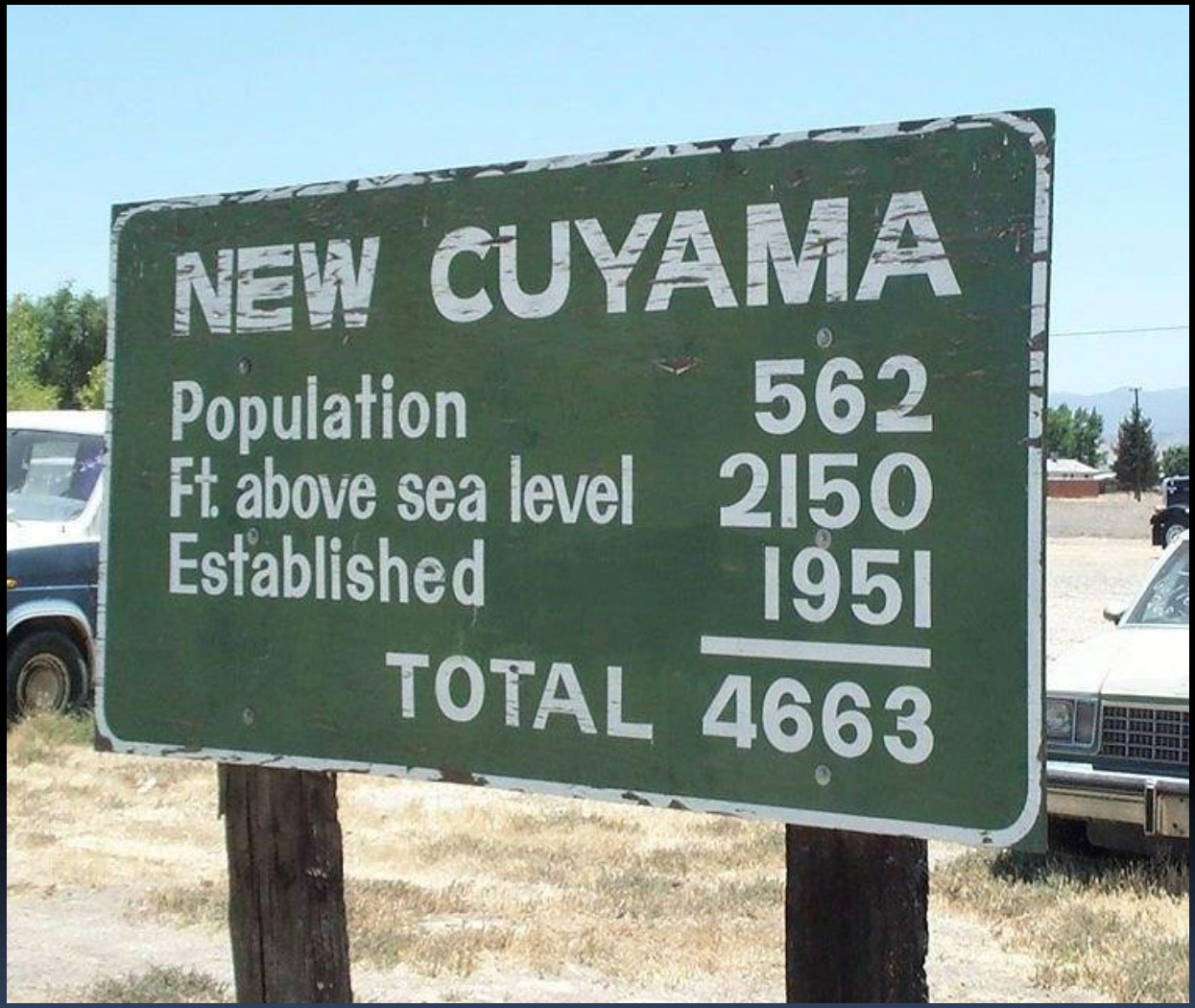
  Report on studies and design of the studies, subject populations and subject characteristics, investigators, outcome measures, study follow-up, etc.

- Meta- Analysis:

  Use quantitative summary from each study.

