Understanding

Evidence Update

A Short Guide

December 2006

Produced by:

Effective Health Care Research Programme Consortium

If you have any comments or suggestions relating to this guide, please email them to Katharine Abba at the Effective Health Care Research Programme Consortium:

K.Abba@liverpool.ac.uk

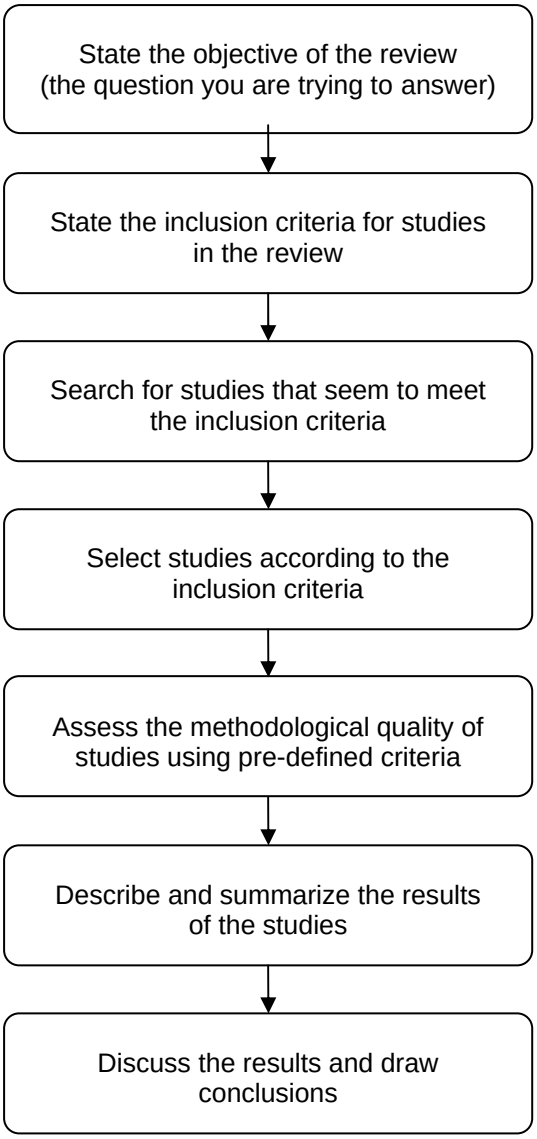
Introduction

Each *Evidence Update* is a two page summary of a Cochrane review. This information sheet describes what a Cochrane Review is, and how to read and interpret *Evidence Update*.

What is a Cochrane Review?

Cochrane Reviews assess available evidence on the effectiveness of interventions in health care and public health. They are designed to answer specific questions, such as whether one drug treatment is more effective than another for treating a defined illness.

All Cochrane Reviews are *systematic reviews*. A systematic review is a review of the methods and results of *all* the individual studies designed to answer a specific question and conforming to set criteria. The process for undertaking any systematic review is summarized below:



If two or more studies with similar outcome measures are included in a systematic review, their results may sometimes be combined (or pooled) statistically using a process called *meta-analysis*. Some reviews may also compare analyses for different subgroups within studies (for example, men and women or people of different age groups).

The discussion and conclusions of a systematic review take into consideration the quality of included studies, the likely impact of bias and chance on the results, and the applicability of the findings to different groups and settings.

What is a randomized controlled trial?

Cochrane Reviews often include only *randomized controlled trials.* This is an experimental design in which participants are allocated to two or more groups at random. One group (the control group) receives no treatment, a placebo, or the old or usual treatment, while the other group(s) (the intervention or treatment groups) receive the intervention(s) being evaluated. Random allocation of participants to intervention or control groups is designed to ensure that the groups are similar in all respects except the intervention they receive, so that differences in outcome can be attributed to the intervention only.

Another kind of study often included in Cochrane Reviews is the *quasi-randomized controlled trial.* These are similar to randomized controlled trials, but participants are not allocated to groups truly randomly. For example, a group may include every second person presenting at a clinic, or every person whose birthday falls on an odd day of the month.

