Public health interventions to promote mental well-being in people aged 65 and over: systematic review of effectiveness and cost-effectiveness

Gill Windle[1], Dyfrig Hughes[2], Pat Linck[2], Rhodri Morgan[1], Vanessa Burholt[1], Rhiannon Tudor Edwards[2], Carla Reeves[1], Seow Tien Yeo[2], Bob Woods[3], Ian Russell[4]

1 **Centre for Social Policy Research & Development, Institute of Medical & Social Care Research, University of Wales Bangor.**
2 **Centre for Economics & Policy in Health, Institute of Medical & Social Care Research, University of Wales Bangor.**
3 **Professor of the Clinical Psychology of the Elderly, Institute of Medical & Social Care Research, University of Wales Bangor.**
4 **Professor of Public Health, Institute of Medical & Social Care Research, University of Wales Bangor.**

List of contents

List of Tables (T) & Figures (F)

Summary

A systematic search of 21 data bases and 11 websites sought evidence, published between January 1993 and February 2007, of the effectiveness or cost-effectiveness of interventions to promote mental well-being in later life. The search was restricted to the English language. In principle all study designs were considered for inclusion. In total 15,388 citation titles and abstracts were screened for relevance. By this process 248 articles were identified for further appraisal for inclusion in either review – 218 for effectiveness and 30 for cost-effectiveness. Application of inclusion criteria selected 96 papers for the review – 94 for effectiveness and two for cost-effectiveness. The 94 effectiveness papers described four meta-analyses, 13 trials of good quality (one of which generated two papers), 68 quantitative studies of poor quality (one of which generated two papers) and seven qualitative papers (including five of good quality).

Thus most included studies were of poor quality. Many used small samples that may not represent the population of interest, and certainly lack statistical power. Many recruited participants through advertisements, probably recruiting more motivated individuals, and again making findings less representative. The frequent use of self selection means that women predominate. Few included studies focused on frail older people or people over 80. Few interventions were targeted at alleviating poverty, and none at older people from ethnic or sexual minorities. Few studies answered sub-questions in full, including who delivered the intervention and where.

We divided the 94 studies into 15 categories – six concerned with different types of exercise, three with different types of health promotion, one each with psychological interventions, computer use, gardening, support groups and volunteering, and one residual category. Three of these categories generated useful evidence statements from meta-analyses, and another three from single rigorous trials. Unfortunately two categories, each with three rigorous trials, did not generate a useful evidence statement because the evidence from these trials was conflicting. Another six categories generated no rigorous evidence. While the four good qualitative studies in the final category – mixed health promotion – are helpful, they cannot estimate strength of evidence. Hence the review has generated six robust positive evidence

statements – nos. 1 to 4 relating to exercise, no. 7 relating to health promotion and no. 10 relating to psychological interventions (pp. 9 to 11). Of the two identified cost-effectiveness papers, one added to the evidence on exercise and the other to the evidence on health promotion. To address the lack of economic papers, this review also shows how economic modelling can extend studies concerned solely with effectiveness so as to throw light on cost-effectiveness.

In summary there is a shortage of robust evidence for the effectiveness and cost-effectiveness of interventions to improve the mental well-being of older people. Better research is needed to estimate the value of most interventions. Research into cost-effectiveness is especially sparse, with little economic research even into programmes with evidence of effectiveness. Nevertheless this review has generated six useful positive evidence statements

Evidence statements

1	**Mixed exercise** Two meta-analyses (Arent et al., 2000, MA**+**; Netz et al., 2005, MA**+**), together comprising 68 controlled trials from many developed countries, since augmented by four other rigorous trials in the Netherlands (2), Norway and the US, together provide strong evidence that mixed exercise programmes generally have small-to-moderate effects on mental well-being. As the reported exercise programmes cover a range of types, settings and countries, firm conclusions about the duration of programmes and the frequency of sessions are difficult. It is clear, however, that exercise of moderate intensity (not well defined in the meta-analyses) has beneficial effects on physical symptoms and psychological well-being. The programmes evaluated were generally community-based, well organised and run by trained instructors. The findings apply to similar populations (relatively healthy and independent, and motivated to take exercise) in similar community settings in the UK. The sole qualitative study (Hardcastle & Taylor, 2001; Q**+**) highlights the importance of appropriate facilities and good supervision.
2	**Strength & resistance exercise** Meta-analysis of four US trials that included a total of 1733 independent frail older people and used four of the SF-36 scales to evaluate similar resistance exercise interventions found a significant small-to-moderate improvement in emotional health (Schechtman & Ory, 2001; MA**+**). The findings are likely to be broadly applicable to frail older people in a range of settings in the UK. Of six smaller controlled studies evaluating the benefit of resistance exercise for older people in general, five reported significant positive effects, mostly on the POMS measure (a self-reported measure of general mood over the past week). As all six were of poor quality, this finding should not be considered robust.