How to construct comparable effect sizes for combining results, how to weight studies.

- Heterogeneity: Study results differ from one another. How to handle heterogeneity?


NEW CUYAMA
Population 562
Ft. above sea level 2150
Established 1951
TOTAL 4663

Notation

Y = effect size for each study

(risk difference, OR, or mean difference, regression coefficient)

s^2 is standard error of Y obtained from each study

Random Effects Model
(Cochran, 1954)

$$Y_i = T_i + e_i$$

e_i is sampling variability for i^{th} study, $\text{var}(e_i) = s_i^2$

$$E(T_i) = \mu$$

$$\text{var}(T_i) = \sigma^2$$

T_i, true study effects, are treated as random

Method of Analysis

$$\begin{aligned} E(Y_i) &= \mu \\ \text{Var}(Y_i) &= \sigma^2 + s_i^2 \end{aligned}$$

Decompose sample variance of Y_i

$$\sum \left[(Y_i - \bar{Y})^2 / s_i^2 \right]$$

To obtain a 'Method of Moments' estimator of σ^2, $\hat{\sigma}^2$.

Estimate μ as

$$\begin{aligned} \hat{\mu} &= \sum \left[Y_i / \left(\hat{\sigma}^2 + s_i^2 \right) \right] \\ \text{var}\hat{\mu} &= \left[\sum \left(\frac{1}{\hat{\sigma}^2 + s_i^2} \right) \right]^{-1} \end{aligned}$$

Assumptions, n_{c_i}, n_{t_i} large, K large, T_i and e_i are independent

Random Effects Approach to Meta-Analysis

- Advantages:

Provides a quantitative measure of how results differ, above and beyond sampling error.

- Criticisms:

Studies should not be combined if effects differ.

- Major purpose of meta-analysis:

Provide an overall summary, characterize and report variation in study results.

Fixed Effects Model

- Set $\sigma = 0$, analysis remains same:
- Estimate μ with a weighted mean:

$$w_i^f = 1/s_i^2$$

- Weight driven by sample size
- Estimates of μ may be similar, but standard errors will be smaller under FE model:

$$w_i^r = 1/(s_i^2 + \sigma^2) < w_i^f$$

$$se^r > se^f$$

- RE confidence intervals and standard errors are smaller than corresponding FE estimates.

Meta-Analysis in Genetic Association Studies

- Genetic Association Study: Usually a case-control study comparing genetic variants in cases and controls.
- Special issue: GWAS tests for millions of genetic variants, each test separately; primary interest is testing
- Effect sizes can be very small and typical significance level is 10^{-8}
- Lot of interest in meta-analysis for increased power and replication
- Most individual studies are large (thousands of individuals, and number of studies is small)

Obesity and the Insig2 Variant

- Insig2 variant found in one of first GWAS in Framingham Study Population
- Replicated in the original paper in 5 other cohorts and not replicated in one cohort
- Studies can be classified by design as: General population study, healthy population study, or comparative (case-control) study

Table 3. Summary of study results and meta-analysis. All values given are for a recessive model. NHS, Nurses Health Study; TDT, transmission disequilibrium test; PBAT, tools for FBATs.

Study	Design	Total genotyped	Obese	Non-obese	Test	Total number of families		*P* value
FHS	Family	694			PBAT	288		0.0030
Maywood	Family	866	361	505	PBAT dichotomous	342		0.0090
Maywood	Family	866			PBAT quantitative	342		0.0700
Essen children/adolescents	Trios	1104	368	–	TDT	368		0.0020
						Adjustments		
KORA	Cohort	3996			Linear regression	Sex, age		0.0080
NHS	Cohort	2726			Linear regression	Age		–
			Obese	*Non-obese*	*Test*	*OR*	*95% CI*	
KORA	Cohort	3996	935	3061	Logistic regression	1.32	1.06–1.65	0.0167
NHS	Cohort	2726	503	2223	Chi-squared test	0.81	0.58–1.13	–
American/Polish Caucasian	Case-control	2761	1835	926	Chi-squared test	1.40	1.08–1.78	0.0200
Maywood	Case-control	398	216	182	Fischer's exact test	2.36	–	0.0400
Pooled OR (all)		9881	3445	6426	2-tailed Mantel-Haenszel	1.22	1.05–1.42	0.0080

