

Evidence Synthesis / Meta-Analysis

Guido Knapp

Department of Statistics, TU Dortmund University

Gerta Rücker and Guido Schwarzer

Center for Medical Biometry and Medical Informatics
University of Freiburg

*DAGStat 2016
Göttingen, March 14, 2016*

Introduction

- *Meta-analysis*, a term coined by Glass (1976), is intended to provide *the statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings.*
- Pearson (1904) used data from five independent samples and computed a pooled estimate of correlation between mortality and inoculation in order to evaluate the efficacy of the vaccine for enteric fever.
- Birge (1932) combined estimates across experiments at different laboratories to establish reference values for some fundamental constants in physics.
- Early works of Cochran (1937), Yates and Cochran (1938), Tippet (1931), and Fisher (1932) dealt with combining information across experiments in the agricultural sciences.

Agenda

- Session 1: 9:00am – 10:30am
Topics: Introduction; Generic Fixed-Effect and Random-Effects Meta-Analysis Models
- Session 2: 11:00am – 12:30pm
Topics: Meta-Analysis with Continuous and Binary Outcome
- Session 3: 1:30pm – 3:00pm
Topics: Meta-Regression; Small Study Effects
- Session 4: 3:30pm – 5:00pm
Topics: Network Meta-Analysis

Introduction

Areas of application

- Social sciences:
Reliability and validity studies, teacher expectancies studies, ...
- Life sciences:
Effectiveness of drugs, second-hand smoking, ...
- Archaeology, astronomy, chemistry,
engineering, environmental sciences, geosciences,
military operations analysis, official statistics,
physics, psychology, ...
- Interlaboratory trials, Metrology

Introduction

Four important stages of research synthesis:

1. **Problem formulation** stage
2. **Data collection** stage
3. **Data evaluation** stage
4. **Data analysis and interpretation** stage

Data analysis: Results from **published** studies

Examples

Four experiments about the percentage of albumin in plasma protein in human subjects

Experiment	n_i	Mean	Variance
A	12	62.3	12.986
B	15	60.3	7.840
C	7	59.5	33.433
D	16	61.5	18.513

Examples

Studies of the relationship between an observation measure of teacher indirectness and student achievement

Study	No. of teachers	Correlation coefficient r
1	15	-0.073
2	16	0.308
3	15	0.481
4	16	0.428
5	15	0.180
6	17	0.290
7	15	0.400

Examples

Number of patients and mortality rate from all causes, for six trials comparing the use of aspirin and placebo by patients following a heart attack

Study	Aspirin		Placebo	
	No. of Pat.	Mort. Rate (%)	No. of Pat.	Mort. Rate (%)
UK-1	615	7.97	624	10.74
CDPA	758	5.80	771	8.30
GAMS	317	8.52	309	10.36
UK-2	832	12.26	850	14.82
PARIS	810	10.49	406	12.81
AMIS	2267	10.85	2257	9.70

Examples

Number of cases of lung cancer in women who did not actively smoke cigarettes and estimated relative risk of lung cancer in relation exposure to environmental tobacco smoke

Study	No. of Cases	Estimated RR (95% CI)	Study	No. of Cases	Estimated RR (95% CI)
1	94	1.52 (0.88 - 2.63)	11	24	0.79 (0.25 - 2.45)
2	19	1.52 (0.39 - 5.99)	12	86	1.55 (0.90 - 2.67)
3	41	0.81 (0.34 - 1.90)	13	199	1.65 (1.16 - 2.35)
4	84	0.75 (0.43 - 1.30)	14	60	2.01 (1.09 - 3.71)
5	22	2.07 (0.82 - 5.25)	15	32	1.03 (0.41 - 2.55)
6	246	1.19 (0.82 - 1.73)	16	67	1.28 (0.76 - 2.15)
7	134	1.31 (0.87 - 1.98)	17	34	1.26 (0.57 - 2.82)
8	54	2.16 (1.08 - 4.29)	18	62	2.13 (1.19 - 3.83)
9	20	2.34 (0.81 - 6.75)	19	28	1.41 (0.54 - 3.67)
10	22	2.55 (0.74 - 8.78)			