3	**Aerobic exercise** A medium-sized RCT in the US showed that both interventions – supervised aerobic brisk walking and 'toning & stretching' – generated similar trajectories of MUNSH and SWLS scores over 12 months in sedentary adults aged 60 to 75; these trajectories showed significant growth in happiness and satisfaction over the six-month exercise period, followed by a significant decrease at 12 months (McAuley et al., 2000, RCT**+**). The findings are likely to be broadly applicable to similar populations in the UK.
4	**Walking interventions** A walking programme delivered to older people in 28 heterogeneous neighbourhoods in Portland, Oregon by trained leaders three times a week over six months improved SF-12 mental health and SWLS life satisfaction scores relative to control neighbourhoods (Fisher & Li, 2004, Cluster RCT**+**). This cluster randomised trial recruited 279 people to the intervention group (of whom 156 completed the intervention) and compared them with 303 controls who received education only. Though recruitment and retention of participants is important for such programmes, the results are likely to be broadly applicable to similar populations in the UK.
5	**Tai Chi** Two out of three rigorous evaluations in the US showed that 3 to 6-month community-based Tai Chi programmes delivered by professionals improve differing mental health measures in older people (Li et al., 2002 & 2004; RCT**+** but not Kutner et al. 1997; NCT**+**). However there was little difference between Tai Chi and less specific exercise programmes. Hence there is no evidence that the distinctive element of Tai Chi confers any benefit.
6	**Other exercise** A US pilot study (Williams et al CBAS-) found that home-based balance-training for 13 older females had no effect on the self-reported SWLS. Another very small study (Tanaka et al., 2002, UBAS-) tested a four-week programme of exercise and short naps on 11 older people in Japan. As only these weak studies were found in this category, the conclusion is that there is no robust evidence that these forms of exercise improve mental well-being.

7	**Group-based health promotion** There is evidence from one well-designed longitudinal trial (Clark et al., 1997, RCT**++**; Clark et al., 2001, RCT**++**) that weekly educational sessions led by occupational therapists promoted and maintained positive changes in the SF-36 mental health score in participants recruited from two federally-subsidised apartment complexes for older adults in the US. Though the findings are likely to be broadly applicable to a similar population in the UK, the findings may not generalise to those in other circumstances (e.g. owner-occupiers & nursing home residents).

8	**Mixed health promotion programmes** There is no quantitative evidence on the effectiveness of mixed health promotion in improving mental well-being. Four qualitative studies suggest that comprehensive health promotion programmes delivered by professionals to homeless, poor or socially isolated older people are acceptable to users and perceived to improve mental wellbeing markedly (Buijs et al., 2003, Q**+**; Greaves & Farbus, 2006, MM**+**; Wilcock, 2006a & 2006b, Q**+**). As three of these studies are British and the fourth Canadian, they are applicable to the UK.
9	**Individually targeted health promotion** There is conflicting evidence from four randomised trials (Halbert et al., 2000, RCT**+**; Kerse, 2005, RCT**+**; Frieswijk et al., 2006, RCT**–**; Markle-Reid et al., 2006, RCT**++**) on the effects on mental well being.of differing health promotion interventions delivered to individuals by professionals In Canada Markle-Reid et al. found that monthly home visits of 1 hour by health promotion nurses significantly improved SF-36 mental health summary scores in the intervention group compared with usual care. The intervention also significantly reduced the costs of prescription drugs in the health promotion group, by enough to offset the costs of the scheme. In the Netherlands Frieswijk et al.. found that a five-part bibliotherapy correspondence course to aid self management in slightly to moderately frail older volunteers living at home, resulted in significant improvements on the