Meta-Analysis Designed to Explain Heterogeneity

- Lot of controversy about original paper
- Every clinical study measures height and weight and BMI
- Can measure one genetic variant easily
- Many, many 'studies' published on this association
- (Heid, et al, 2009) Collected results from 34 publications, 74,000 subjects to look at effect of study population

Study Design

- General Population Study (GP)

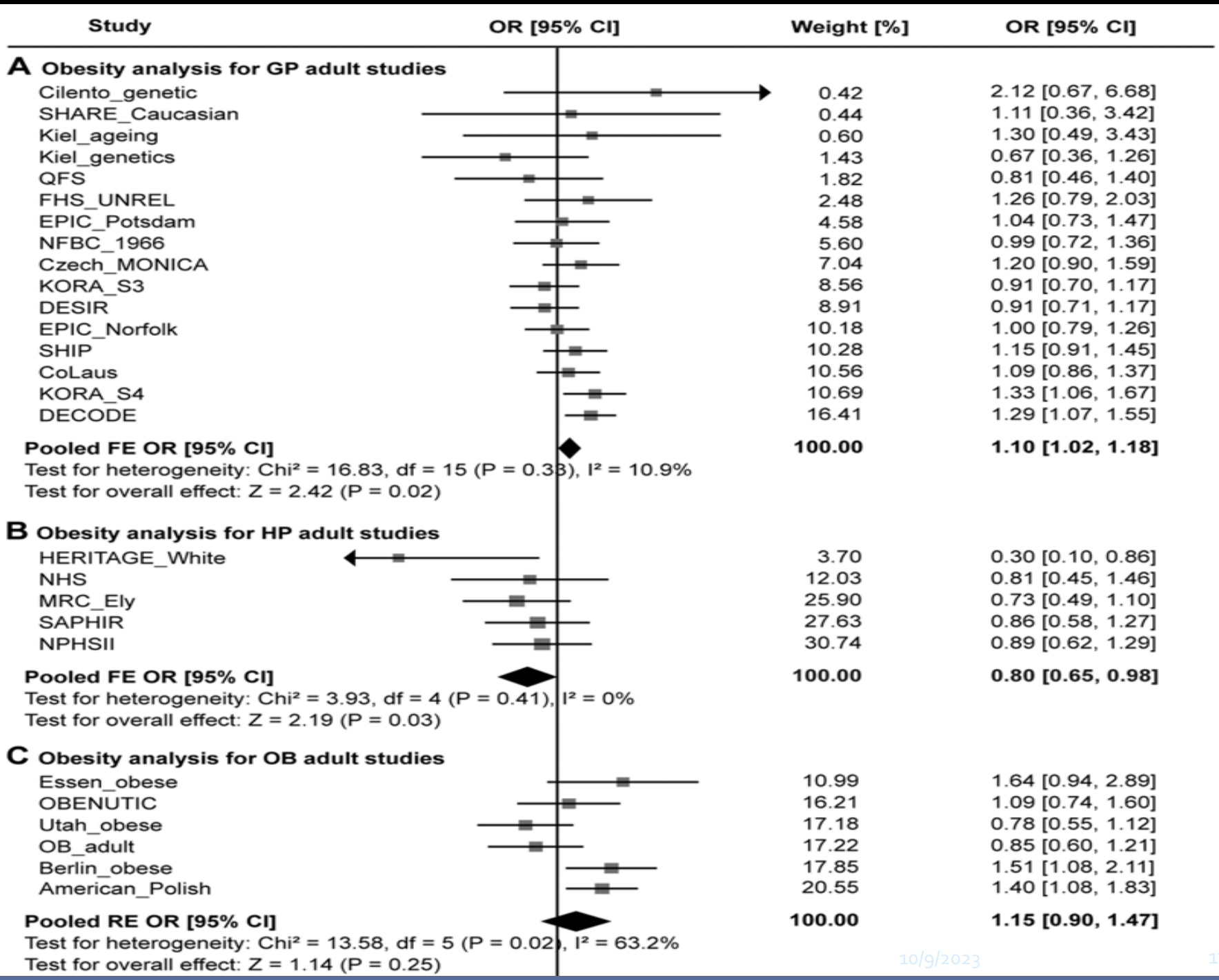
Subjects included regardless of health status.