What is Evidence Update?

Each *Evidence Update* is a two page summary of a Cochrane Review that includes the methods and findings. *Evidence Updates* are produced for reviews that are particularly relevant to people in low and middle-income countries. The *Evidence Update* series is available to download at: http://www.liv.ac.uk/evidence/evidenceupdate/home.htm

***Evidence Update* is structured using the following headings:**

- Review question
- Statement conclusion
- Inclusion criteria
    - Types of studies
    - Types of participants
    - Interventions
    - Outcomes
- Results
- Authors conclusions
    - Implications for practice
    - Implications for research

These headings cover:

Review question:

The question the review was designed to answer

Statement conclusion:

The main or most important conclusions of the review

Inclusion criteria:

The characteristics of studies eligible to for inclusion in the review, as specified by the authors at the beginning of the review process.

Types of studies included in the review are selected so that only studies using the most objective research methods that are practical to answer the question are included. Often this means randomized controlled trials only. Some reviews also include quasi-randomized controlled trials, controlled before-and-after-studies, or interrupted time series analyses.

The *types of participants, interventions,* and *outcomes* are selected so that the included studies are designed to answer the specific research question of the review. In the case of the intervention, the comparison or control group is also stated.

All identified studies that match these inclusion criteria are eligible for the review.

Results:

A summary of the main findings of the review, presented as bullet points:

- The first bullet point usually describes the number of studies included, the total number of participants, and other relevant information such as study locations. If randomized controlled trials are included, there is a statement on the number of trials that had adequate *allocation concealment*, as an indicator of their methodological quality. Allocation concealment is the process of shielding those involved in the study from knowing upcoming group assignments, to prevent selection bias caused by intentional changes in which participant gets the next assignment.
- Further bullet points describe the most important findings of the review. Where a study, or the combined results of more than one study, detects a significant difference between two interventions, summary statistics are presented in brackets. These include the point estimate of the effect size, the 95% confidence interval around the estimate, the number of trials, and number of participants included in the analysis.

Where appropriate, meta-analyses for the main results are also presented as a graph.

(A guide to interpreting the summary statistics and graphs is presented in the next section).

Authors' conclusions:

Implications for practice describe whether there was a significant effect of intervention compared with the control or comparison treatment, which specific groups this effect relates to, and sometimes how the findings might be applicable in the context of current practice and situations.

Implications for research highlight areas where evidence is lacking and new research might be useful.

Interpreting the results presented in Evidence Update

Measures of difference between groups: effect sizes

The results of individual trials, or the combined results of two or more trials comparing the same interventions, are summarized using a point estimate of the ***effect size*** and the ***95% confidence interval*** *(95% CI)* around this point estimate. Effect sizes compare results in two comparison groups (usually the intervention and control groups) in relation to a specific outcome.

For continuous outcomes (such as height or weight), the effect size is usually presented as the ***mean difference***, or ***standardized mean difference (SMD)*** between groups. It is expressed in the units in which it was measured (for example, kilogrammes or millimetres). The 95% CI around a mean difference or SMD is the range in which you can be 95% confident that the true value lies. Mean differences or SMDs are only considered *statistically significant* where the 95% CI does not include 0.

For events outcomes (such as deaths or treatment failures), the effect size is usually presented as ***relative risk (RR)*** or ***odds ratio (OR).*** Relative risk is the risk of a stated outcome for the intervention group compared with (divided by) the control group. The odds ratio is the odds of a stated outcome in the intervention group compared with the odds of the same event in the control group. Odds are the number of events divided by the number of non-events. If the RR or OR is greater than 1, the intervention group has a higher risk than the control group. If the RR or OR is less than 1, the intervention group has a lower risk. The 95% CI around an RR or OR is the range in which you can be 95% confident that the true value lies. An RR or OR can only be said to be *statistically significant* if the 95% CI does not include 1.