	Sense of Mastery Scale in the short term, but not at six months.. In New Zealand Kerse found that a primary care intervention in which independent sedentary older patients received monthly phone calls from exercise specialists improved SF-36 vitality subscale scores, but had no effect on the mental health scores. In contrast in Australia Halbert et al found that both the provision of 20 minutes of advice on physical activity by an exercise specialist to older patients, and no treatment for controls, in two general practices significantly reduced mental well-being in two SF-36 dimensions – vitality and emotional limitations on role.
10	**Psychological interventions** A meta-analysis (Pinquart & Sörensen, 2001; MA**+**) covering a total of 84 studies from many developed countries provides strong evidence for the effectiveness of cognitive training, control-enhancing interventions, psycho-education, relaxation and supportive interventions in improving the subjective well-being of older people. Their findings apply to similar populations – relatively healthy and independent – in similar community settings in the UK. There is no robust evidence on the effectiveness of more specific psychological interventions – notably dream telling, memory tapping, mental fitness training, resourcefulness training and visual stimulation.
11	**Computer use** Four trials examined the effect of computer training or use (Shrerer et al., 1996, NCT–; White et al., 1999, NCT–; Billipp et al., 2001, NCT-; White et al, 2002, RCT–). As all were of poor quality, there is no robust evidence on the effectiveness of computer use in improving mental well-being.
12	**Gardening interventions** Three studies examined the role of gardening in the mental wellbeing of older people (Barnicle & Midden 2003, CBAS-, Milligan et al, 2004, MM-, Heliker et al., 2000; UBAS-). As there were critical flaws in each study, there is no robust evidence on the effectiveness of gardening interventions in improving mental well-being.

13	**Support groups** Three studies reported the effect of support groups on mental well-being (Barnes & Bennett; 1998, Q-; Stewart et al., 2001, UBAS–; Powers & Wisocki, 2006, UBAS–). As each was of poor quality, there is no robust evidence that support groups improve mental well-being.
14	**Volunteering interventions** Three studies reported the effect of volunteering interventions on older people (Wheeler et al., 1998, MA–; Rabiner et al., 2003, CBAS-; Butler, 2006, MM–). As all were of poor quality, there is no robust evidence on the effectiveness of volunteering in improving the mental well-being of older volunteers or older clients.

15	**Other interventions** A range of other interventions have been evaluated in poor quality studies. It is concluded that there is no robust evidence on the effectiveness of altruistic activity, art therapy, catering redesign in long-term care, home massage, occupational therapy, pet therapy, sleep management, video games and wheelchair modification.
16	**Published studies of cost-effectiveness** Two studies provided good evidence about the cost-effectiveness of interventions to improve the mental well-being of older people. First Hay et al. (2002; RCT**+**) showed that a two-hour group session of preventive advice from an occupational therapist per week is cost-effective in the US with an incremental cost per QALY of $10,700 (95% CI $6,700 to $25,400). Secondly Munro et al (2004; RCT**+**) showed that twice-weekly exercise classes led by qualified instructors are probably cost-effective in the UK with an incremental cost per QALY of £12,100 (95% CI = £5,800 to £61,400). While both studies are sound, one cannot be confident that such sparse findings will apply to similar populations (relatively healthy, living independently, and motivated to take advice and exercise) in similar community-based settings in the UK.

17	**Economic modelling of cost-effectiveness** There are only two published economic analyses of interventions to improve the mental well-being of older people (evidence statement 16). To complement these sparse data needs economic modelling based on the integration of existing studies of effectiveness and existing sources of data about patient utilities and resource costs. The most cost-effective intervention was a thrice-weekly community-based walking programme, delivered to sedentary older people who are able to walk without assistance (Fisher & Li, 2004; Cluster RCT**+**). Modelling yielded an incremental cost per QALY of £7,400 after six months, which is comparable with the two published economic analyses. Modelling was also used to enhance three RCTs of advice about physical activity. Such advice had an estimated incremental cost per QALY of £26,200 when modelled from Kerse et al. (2005; NCT**+**), who estimated the effects of the primary care 'green prescription' counselling programme in New Zealand. The estimated incremental cost per QALY rose to £45,600 when modelled from Markle-Reid et al. (2006; RCT**++**), who evaluated proactive health promotion by nurses in Canada in addition to usual home care for people over 75. However Halbert et al. (2000; RCT**+**) reported decreased mental well-being in response to 20 minutes of individual advice on physical activity by an exercise specialist in general practice in Australia. Thus the advice was dominated by the control group to whom no advice was given.

