BOND UNIVERSITY
BRINGING AMBITION TO LIFE

Centre for Research in Evidence-Based Practice
CREBP

# Meta-analysis when the numbers get tricky

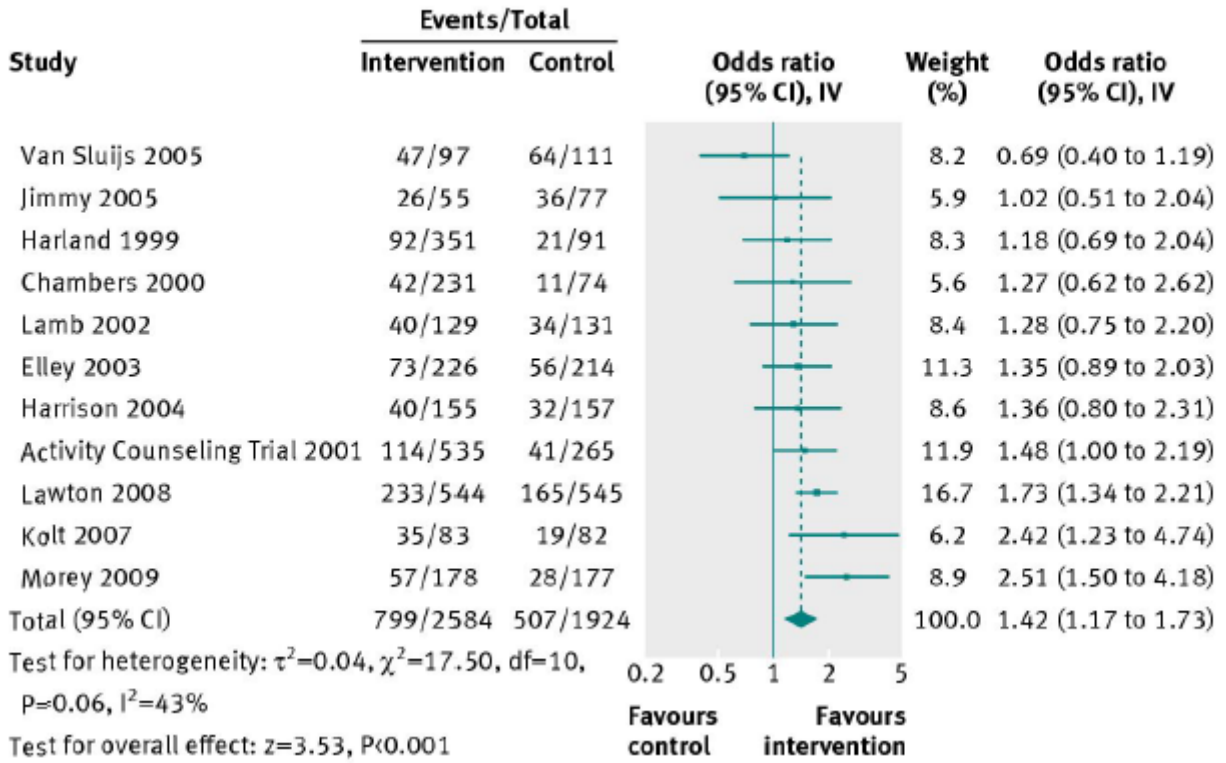
Elaine Beller

# Outline

- recap “standard” meta-analysis methods
  - effect measures for dichotomous data
  - for continuous data
- some examples of other types of data
- ways to handle these
- what we DON’T know how to handle
- summary

# Dichotomous data

Proportion of participants achieving physical activity target



| Study | Events/Total Intervention | Events/Total Control | Odds ratio (95% CI), IV | Weight (%) | Odds ratio (95% CI), IV |
|---|---|---|---|---|---|
| Van Sluijs 2005 | 47/97 | 64/111 | | 8.2 | 0.69 (0.40 to 1.19) |
| Jimmy 2005 | 26/55 | 36/77 | | 5.9 | 1.02 (0.51 to 2.04) |
| Harland 1999 | 92/351 | 21/91 | | 8.3 | 1.18 (0.69 to 2.04) |
| Chambers 2000 | 42/231 | 11/74 | | 5.6 | 1.27 (0.62 to 2.62) |
| Lamb 2002 | 40/129 | 34/131 | | 8.4 | 1.28 (0.75 to 2.20) |
| Elley 2003 | 73/226 | 56/214 | | 11.3 | 1.35 (0.89 to 2.03) |
| Harrison 2004 | 40/155 | 32/157 | | 8.6 | 1.36 (0.80 to 2.31) |
| Activity Counseling Trial 2001 | 114/535 | 41/265 | | 11.9 | 1.48 (1.00 to 2.19) |
| Lawton 2008 | 233/544 | 165/545 | | 16.7 | 1.73 (1.34 to 2.21) |
| Kolt 2007 | 35/83 | 19/82 | | 6.2 | 2.42 (1.23 to 4.74) |
| Morey 2009 | 57/178 | 28/177 | | 8.9 | 2.51 (1.50 to 4.18) |
| Total (95% CI) | 799/2584 | 507/1924 | | 100.0 | 1.42 (1.17 to 1.73) |

