SELECTING THE APPROPRIATE EFFECT SIZE FOR NETWORK META-ANALYSIS

Overview of empirical evidence

INTRODUCTION

- Dichotomous data is the most frequent type of data in network meta-analyses
  - Database including 186 networks:
    - 60% with dichotomous data
    - 28% with continuous data
    - 9% and 3% with survival and rate data

    - Effect measures used in dichotomous networks:

    - 59% OR
    - 40% RR
    - 0% RD
    - 1 network used all three effect measures [Ballesteros 2005]

    We could not find any rationale provided for the choice of the effect measure!

DICHOTOMOUS EFFECT MEASURES IN META-ANALYSIS

- Choosing an effect measure (OR, RR, RD) based on:
  - the mathematical properties
  - the level of interpretability
    - Do the analysis in one measure, and transform it into another!
    - Focus here is measure of analysis, not measure of presentation (see Stream 3)
- Heterogeneity
  - RD more heterogeneous than RR and OR [Deeks et al. 2002]
- Remember that RR…
  - RRb & RRh can give different results (magnitude of effect and precision)

IMPLICATIONS FOR INDIRECT COMPARISONS

- Because RRb and RRh can give different results
  - Indirect effects using RRb and RRh can differ in magnitude & direction of effect! [Eckerman et al. 2009]
    - two illustrative examples of indirect comparisons:

    - RR(no stroke) warfarin vs. aspirin – warfarin marginally more effective
      RR(stroke)– 56% reduction in risk of stroke with warfarin
    - **RR(no progression) natalizumab vs. interferon – natalizumab 16% less effective**
      RR(progression) – natalizumab 30% more effective

- Inconsistency
  - No important 'a priori' differences between the different measures (loop-specific & design-by-treatment approach) [Veroniki et al. 2013]
    - empirical data on 40 networks of trials

- Treatment ranking
  - The three effect measures (OR, RR, RD) can give different results [Norton et al. 2012]
    - Graphical presentation & mathematical proof of the issue
  - This study prompted a reply…

HOW TO SELECT THE APPROPRIATE MEASURE?

- Exchangeability & additivity of treatment effects
  - The assumptions cannot hold for all measures simultaneously
    [van Valkenhoef & Ades 2013]

- Choice of effect measure should be based
  - not on convenience and interpretation criteria
  - on scientific grounds; heterogeneity and goodness-of-fit measures [Caldwell et al. 2012]

- Scale of analysis is
  - specific for each dataset
    - greater consideration on HR for time-to-event data
    - OR sometimes gives larger effects and can be misinterpreted
  - a different issue than the scale of reporting
    [Caldwell et al. 2012]

More details and discussion (see Tony Ades' presentation that follows)

REFERENCES

- Ballesteros J: Orphan comparisons and indirect meta-analysis: a case study on antidepressant efficacy in dysthymia comparing tricyclic antidepressants, selective serotonin reuptake inhibitors, and monoamine oxidase inhibitors by using general linear models. *J Clin Psychopharmacol* 2005, 25: 127-131.
- Caldwell DM, Welton NJ, Dias S, Ades AE: Selecting the best scale for measuring treatment effect in a network meta-analysis: a case study in childhood nocturnal enuresis. *Res Synth Meth* 2012, 3: 126-141.
- Deeks JJ: Issues in the selection of a summary statistic for meta-analysis of clinical trials with binary outcomes. *Stat Med* 2002, 21: 1575-1600.
- Eckermann S, Coory M, Willan AR: Indirect comparison: relative risk fallacies and odds solution. *J Clin Epidemiol* 2009, 62: 1031-1036.
- Norton EC, Miller MM, Wang JJ, Coyne K, Kleinman LC: Rank reversal in indirect comparisons. *Value Health* 2012, 15: 1137-1140.
- van Valkenhoef G, Ades AE: Evidence synthesis assumes additivity on the scale of measurement: response to "Rank reversal in indirect comparisons" by Norton et al. *Value Health* 2013, 16: 449-451.
- Veroniki AA, Vasiliadis HS, Higgins JP, Salanti G: Evaluation of inconsistency in networks of interventions. *Int J Epidemiol* 2013, 42: 332-345.