- Healthy Populations (HP)

Subjects deselected on basis of health conditions

- Subjects selected on basis of obesity

More powerful, and specific design


Study
OR [95% CI]
Weight [%]
OR [95% CI]
A Obesity analysis for GP adult studies
Cilento_genetic 0.42 2.12 [0.67, 6.68]
SHARE_Caucasian 0.44 1.11 [0.36, 3.42]
Kiel_ageing 0.60 1.30 [0.49, 3.43]
Kiel_genetics 1.43 0.67 [0.36, 1.26]
QFS 1.82 0.81 [0.46, 1.40]
FHS_UNREL 2.48 1.26 [0.79, 2.03]
EPIC_Potsdam 4.58 1.04 [0.73, 1.47]
NFBC_1966 5.60 0.99 [0.72, 1.36]
Czech_MONICA 7.04 1.20 [0.90, 1.59]
KORA_S3 8.56 0.91 [0.70, 1.17]
DESIR 8.91 0.91 [0.71, 1.17]
EPIC_Norfolk 10.18 1.00 [0.79, 1.26]
SHIP 10.28 1.15 [0.91, 1.45]
CoLaus 10.56 1.09 [0.86, 1.37]
KORA_S4 10.69 1.33 [1.06, 1.67]
DECODE 16.41 1.29 [1.07, 1.55]
Pooled FE OR [95% CI] 100.00 1.10 [1.02, 1.18]
Test for heterogeneity: Chi² = 16.83, df = 15 (P = 0.33), I² = 10.9%
Test for overall effect: Z = 2.42 (P = 0.02)
B Obesity analysis for HP adult studies
HERITAGE_White 3.70 0.30 [0.10, 0.86]
NHS 12.03 0.81 [0.45, 1.46]
MRC_Ely 25.90 0.73 [0.49, 1.10]
SAPHIR 27.63 0.86 [0.58, 1.27]
NPHSII 30.74 0.89 [0.62, 1.29]
Pooled FE OR [95% CI] 100.00 0.80 [0.65, 0.98]
Test for heterogeneity: Chi² = 3.93, df = 4 (P = 0.41), I² = 0%
Test for overall effect: Z = 2.19 (P = 0.03)
C Obesity analysis for OB adult studies
Essen_obese 10.99 1.64 [0.94, 2.89]
OBENUTIC 16.21 1.09 [0.74, 1.60]
Utah_obese 17.18 0.78 [0.55, 1.12]
OB_adult 17.22 0.85 [0.60, 1.21]
Berlin_obese 17.85 1.51 [1.08, 2.11]
American_Polish 20.55 1.40 [1.08, 1.83]
Pooled RE OR [95% CI] 100.00 1.15 [0.90, 1.47]
Test for heterogeneity: Chi² = 13.58, df = 5 (P = 0.02), I² = 63.2%
Test for overall effect: Z = 1.14 (P = 0.25)

INSIG2 rs7566605 association with obesity
Main meta-analysis results

	# cases/controls (# studies)	OR (p-value) fixed effect	OR (p-value) random effect	I^2 (p-value)
All	16,365/49,848 (27)	1.076 (0.023)	1.051 (0.268)	41.0 (0.015)
GP	9162/39,682 (16)	1.097 (0.015)	1.092 (0.035)	10.9 (0.329)
HP	1307/6333 (5)	0.796 (0.028)	0.796 (0.028)	0.0 (0.415)
OB	5896/3833 (6)	1.163 (0.018)	1.152 (0.253)	63.2 (0.018)

Summary

- Heterogeneity does not preclude combining results. Important to report measure of heterogeneity, presence required caution in interpretation
- Heterogeneity can suggest important insights as to the nature of interventions, of populations, or both