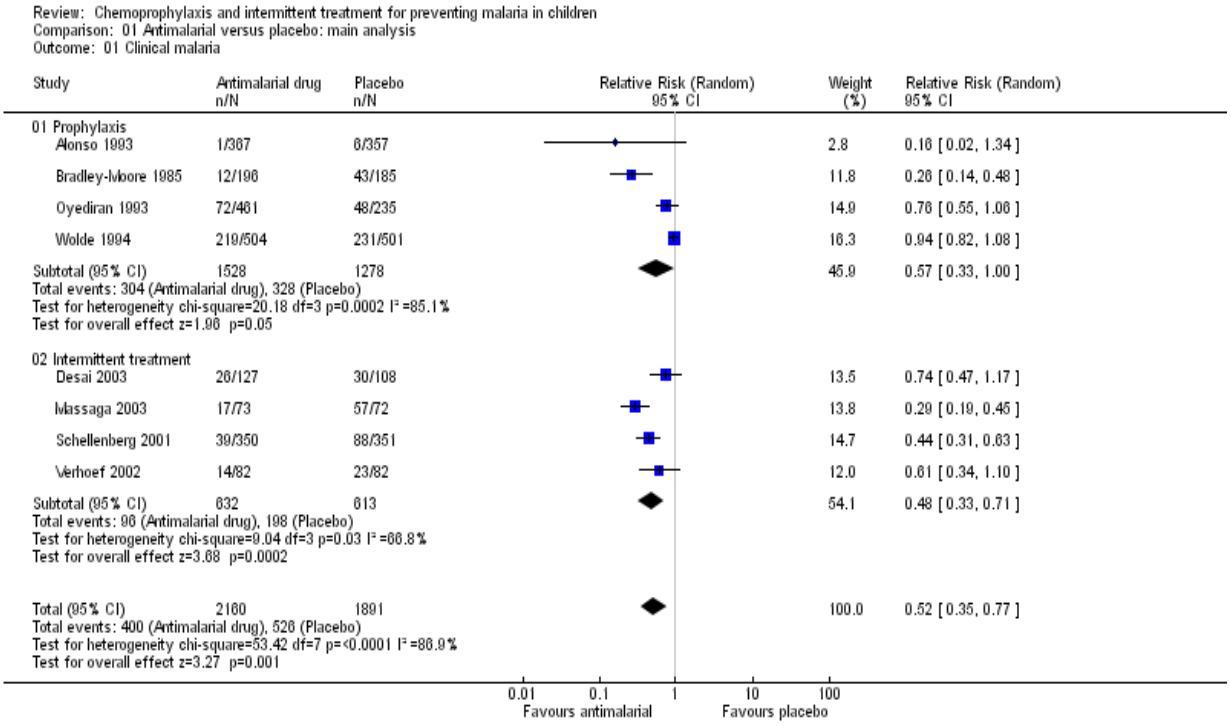
Statistically significant differences are not the same as *clinically significant* differences; a very small clinical effect can be statistically significant when observed in a large number of people.

Combining the results of studies (meta-analysis)

If visual inspection of the data and the results of statistical tests indicate that two or more studies in the review seem to be providing estimates of the same effect, their results may be combined to give an overall estimate of the intervention's effect size (RR, OR, or SMD), and its 95% CI. It is a summary of the results of all the trials, with more weight given to more precise results (usually from the larger trials). The pooled estimate is always more precise than the individual trial estimates, and it sometimes detects a significant effect that was not apparent from the results of individual studies.

The meta-analysis

The meta-analysis sometimes called a 'forest plot' or 'blobogram'. The example below is from a review of chemoprophylaxis and intermittent treatment with antimalarial drugs for preventing malaria in children, comparing the number of children who developed clinical malaria in groups who received antimalarial drugs (the intervention group) with those who received placebo (the control group).