Glossary of specialist terms

Term	Definition
Bias	Any process in the collection, analysis, interpretation, publication or review of data or studies that can lead to conclusions that deviate systematically from the truth.
Cluster RCT	RCT in which the unit of randomisation is a cluster of participants, e.g. a class, practice or Primary Care Trust
Concurrent validity	Concurrent validity is demonstrated where a test correlates well with a measure that has previously been validated. The two measures may be for the same or related constructs.
Confidence interval	An interval around a statistical estimate to show where the true parameter lies with specified probability or 'confidence'.
Controlled Before and After study (CBAS)	Intervention & control groups are defined and data collected before & after the intervention is administered. CBASs differ from controlled trials in that participants are not allocated to intervention or control groups, but an opportunistic control group is used.
Cost-effectiveness analysis	The consequences of the alternatives to particular interventions are measured in natural units, such as years of life gained. The consequences are not given a monetary value.
Effect size	Magnitude of the effect of an intervention or a relationship between variables, calculated as the ratio of the net effect divided by the population standard deviation of the relevant outcome measure. Since this index is independent of sample size, unlike statistical tests of significance, it is useful in meta-analysis. In this review we describe effect sizes less than 0.2 as trivial, those between 0.2 & 0.4 as small, those between 0.4 & 0.6 as 'small to moderate', those between 0.6 & 0.8 as moderate, those between 0.8 & 1.0 as 'moderate to large' and those greater than 1.0 as large.
External validity (also known as generalisability)	A study is externally valid, or generalisable, if it yields unbiased inferences about a specified target population beyond the subjects in the study (Last, 2001)
Forest plot	Common method of displaying the results from a meta-analysis. The results of each study are displayed graphically as squares centred on each study's point estimate of the intervention effect with horizontal lines representing confidence intervals (usually 95%) for that effect.

Heterogeneity	Differences in study design or methods, or in the characteristics or distributions of populations to be compared, rendering comparison invalid.
Homogeneity	In contrast to the previous term, this describes consistency in study design or methods, or in the characteristics or distributions of populations to be compared, rendering comparison valid.
Intention-to-treat analysis	Method of data analysis in which participants are analysed in the group to which they were allocated regardless of whether they complied with their allocated intervention or treatment.
Internal consistency	Internal consistency is an estimate of how much a measure is based on systematic experimental technique, so that reliable inferences about cause-consequence relations may be made
Internal validity	A study is internally valid if it yields unbiased comparisons of cases and controls within the study (or intervention and control groups) apart from sampling error (after Last, 2001)
Meta-analysis (MA)	A mathematical procedure that combines quantitative evidence from a number of different studies, facilitating combination of these studies and comparison between them.
Mixed methods (MM)	The combination of two or more distinct research methods, typically quantitative & qualitative, to validate findings through triangulation, i.e. showing consistent results from the independent sources.
Non-Randomised Controlled Trial (NCT)	Trials in which individuals are allocated between intervention and control groups but the allocation is not randomised (for example alternate allocation).
Qigong	The form of traditional Chinese medicine that coordinates breathing patterns with physical postures and body movements. Though it is mostly taught for health maintenance, some teach it as a therapy.
Randomised Controlled Trial (RCT)	Individuals or defined groups of individuals are randomised to either an intervention or a control group. If well implemented randomisation should ensure that intervention and control groups only differ in their exposure to treatment.
Reliability	The consistency of measurements or measuring instruments. Reliability does not imply validity, because a reliable measure can be consistent without measuring what it is supposed to measure.
Systematic review (SR)	Method of finding & selecting primary studies relating to a defined topic, and appraising & synthesising the resulting evidence.

Tai Chi	The slow motion routines practiced every morning in parks around the world, notably in China, to promote health and longevity.
Uncontrolled Before and After study (UBAS)	Intervention group is defined and data collected before & after the intervention is administered. UBASs differ from CBASs in that there is no control group.
Validity	Achieved when the measuring instrument (e.g. questionnaire) measures what it is designed to measure.