Test for heterogeneity: $\tau^2=0.04$, $\chi^2=17.50$, df=10, P=0.06, $I^2=43\%$

Test for overall effect: z=3.53, P<0.001

# Effect measures for dichotomous data

- Risk ratio

$$\frac{Proportion\ with\ event\ in\ intervention}{Proportion\ with\ event\ in\ control}$$

- Odds ratio

$$\frac{Events\ in\ intervention \Big/ No\ events\ in\ intervention}{Events\ in\ control \Big/ No\ events\ in\ control}$$

- Risk difference

$$Proportion\ with\ event\ in\ intervention - Proportion\ with\ event\ in\ control$$

# Effect measures for continuous data

- mean difference
  - need mean and SD in each group
  - assumes approximately normally-distributed
- standardised mean difference
  - different measurement scales

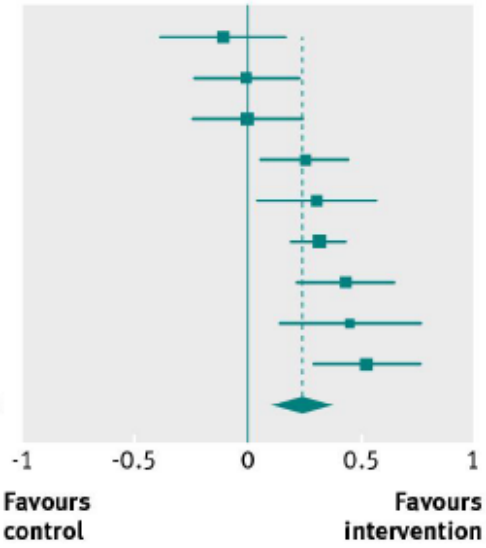
# Continuous data

**Fig 4** Individual study and pooled effects of physical activity promotion on self reported physical activity at 12 months (continuous data). Random effects model used. SD=standard deviation; 95% CI=95% confidence intervals; IV=inverse variance

| Study | Intervention Mean (SD) | Intervention Total | Control Mean (SD) | Control Total | Weight (%) | Standardised mean difference (95% CI), IV |
|---|---|---|---|---|---|---|
| Van Sluijs 2005 | 484.6 (504.7) | 97 | 539.1 (510.1) | 111 | 9.7 | -0.11 (-0.38 to 0.17) |
| Kinmonth 2008 | 101.23 (48.96) | 214 | 101.4 (58.6) | 110 | 11.1 | 0.00 (-0.23 to 0.23) |
| Lamb 2002 | 60 (133.33) | 131 | 60 (133.33) | 129 | 10.6 | 0.00 (-0.24 to 0.24) |
| Elley 2003 | 247.26 (42.2) | 226 | 236.07 (45.3) | 214 | 12.4 | 0.26 (0.07 to 0.44) |
| Chambers 2000 | 2.61 (12.45) | 231 | -1 (9.2) | 74 | 10.0 | 0.31 (0.04 to 0.57) |
| Lawton 2008 | 120 (155.56) | 544 | 75 (125.93) | 545 | 14.6 | 0.32 (0.20 to 0.44) |
| Morey 2009 | 126.6 (142.9) | 178 | 69.6 (116.1) | 177 | 11.7 | 0.44 (0.23 to 0.65) |
| Kolt 2007 | 244 (365.7) | 83 | 117.3 (138.8) | 82 | 8.7 | 0.45 (0.15 to 0.76) |
| Halbert 2000 | 3 (2.22) | 149 | 2 (1.48) | 150 | 11.0 | 0.53 (0.30 to 0.76) |
| Total (95% CI) | | 1853 | | 1592 | 100.0 | 0.25 (0.11 to 0.38) |

Test for heterogeneity: $\tau^2$=0.03, $\chi^2$=26.84, df=8, P<0.001, $I^2$=70%

Test for overall effect: z=3.66, P<0.001



# What do we need from the study report?

- measure of effect in each group (e.g. mean, proportion of events)
  - with estimate of variability (e.g. SD, SE, 95%CI)