The included trials are listed down the left hand side. The next two columns show, for each trial, the number of children in the antimalarial and placebo groups who had clinical malaria during the trials' follow-up period, and total number of children in the group. The horizontal lines in the centre indicate the *treatment effect* of intervention compared with control for each trial (read from the log scale along the bottom). In this case it is a relative risk (RR); in other reviews with different types of outcomes it may be an odds ratio (OR) or standardized mean difference (SMD). ORs and RRs usually (but not always) apply to an undesired outcome, in this case malaria. For an undesired outcome, an OR or RR less than 1 favours the intervention, an OR or RR greater than 1 favours the control. The vertical line down the middle is the *line of no effect* (OR = 1, RR = 1 or SMD = 0) where intervention and control have exactly the same outcomes.

The centre square on each horizontal line represents the *point estimate* of the effect size. The bigger the square, the more precise the point estimate and the more weight is given to the results of that trial. The line itself shows the *95% confidence interval* around the point estimate. Differences between intervention and control are *statistically significant* if the confidence interval does not cross the line of no effect.

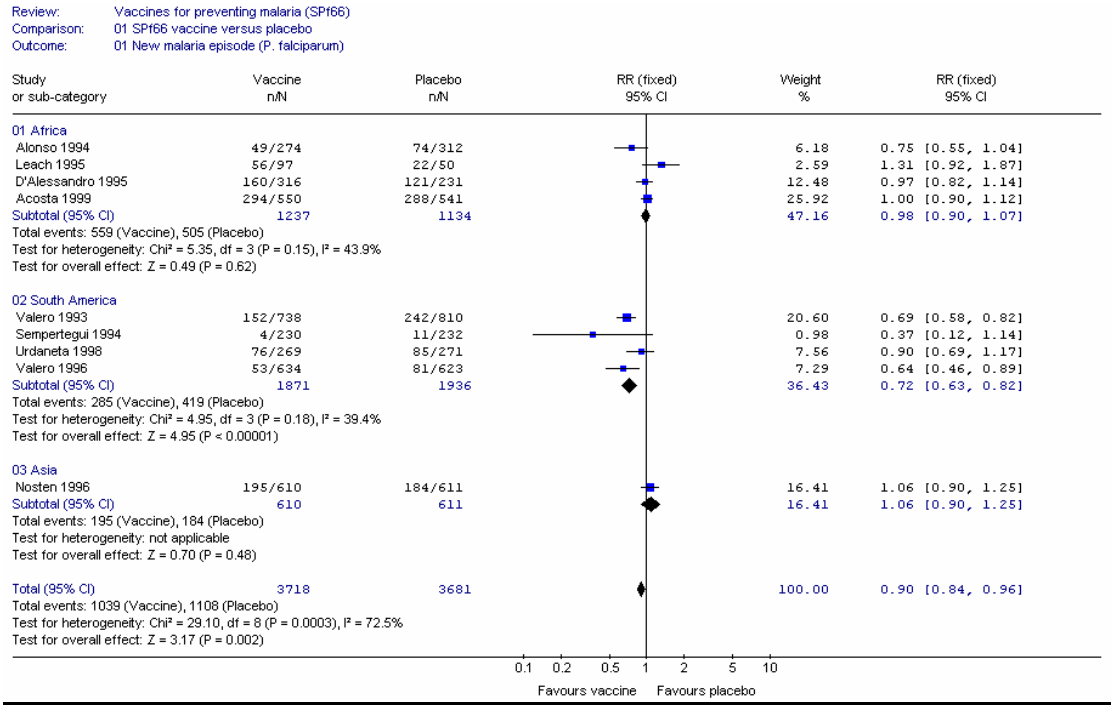
The small diamond at the bottom indicates the combined relative risk and 95% CI for all the trials included. In this case, the combined results show a significant benefit for the intervention, which was not apparent in all of the individual trials. This example also contains separate analyses for chemoprophylaxis and intermittent treatment; these were combined together because their results were similar.

The two columns on the right of the graph list the weight given to each trial in the meta-analysis, and the point estimate and 95% CI of the effect size.

The example above shows a consistent positive effect for the intervention. The examples below show what the meta-analysis looks like when there is an inconsistent effect across studies, no significant effect of intervention, and where there is not enough evidence to say whether there is a significant effect or not.

