

Appendix 1: Statistical Methods

Calculation of Effect Size

Effect sizes were calculated following methods described by Hedges⁵⁵. Hedges' 'g' is a measure of standardised mean difference that can be used with pretest-posttest-control group designs, as is the case in cognitive intervention RCTs. Calculation of mean change scores ($M_{\text{post}} - M_{\text{pre}}$ or $M_{\text{followup}} - M_{\text{pre}}$) between the intervention and comparator conditions (control or other treatment groups) allows an estimate of effectiveness even when the intervention and control groups are non-equivalent.

Hedges' g is preferable to Cohen's d as it is calculated using the pooled pre-intervention SD as the denominator, which has been suggested as the most precise estimate of ES, as the SD of the pre-treatment population is more likely to be comparable across studies⁵⁶. As estimates for g may show positive bias with small sample sizes, a correction was applied to provide a more accurate estimate with small sample sizes, as was the case with many of the studies examined^{56 57}.

Hedges' g was calculated as follows:

$$g = C_p [(M_{\text{post intervention}} - M_{\text{pre intervention}}) - (M_{\text{post comparator}} - M_{\text{pre comparator}})] / SD_{\text{pre}}$$

$$\text{where } SD_{\text{pre}} = \sqrt{[(N_{\text{intervention}} - 1) SD_{\text{preintervention}}^2 + (N_{\text{comparator}} - 1) SD_{\text{precomparator}}^2] / (N_{\text{intervention}} + N_{\text{comparator}} - 2)}$$

$$\text{and } C_p = 1 - [3/4(N_{\text{intervention}} + N_{\text{comparator}} - 2) - 1]$$

A single study may contain two or more intervention or control groups and it may be appropriate to include more than one comparison from the same study in the same meta-analysis. In these circumstances, double-counting of participants was avoided by dividing the value of n in the group used more than once by the number of times it was included in the same meta-analysis, (e.g. if the same control group was included twice in the same meta-analysis the number of control subjects was divided by 2 and this value of n/2 was used in the analysis of each of the comparisons)⁵⁸.

Meta- analyses

Random-effects meta-analyses using a DerSimonian and Laird estimator based on inverse variance weights were employed. Random-effects meta-analysis was chosen as heterogeneity in treatment effects was anticipated because of between-study variations in clinical factors (e.g. content of intervention). The DerSimonian and Laird method incorporates a measure of the heterogeneity between studies⁵⁹.

For each meta-analysis, the overall effect size was calculated by weighing the average effect size for each study according to sample size and then pooling across studies. The z statistic was employed to test whether the pooled effect size was significantly different from 0. The I^2 statistic was used to examine variability in effect sizes between studies. The I^2 statistic estimates the proportion of variation in effect sizes due to heterogeneity, whereby values of 25%-49%, 50%-74% and >75% indicate low, moderate and high heterogeneity respectively⁶⁰. High levels of heterogeneity in effect sizes between different studies can result in potentially misleading conclusions being drawn. If there was evidence of low to high heterogeneity, and greater than 3 studies were included in the meta-analysis, 95% prediction intervals were calculated in order to provide an estimate of the range of treatment effects within an individual study setting. If a prediction interval lies entirely above zero, then it can be concluded that the intervention is beneficial in at least 95% of the individual study settings⁶¹. Finally, publication bias was estimated using funnel plots and the Egger regression asymmetry test. If publication bias was detected, a non-parametric trim and fill method was used to impute missing studies and re-estimate the pooled effect size⁶². An alpha level of 0.05 was used for tests of the estimated average treatment effect and publication bias. Data were analyzed using the 'metan' function in Stata 10 (StataCorp, College Station, TX).

Sensitivity analyses were also conducted by repeating random-effects meta-analyses of the main comparisons using SDs of mean change scores, without correction for upward bias, to calculate weighted mean difference scores. This was performed to examine whether this method of calculating effect size, as used in other meta-analyses, produced differing results than when corrected g values were used as described in the main analyses.

Meta- Regression

Variables examined in the meta-regression were format of intervention (group or individual), study quality (sequence generation, allocation concealment, blinding of outcome assessors), as these have been suggested by previous analyses to influence effect size⁷. Other variables examined were setting of intervention (outpatient/community vs. inpatient/care home facilities), intensity of intervention (hours per week), length of intervention (weeks) and severity of dementia (as determined by mean MMSE score).

If more than 30% of data were missing, the variable was excluded from analyses. The above variables, together with effect sizes, were entered into separate random-effects univariate meta-regression analyses using restricted maximum likelihood estimation. The Knapp-Hartung adjustment was employed to control for risk of false positives with multiple covariates, as incorporated in the 'metareg' command in STATA. Separate meta-regression analyses were conducted for the different outcome measures (MMSE and ADAS-Cog). Any factor that was significant in univariate analyses was entered into a random-effects multivariate meta-regression analysis that corrected for multiple comparisons (thus controlling for the risk of false positives). Data were analyzed using the 'metareg' function in Stata 10.

References

56. Morris SB. Estimating effect sizes from pretest-posttest-control designs. [Organ Res Methods](#) 2008;11:364–86.
57. Durlak JA. How to select, calculate, and interpret effect sizes. [J Pediatr Psychol](#) 2009;34:917–28.
58. Higgins JP, Green S. eds. Cochrane handbook for systematic reviews of interventions: Cochrane Book Series. The Cochrane Collaboration, 2008.
59. DerSimonian R, Laird N. Meta-analysis in clinical trials. [Control Clin Trials](#) 1986;7:177–88.
60. Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. [BMJ](#) 2003;327:557–60.
61. Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. [BMJ](#) 2011;342:d549.
62. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. [Biometrics](#) 2000;56:455–63.