Department of Health Sciences

M.Sc. Module: Systematic Reviews

Suggested answers to exercise on meta-analysis: house dust mite control

1. *Comment on the review question (Methods first paragraph) in terms of the patient or problem, intervention, comparison treatment, and outcomes.* The patients are defined as 'patients with asthma who were sensitised to house dust mites'. The authors specified that 'Asthma had to have been diagnosed by a doctor and sensitisation to mites had to have been assessed by skin tests, bronchial provocation tests, or serum assays for specific IgE anti-bodies.' The problem is well-defined, at the risk of excluding studies of mite removal where anti-body tests had not been performed. The intervention was any 'measure designed to reduce their exposure to mite antigen in the home'. The method of removing the mites is not important, nor whether they were successful in removing mites. The comparison is no treatment. No control treatment or placebo is specified. No outcome is specified. We are told only that 'effects on patients with bronchial asthma' so presumably any outcome measurement will be included. Hence although the patient and problem is precisely specified, intervention and comparison treatment are very loosely specified, and outcome is not specified at all.

2 *Comment on the search strategy. How does this compare with current recommendations on identifying clinical trials?* They searched only four data bases, rather a small number, although they appear to those likely to contain the trials they seek. They searched two journals by hand, though they do not say why these particular journals were chosen. They do not appear to have looked for unpublished trials or grey literature. Although their search strategy for trials looks fairly sensitive, it does not contain any publication types or MeSH headings (e.g. randomised controlled trial as publication type). The Cochrane Library provides more comprehensive standard searches. A good student might point out that this paper was published in 1998, so these may not have been available then.

3. *Two authors each identified trials and extracted data, then discussed any ambiguities. Why did they do this?* Identifying trials and extracting information from papers is subjective. Judgements have to made as to whether trials match the criteria and how the trial was done. Different investigators may come to different decisions. Researchers may be biased. Also, simple human error may lead to data been transcribed incorrectly. Having two authors identify trials and extract data helps us avoid errors in trial selection and data extraction. Ideally, they should present information on their agreement.

4 *In the methods we read that 'Two [of the authors] (PCG and CH) extracted data on the following outcomes: subjective wellbeing, improvement in asthma symptoms, use of drugs to control asthma, number of days of sick leave taken from school or work, number of unscheduled visits made to a doctor or hospital, forced expiratory volume in 1 second, peak expiratory flow rate, provocative concentration that causes a 20% fall in forced expiratory volume in 1 second, and results of skin prick testing.' What other kinds of information would you expect them to extract?* They should collect information about the quality of the study, such as randomisation and blinding, the definition of the patients, and about the intervention and control treatments used.

5 *The authors calculated the standardised mean difference in the analysis of some data (Statistical methods). What does this mean and why did they do it?* The standardised mean difference is the difference between the means of the observations in the two treatment groups, or the mean difference between treatments in a paired trial, divided by the standard deviation of the measurements. This is done when we want to combine or compare treatment differences for different variables, measured in different units. Here they had peak expiratory flow presented in litres/minute in some papers and as a percentage of that expected in others. To combine them for meta-analysis each was divided by the standard deviation so that they were both in standard deviation units.

6. *What kind of graph is Figure 1? What do the squares and horizontal lines mean? Why are the squares of different sizes?* This is a forest plot. The squares represent the point estimate of the odds ratios for the individual studies. The horizontal lines represent the confidence intervals for these estimates. The squares are of different sizes because studies which provide more information carry more weight in the analysis. The area of the square should be proportional to the weight. Hence the square for Carswell, which has very few patients improving and so a low weight, is smaller than the others. If the squares were all shown of equal size, the smaller studies with wider confidence intervals would have more visual impact than larger studies with narrow confidence intervals and the picture would be misleading.

7 *What kind of horizontal scale (x axis) is used in Figure 1? Why is this?* This is a logarithmic scale. It is used because the treatment difference is expressed in the form of an odds ratio. Confidence intervals for odds ratios are calculated for the log odds ratio and then anti-logged. Using a logarithmic scale makes them appear symmetrical on the plot. If we reverse the order of the treatments, the odds ratio becomes the reciprocal. For the log odds ratio this only changes the sign of the treatment effect. This means that equal but opposite treatment effects are represented by the same distance on the graph when a logarithmic scale is used.

8 *In Figure 2, what is 'χ^2, df=6, z=0.27' testing?* This is testing two things: heterogeneity and the pooled treatment difference. The 'χ^2' is a chi-squared test of the null hypothesis that there is no heterogeneity between the trials. If there is no heterogeneity, we expect this to be equal to the degrees of freedom. In Figure 2, $\chi^2 = 17.84$, which is much greater than the 6 degrees of freedom, so there is good evidence for heterogeneity ($P<0.01$, not $P<0.0001$ as in the paper, and if any student spotted that give them full marks!). The 'z = 0.27' is a test of the null hypothesis that the treatment difference in the whole population is zero. This is not significant.

9 *In Figure 2, what do the diamond or lozenge shapes represent? What suggests that they have been drawn incorrectly?* They represent the pooled meta-analysis estimates of the treatment effect. The deepest point represents the point estimate and the width of the diamond represents the confidence interval. They appear incorrect because for the first two meta-analysis estimates the confidence interval includes zero, but the diamonds do not cross the zero line.

10 *The authors say that 'If P<0.10 in the test for heterogeneity a random effects analysis was carried out' (Statistical methods). What is heterogeneity? What is 'a random effects analysis'? Why is P<0.10 taken as the decision point, rather than P<0.05?* Heterogeneity means that the trials are not all estimating the same treatment effect, but that the treatment effect varies from trial to trial. This may be clinical heterogeneity, where trials vary in patient definition, intervention, etc., or statistical heterogeneity, where for whatever reason trials vary more than we would expect given their sample sizes and the variability between patients. A random effects analysis is one which takes into account heterogeneity by not assuming that the treatment effect is uniform and estimating the variance between trials. This variance is used in calculating the weights. Small trials get relatively greater weight in a random effects analysis than in a fixed effects analysis, confidence intervals are wider and P values are larger. 0.10 is used as the cut-off for the test of heterogeneity because this test has low power to detect heterogeneity, particularly in meta-analyses with small numbers of trials. Increasing the P value increases the power of the test at the cost of increasing the Type I error rate.