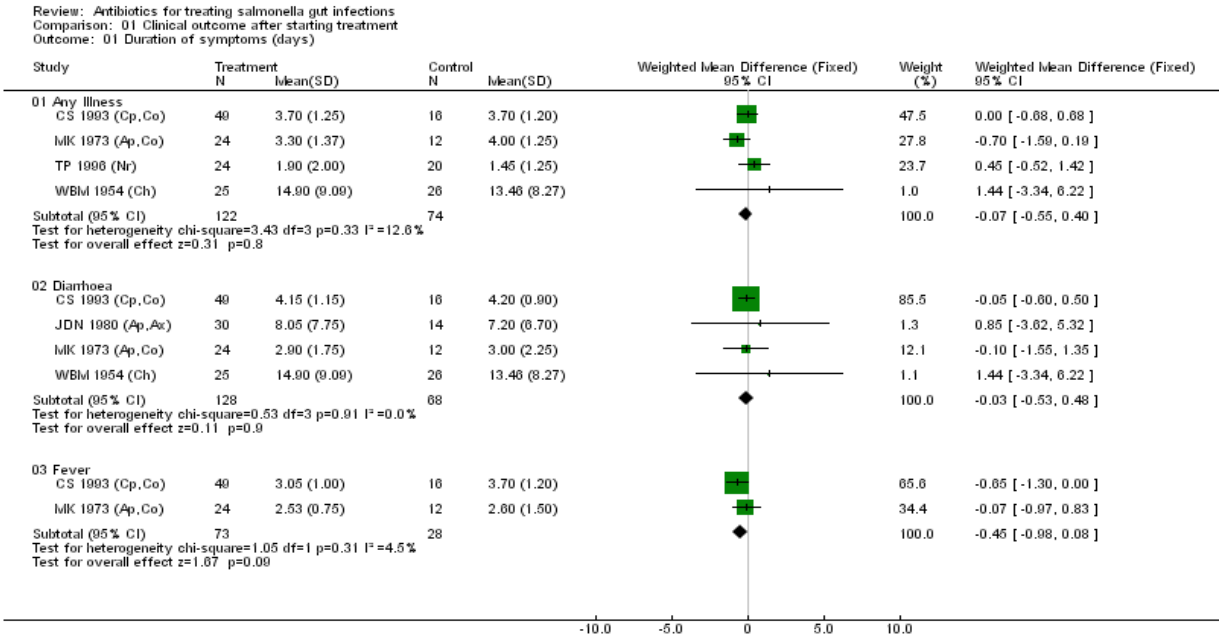
Inconsistent effects across studies

Nine trials compared SPf66 vaccine against placebo for preventing new episodes of *Plasmodium falciparum* malaria. The analysis was stratified by continent. Trials in South America showed an apparent benefit of the vaccine, while those in Africa and Asia showed no benefit.



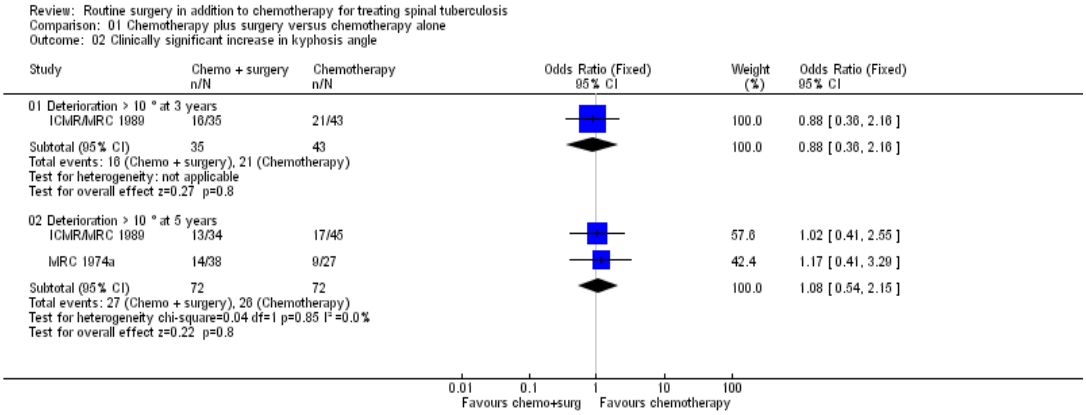
No significant effects

Five trials compared duration of fever, diarrhoea, or any illness in people with salmonella gut infections who were treated with antibiotics or who were not treated. It is apparent that if there was any effect of antibiotics on duration of illness, it was very small, and therefore probably unimportant.