Measures of mental well-being used by studies reported in this review

Measure	Description
Affect Balance Scale (Bradburn, 1969)	A 10-item rating scale containing five statements reflecting positive feelings and five statements reflecting negative feelings. Administered to determine overall psychological well-being at a given point in time. Positive affect questions receive a rating of 1 for yes and 0 for no. A Positive Affect Scale score is obtained by summing ratings for the five positive affect questions. Scores range from 0 to 5. A Negative Affect Scale score is obtained by summing the ratings for the five negative affect questions. Scores range from 0 to 5. The Affect Balance Scale score is computed by subtracting Negative Affect Scale scores from Positive Affect Scale scores and adding a constant of 5 to avoid negative scores. Scores range from 0 (lowest affect balance) to 10 (highest affect balance). Satisfactory levels of reliability have been reported ranging from 0.47 to 0.73 for the positive scale and 0.48 to 0.73 for the negative scale (Cherlin & Reeder, 1975; Warr, 1978). Barnicle & Midden (2003) used a derived five-point scale (strongly disagree, disagree, neutral, agree and strongly agree). This gave a score range from -20 (lowest level of psychological well-being) to +20 (highest psychological well-being). A score of 0 indicates neutral psychological well-being. This scale was found to be reliable in this population over a 5-7 day test-retest period with Pearson correlation = 0.72. Thus the adjustment of the measure did not compromise reliability.
Apathy Scale (Starkstein et al., 1995)	The scale consisted of 14 headings, with points between 0 & 3 allotted to each. Higher scores reflect apathetic mood. The scale was validated for use in stroke, Parkinson's and Alzheimer's disease.
Campbell's index of well-being (Campbell et al., 1976)	Self-report measure with multiple choice items ranging from 1 to 7. The index is the sum of 2 measurements: the average score on an Index of General Affect (8 items on semantic differential scales); and a single-item assessment of life satisfaction. The measure is reliable and valid: the index has a Cronbach's alpha of 0.89, and a correlation of 0.55 with the life satisfaction question (Robinson, 1981).

Coopersmith Self-Esteem Inventory (1981)	The scale ranges from 25 to 50. Reliability and validity have been reported as adequate in several studies.
Emotional well-being scale (Campbell & Aday, 2001).	Campbell & Aday (2001) developed several measures to explore health and well-being, including a 7-item emotional well-being subscale that asked participants whether they felt the intervention under evaluation had brought about any changes in this dimension. Life satisfaction and mental health were also assessed in this study, but there is no reference to how scales were constructed. Cronbach's alpha of 0.86 indicated strong internal consistency. The scale appears to require further validation.
Emotional well-being scale (Hermans & Tak-van der ven, 1973)	A measure of the relative proportion of positive to negative feelings experienced by an individual. Positive and negative feelings are treated as independent of each other. Respondents agree or disagree with positive and negative statements. Goldstein et al. (1997) selected 10 items, 5 positive and 5 negative, from the 36-item scale by Hermans & Tak-van der Ven (1973). There is no mention of validating this alternative version.
Exercise induced feeling inventory (Gauvin & Rejeski, 1993)	A 12-item inventory that assesses 4 dimensions: positive engagement (enthusiastic, happy, and upbeat); revitalisation (refreshed, energetic, and revived); tranquillity (calm, relaxed and peaceful); and physical exhaustion (fatigued, tired and worn-out). On the 5-point scale subjects indicate how strongly they experience each feeling state immediately after one period of exercise. Anchors were: 0 = do not feel and 4 = feel very strongly. Internal consistency exceeded 0.70 for each subscale (McAuley & Courneya, 1994). In Matsouka et al. (2003) Cronbach's alpha exceeded 0.85. The subscales have good internal consistency, share expected variance with related constructs, are sensitive to different exercise settings and responsive to different social contexts.
General Health Questionnaire (GHQ-28; Goldberg and Hillier, 1979)	A 28-item version of the GHQ commonly used as a measure of psychological well-being (low scores, while poor health scores high). There has been extensive testing of validity, reliability, and sensitivity (Bowling, 1995). It has been widely used and found to be acceptable by participants over 65 years. The scale has four subscales: A – somatic symptoms; B – anxiety & insomnia; C – social dysfunction; D – severe depression.

