Footnote for section 5.5

Two-sided P-values (2P) are used throughout, estimated from the "normal" approximation that if treatment had no effect whatever on outcome then (O-E)/sd would be distributed approximately like the standard normal (bell-shaped) distribution. In the standard normal distribution the probability of a result being, just by chance, less than -2 is about 0.025. The probability of it being bigger than +2 is also about 0.025, so the total probability (which is written "2P") of it differing from zero by more than 2 is about 0.05 (i.e. 0.025 + 0.025). If, therefore, (O-E) is negative, indicating a favorable effect of treatment, and is about equal to -2 sd then the two-sided P-value is about 0.05. (Values of -2.6 sd, -3.3 sd and -3.9 sd would correspond to 2P=0.01, 2P=0.001 and 2P=0.0001). Hence, "2P=0.05" means that if treatment does nothing at all then 0.05 is the approximate probability of getting, just by chance, a result at least as extreme as that actually observed (i.e. at least as good as -2 sd in favor of treatment or at least as bad as +2 sd against treatment). In the Figures, the estimated significance levels are printed to 2, 3, 4 or 5 decimal places, according to whether 2P<0.1, 0.01, 0.001 or 0.0001. The abbreviation NS (i.e. Not Significant) is used to denote 2P>0.1.

Footnote for section 5.7

A previous review [3] of some of the mature trials of radiotherapy after mastectomy differs in two ways from Table 3M. First, Table 3M includes many trials that are not yet mature and which, therefore, contribute data only on early survival. Second, Table 3M excludes the early Manchester trials from the main analysis because treatment allocation in those trials was based on odd/even birth dates and may, at least in principle, have been subject to some bias. (For the sake of comparison with the previous review, overall results with the Manchester trials included are given in footnotes to these tables.) Nevertheless, the two reviews have much in common, and neither provides good evidence of any favourable net effect of radiotherapy on medium-term (e.g. 10-15 years) survival after mastectomy.

Footnote for Section 5.8

It can be shown that z/sd is the "one-step" estimator of the log odds ratio, i.e. the first step from a log odds ratio of zero towards the "maximum-likelihood" estimator in a standard (Newton-Raphson) iterative search for the maximum of the log-likelihood function. Hence, exp(z/sd) is called the one-step estimator of the odds ratio. [32] Any bias in this one-step estimator will be negligible in overviews of randomized trials involving **small** treatment effects and **reasonably large** numbers of outcome events.[32] In practical analyses of substantial trial results, it appears that (as long as there is less than a twofold difference in odds and at least several dozen endpoints) the one-step and the maximum-likelihood are about as accurate as each other as estimators of the true odds ratio: for example, from the ISIS-2 data in Table 1 they yield 0.772 and 0.771 respectively.

Footnotes for Section 5.9

* Formally, the area of each solid square has been made proportional to the variance of O-E, since when trying to compare two treatments the size of the variance of O-E can be used in a statistical sense as an estimate of the "information content" of the data (i.e. as the local curvature of the log-likelihood[5])

** The area of the black square describes the amount of information in the overview. It is simply the sum of the areas of the solid squares plotted for the individual trials that contribute to the overview, since the sizes of all the squares throughout this report involve the same scale factor. (The scale factor that was chosen makes the length of the base of each square equal 0.01 times the square root of the corresponding variance.)

Footnote for Section 5.12

In one particular time period, if r, the estimated annual failure rate, equals [no. of failures/no. of woman-years], and b, the estimated log odds ratio (treatment : control), equals [total of (O-E) values/total of their variances] then p, the estimated probability of avoiding failure for one year, equals exp(-r), and the separate probability estimates would be p+0.5p(p-1)b for treated patients and p-0.5p(p-1)b for the control patients.

Footnote for Section 5.17

Moreover, such methods may be of limited statistical sensitivity, particularly when just a few major studies provide most of the evidence. Hence, the loss of statistical power may be illustrated in a particularly extreme form by an important practical example involving only two major trials, where the "assumption-free" (fixed-effects) methods yield a statistically definite (P<0.0001) answer that is strongly supported by a wide range of indirect evidence, but where various "assumed-representativeness" methods might inappropriately fail to do so. There have been only two major randomized trials of the primary prevention of non-fatal myocardial infarction by long-term antiplatelet therapy.[33] The results of one (129/11037 (1.2%) aspirin versus 213/11034 (1.9%) control, P<0.0001) indicated a highly significant 39% ± 9 reduction in the odds of suffering a non-fatal myocardial infarction, but the results of the other (80/3429 (2.3%) aspirin versus 41/1710 (2.4%) control, NS) indicated a reduction of only 3% ± 19, i.e. virtually no difference. (Although the discrepancy between their results appears striking it is in fact only a 2.1 standard deviation discrepancy, so it could well be largely or wholly attributed to the play of chance and/or the data-dependent early closure of the trial with the significant result.) As is generally the case, the overview of the two results may be more reliable than either considered in isolation, and the "assumption-free" overview methods used in the present report indicate an overall reduction of 32% ± 8 (P<0.0001) in the odds of non-fatal myocardial infarction. [33] A reduction of about this size is rendered extremely plausible by its similarity to the significant reductions in non-fatal myocardial infarction (or reinfarction) that have been shown for antiplatelet therapy in other circumstances by randomized trial overviews (e.g. 31% ± 5 among patients with previous myocardial infarction, 35% ± 12 among patients with previous

stroke, 35% ± 17 among patients with unstable angina, or 49% ± 9 among patients who were in hospital because of a suspected acute heart attack). [33] Hence, the use of "assumption-free" (i.e. fixed-effects) methods in a standard overview of these two primary prevention trials yields an extremely definite answer that is almost certainly qualitatively correct. In contrast, since only one of the trials yields a significantly favourable result while the other yields a completely null result, some "assumed-representativeness" (i.e. random-effects) methods might misleadingly have concluded that the two studies together showed no clear evidence of benefit.