Not enough evidence to say whether or not there is an effect

Three trials reported on increase > 10^{o} in kyphosis (spinal curve angle) at follow up in people with spinal tuberculosis who were given routine surgery plus chemotherapy or chemotherapy only for spinal tuberculosis. One trial assessed this outcome after five years, another after three and five years. Results at three and five years follow up are presented separately; results from the two trials with five years follow up are combined. The 95% CIs at both three and five years cross the line of no effect, meaning that there was no *statistically significant* difference between the two groups. This may be because there was actually no effect of surgery on this outcome, or that there was an effect but that the sample sizes used in the trials were too small to detect it. The 95% CI around the odds ratio is 0.54 to 2.15, meaning that clinically significant effects favouring either surgery or no surgery cannot be ruled out using the available data.



Applying the findings of Evidence Update in practice

The findings of Cochrane Reviews, as presented in *Evidence Update* or elsewhere, represent the best available evidence on a particular intervention for a particular problem at a global level. However, the results need to be interpreted and applied within the local context, taking local situations into account. Techniques to facilitate the local implementation of research findings may include the agreement and production of local guidelines or policies, raising awareness among staff or patients, and processes of clinical audit.

APPENDIX

Statistical terms common in Cochrane Reviews

Allocation concealment – Those involved in the study are shielded from knowing upcoming group assignments, to prevent selection bias caused by intentional changes of who gets the next assignment.

Blinding – The allocation group of participants is hidden from the participants, healthcare staff, and/or study personnel. The study or review should describe exactly who was blinded. Blinding helps prevent bias in care received or observations made.

Confidence Interval (CI) – A 95% confidence interval is the range in which we can be 95% confident a true value lies, that is there is only a 5% chance that it lies outside this range. Cochrane reviews always use 95% CI, but trial reports may also occasionally use 90% or 98% CI. Confidence intervals can be calculated for many statistics, including standardized mean differences, relative risks and odds ratios.

Controls, control group – Trial participants who do not receive the intervention being tested, for comparison. Controls often receive *placebo* or alternative treatment**.**

Heterogeneity – Greater difference between studies in magnitude of results than could have been expected due to chance alone.

Mean, weighted mean, standardised mean – The average value of a continuous variable, such as weight or blood pressure. Weighted or standardised means combine means of different groups taking their different sizes into account, with a greater weight being given to larger groups.

Mean difference, weighted mean difference (WMD), standardised mean difference (SMD) – The difference between the means or standardised means of two groups or categories, for example, those receiving treatment and placebo.

Meta-analysis – The statistical combination of data from studies included in a systematic review.

Odds ratio (OR) – The odds of an event is the number of events divided by the number of non-events. The odds ratio is the odds of a stated undesirable outcome in the intervention group compared with the odds of the same event in the control group. As with RR, if the OR is greater than one, the intervention group has a higher risk than the control group and if the OR is less than 1, the intervention group have a lower risk.

Quasi-randomization – Group allocation that is not random but still has a low probability of introducing bias, for example, by odd or even numbered dates of birth.

Randomization – Participants are allocated to different study intervention groups using an unpredictable, random method, such as coin tossing or computer-generated random sequences. This prevents selection bias in the allocation.

Relative risk (RR) – The risk of a stated undesirable outcome for the intervention group compared with the control group. If the RR is greater than one, the intervention group has a higher risk than the control group (the intervention is harmful). If the RR is less than 1, the intervention group has a lower risk (the intervention is beneficial).