Examples

Studies on the effectiveness of anti-TNF- α inhibitors in the treatment of rheumatoid arthritis. Data given are mean change from baseline under treatment for two outcomes (DAS-28 and HAQ score)

Author	Year	mean.das	se.das	mean.haq	se.haq
Bennet	2005	-1.7	0.25	-0.31	0.13
Bingham	2009	-1.6	0.10	-0.35	0.05
Bombardieri	2007	-1.9	0.05	-0.48	0.02
Navarro-Sarabia	2009	-1.1	0.18	-0.21	0.07
Van der Bijl	2008	-1.5	0.25	-0.21	0.08

Examples

13 trials on the prevention of tuberculosis using BCG vaccination

Trial	Vaccinated		Not vaccinated		Latitude
	Disease	No disease	Disease	No Disease	
1	4	119	11	128	44
2	6	300	29	274	55
3	3	228	11	209	42
4	62	13536	248	12619	52
5	33	5036	47	5761	13
6	180	1361	372	1079	44
7	8	2537	10	619	19
8	505	87886	499	87892	13
9	29	7470	45	7232	27
10	17	1699	65	1600	42
11	186	50448	414	27197	18
12	5	2493	3	2338	33
13	27	16886	29	17825	33

Examples

Studies designed to reduce blood glucose levels. Patients enrolled were treated with one of ten diabetes treatments. (Excerpt of dataset)

Study	TE	seTE	Treat1	Treat2
DeFronzo1995	-1.90	0.1414	metf	plac
Lewin2007	-0.82	0.0992	metf	plac
Wilms1999	-0.20	0.3579	metf	acar
Davidson2007	-1.34	0.1435	rosi	plac
Wolffenbuttel1999	-1.10	0.1141	rosi	plac
Kipnes2001	-1.30	0.1268	piog	plac
Kerenyi2004	-0.77	0.1078	rosi	plac
Hanefeld2004	0.16	0.0849	piog	metf
Derosa2004	0.10	0.1831	piog	rosi
:	:	:	:	:

Data

Possible data available from the published study:

- Raw data (very rare) or summary statistics (quite often)
- *P*-value only
- Effect size estimate plus standard error
- Effect size estimate plus confidence interval
- Effect size estimate plus *P*-value
- Effect size estimate plus sample size

Examples

- Validity studies: Correlation between student ratings of instructor with student achievement

Study	Sample	<i>n</i>	<i>r</i>
Bolton et al. (1979)	General psychology	10	0.68

- Studies of the effects of teacher expectancy on pupil IQ

Study	<i>d</i>	SE(<i>d</i>)
Rosenthal et al. (1974)	0.03	0.125

- Second-hand smoking

Study	RR (95% CI)
Akiba, Kato, and Blot (1986)	1.52 (0.88 – 2.63)

Levels of Evidence

Oxford Centre for Evidence-based Medicine - Levels of Evidence

Source: <http://www.cebm.net/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/>

Levels of evidence for therapy / prevention (excerpt)	
Level	Description
1a	Systematic review (with homogeneity*) of randomized controlled trials
1b	Individual randomized controlled trial (with narrow confidence interval)
...	...
2a	Systematic review (with homogeneity*) of cohort studies
2b	Individual cohort study (incl. low quality RCT; e.g., 80% follow-up)
...	...
3a	Systematic review (with homogeneity*) of case-control studies
3b	Individual case-control study
...	...

Level of Evidence

*By homogeneity we mean a systematic review that is free of worrisome variations (heterogeneity) in the directions and degrees of results between individual studies. Not all systematic reviews with statistically significant heterogeneity need be worrisome, and not all worrisome heterogeneity need be statistically significant. Studies displaying worrisome heterogeneity should be tagged with a "-" at the end of their designated level.