OR

- ratio of effects in the groups (for dichotomous)
  - with variability (e.g. 95% CI)
- difference in effects (for continuous)
  - with variability (e.g. 95% CI)

# Other data types

- time to event
- multiple episodes of the event
- cluster and crossover randomised trials

# What can we do here?

Congestive Heart Failure

## Effects of carvedilol early after myocardial infarction: Analysis of the first 30 days in Carvedilol Post-Infarct Survival Control in Left Ventricular Dysfunction (CAPRICORN)

Gregg C. Fonarow, MD,[a] Mary Ann Lukas, MD,[b] Michele Robertson, BSc,[c] Wilson S. Colucci, MD,[d] and Henry J. Dargie, MD[e] *Los Angeles, CA; Philadelphia, PA; Glasgow, United Kingdom; and Boston, MA*

*death*, or *death*. A composite end point of death or nonfatal MI, which was an end point used in the primary CAPRICORN study,[6] was also evaluated in this analysis. We also assessed the

Cumulative incidence curves for the occurrence of these events were constructed by the Kaplan-Meier method, using a time-to-first-event approach. The analyses of major outcome

# Time-to-event - possibilities

- use hazard ratio (and 95% CI)
  - similar interpretation to risk ratio

OR

- convert to dichotomous at a time point
  - e.g. proportion surviving at 36 months
  - (can read from survival curves if necessary)

BUT

- do NOT use mean time to event

# Example – time-to-event

**Table 3. Adjusted Risk (Hazard Ratios) for Primary and Secondary Outcomes in HF Patients Within CKD and Non-CKD Groups, Based on Treatment With Carvedilol**

| | CKD, eGFR ≤60 mL/min/1.73 $m^2$ (Carvedilol Versus Placebo) (n=1293 Versus 1273) | | Non-CKD, eGFR >60 mL/min/1.73 $m^2$ (Carvedilol Versus Placebo) (n=822 Versus 829) | |
|---|---|---|---|---|
| Outcome | HR* | 95% CI | HR* | 95% CI |
| All-cause mortality | 0.76 | 0.63–0.93 | 0.59 | 0.43–0.81 |
| Cardiovascular mortality | 0.77 | 0.62–0.94 | 0.59 | 0.42–0.82 |
| HF mortality | 0.68 | 0.52–0.88 | 0.58 | 0.34–0.99 |
| First hospitalization for HF | 0.74 | 0.62–0.88 | 0.83 | 0.63–1.09 |
| Composite of cardiovascular mortality or hospitalization for HF | 0.75 | 0.65–0.87 | 0.78 | 0.63–0.98 |
| Sudden cardiac death | 0.76 | 0.56–1.05 | 0.58 | 0.37–0.92 |

*HR is based on the Cox model after adjusting for treatment arm and study type.
There was no significant interaction of treatment and CKD/non-CKD for any of the outcomes.

# What can we do here?

| | All Exercisers (N = 53) | All Stretchers (N = 62) |
|---|---|---|
| Number of colds over 12 mo | n (%) | n (%) |
| 0 | 26 (49) | 19 (31) |
| 1 | 11 (21) | 15 (24) |
| 2 | 3 (6) | 13 (21) |
| 3 | 2 (4) | 2 (3) |
| Number of colds between | RR (95% CI) | |
| 0-3 mo | 0.62 (0.24-1.59) | |
| 3-6 mo | 1.26 (0.56-2.86) | |
| 6-9 mo | 0.51 (0.20-1.28) | |
| 9-12 mo | 0.32 (0.13-0.81) | |
| *P* value† | 0.02 | |
| Number of upper respiratory tract infections†‡ over 12 mo | n (%) | n (%) |
| 0 | 14 (26) | 11 (18) |
| 1 | 9 (17) | 18 (29) |
| 2 | 11 (21) | 11 (18) |
| 3 | 1 (2) | 6 (10) |
| 4 | 5 (9) | 3 (5) |
| 5 | 1 (2) | 0 (0) |
| 6 | 1 (2) | 0 (0) |
| Number of upper respiratory tract infections†‡ between | RR (95% CI) | |
| 0-3 mo | 0.95 (0.49-1.86) | |
| 3-6 mo | 2.19 (1.04-4.61) | |
| 6-9 mo | 0.94 (0.58-1.55) | |
| 9-12 mo | 0.71 (0.39-1.28) | |
| *P* value† | 0.16 | |

# Multiple episodes - possibilities

- use rate ratio (and 95% CI)
  - similar interpretation to risk ratio
- calculate rate ratio
  - need follow-up time (average in each group)
  - assumes constant risk over time
  - estimate standard error using number of events

OR

- use mean number of episodes if common event
  - methods for continuous data
  - need standard deviation, 95%CI etc.

# What can we do here?

**Table 5** Incidence and Duration of URTI *.

| Variable | EG (N = 16) | CG (N = 16) |
|---|---|---|
| URS (No.) | 4 | 4 |
| duration of incidence (days) | 5.3 ± 1.5 | 6.3 ± 2.2 |

M ± SD

*p > 0.05

# What can we do here?

BMJ

BMJ 2012;344:d8173 doi: 10.1136/bmj.d8173 (Published 2 February 2012) Page 1 of 13

RESEARCH

## Effectiveness of multifaceted educational programme to reduce antibiotic dispensing in primary care: practice based randomised controlled trial



**Table 3** | Effects of interventions on antibiotic prescribing at index consultation and antibiotic prescribing and reconsultation during 28 days' follow-up

| Variables | Intervention groups | | Control groups | | P value† | Intracluster coefficient |
|---|---|---|---|---|---|---|
| | No of patients | Percentage (crude 95% CI*) | No of patients | Percentage (crude 95% CI*) | | |
| C reactive protein test: | n=227 | | n=204 | | | |
| Antibiotics at index consultation | 70 | 30.8 (21.8 to 39.8) | 108 | 52.9 (43.0 to 62.8) | 0.02 | 0.12 |
| Antibiotics at days 1 to 28 | 102 | 44.9 (35.2 to 54.6) | 119 | 58.3 (48.5 to 68.1) | <0.01 | 0.12 |
| Reconsultation within 28 days | 79 | 34.8 (28.3 to 41.3) | 62 | 30.4 (23.8 to 37.0) | 0.50 | 0.01 |
| Communication skills training: | n=201 | | n=230 | | | |
| Antibiotics at index consultation | 55 | 27.4 (25.6 to 36.6) | 123 | 53.5 (43.8 to 63.2) | <0.01 | 0.12 |
| Antibiotics at days 1 to 28 | 76 | 37.8 (28.1 to 47.5) | 145 | 63 (53.6 to 72.4) | <0.001 | 0.12 |
| Reconsultation within 28 days | 56 | 27.9 (21.4 to 34.4) | 85 | 37.0 (30.4 to 43.6) | 0.14 | 0.01 |

*Calculated and inflated for clustering by using standard deviation inflated by variance inflation factor.[53]
†Calculated from second order penalised quasi-likelihood multilevel logistic regression model adjusted for variance at general practitioner and practice level (random intercept at practice and general practitioner level). Models included both interventions and interaction term of interventions. See web extra for corresponding β coefficients.

# Including Cluster RCTs

- design effect:

$$1 + (M - 1)ICC$$

M=average cluster size, ICC=intraclass correlation coefficient

- can "borrow" ICC from similar study
- effective sample size:

$$\frac{sample\ size}{design\ effect}$$

- dichotomous outcome: divide events and totals by design effect (may need to round to whole no.)
- continuous outcome: divide totals by design effect

# What can we do here?

Table 3| **Re-consultation rates for respiratory tract infections within seven, 14, and 31 days in practices undergoing educational programme aimed at reducing antibiotic prescribing in primary care**

| | Median (IQR) | | | |
|---|---|---|---|---|
| **Re-consultation rate*** | **Intervention (n=20)†** | **Control (n=17)†** | **Median difference (95% CI‡)** | **P value§** |
| Within 7 days | 2.66 (1.88-4.25) | 3.35 (2.16-4.31) | −0.65 (−1.69 to 0.55) | 0.446 |
| Within 14 days | 5.10 (4.70-7.92) | 6.43 (4.04-7.84) | −1.33 (−2.12 to 0.74) | 0.411 |
| Within 31 days | 9.06 (7.53-12.62) | 11.38 (7.39-14.05) | −2.32 (−4.76 to 1.95) | 0.503 |

IQR=interquartile range.

*Median No of individuals who reconsulted for respiratory tract infection per 1000 registered patients.

†Values in each group refer to subset of intervention practices for which data on reconsultation were available.

‡Computed with bootstrapping methods.

§From Mann-Whitney U test.

# What can't we meta-analyse?

- Medians
  - IQR is not a 'statistical' measure of variability
  - i.e. can't be used for weighting studies in MA

# Summary

- often possible to use measures other than proportion and mean
- need to understand ‘generic inverse variance’
- generally ln(measure) and ln(SE)
  - (natural logarithm)

- consult a statistician
- don’t discard ‘weird’ data immediately!