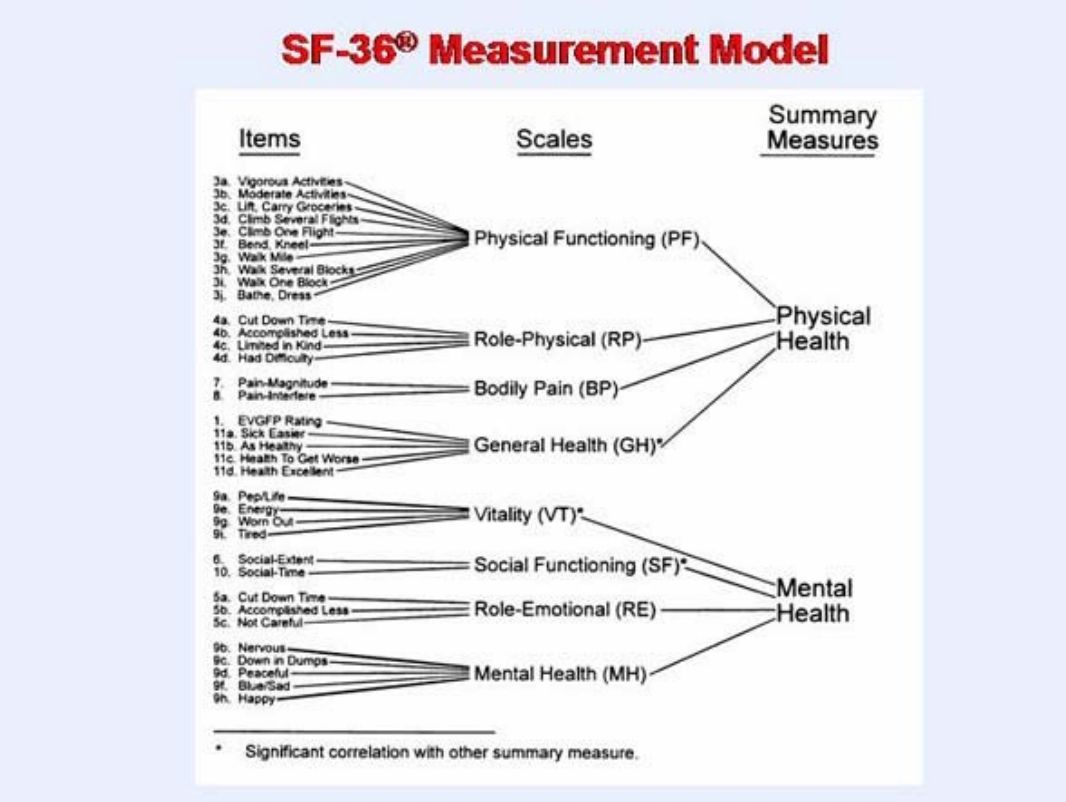
General self efficacy (Sherer et al., 1982)	A 16-item scale that measures self-efficacy expectations across a variety of situations. The scale consists of two components – initiation & persistence of behaviour, and efficacy in the face of adversity.
LEIPAD II Short version (LEIPAD SV; De Leo et al., 1998)	The measure is designed to gauge subjective perceptions of quality of life in the elderly. It comprises 25 items in 6 scales – cognitive functioning scale (CFS), depression & anxiety scale (DAS), life satisfaction scale (LSS), physical function scale (PFS), self-care scale (SCS) and social functioning scale (SFS). Each item is scored from 0 (best) to 3 (worst). Validity and reliability of this short version of the scale are not reported.
Life Attitude Profile (LAP-R; Reker & Peacock, 1981)	A measure comprising 6 sub-scales (choice responsibility; coherence; death acceptance; existential vacuum; goal seeking, life purpose) and two composite scales (personal meaning index [PMI) and existential transcendence [ET]). Each of the 8 items per subscale is rated on a 7-point Likert-type scale (1-7) from "strongly disagree" (1) to "strongly agree" (7).
Life Satisfaction Index-A (LSI-A; Neugarten et al., 1961).	A 20 item self-report instrument used to measure subjective well-being (reflecting satisfaction with life) among individuals aged 65 or over. Respondents are asked to either agree, disagree or express neutrality on each items. Twelve items are positively worded, and eight are negatively worded. Each agreement with a positively worded item receives 1 point and each disagreement with a negatively worded item also receives 1 point. Other responses are scored 0. Totalling the number of points creates a score ranging from 0 to 20, with higher values indicating greater life satisfaction. Neugarten et al. (1961) report a mean score of 12.4 (SD 4.4). Reported internal consistency ranged from 0.73 (Hooker & Ventis, 1984) to 0.84 (Wolk & Kurtz, 1975). Concurrent validity is also reported. This measure has been extensively used in gerontological research.
Life Satisfaction Index-Z (Wood et al, 1969)	A global measure of past, present and future states, this scale was developed for use with older populations in different ethnic groups. It is commonly used to measure well-being in gerontology research and is considered to indicate successful aging (Bowling, 1991). Respondents agree or disagree with each of 13 items about satisfaction with life scoring 0, 1 or 2, yielding a total score ranging between 0 & 26. In initial scale development, split-half reliability was 0.79.

Life Satisfaction Ladder Scale (Cantril, 1965)	A 10-item vertical self-report scale with item responses from 1 = very dissatisfied to 10 = very satisfied. Scale scores therefore range from 10 to 100, with higher scores indicating higher satisfaction. Adequate reliability and validity have been reported.
Life Satisfaction Questionnaire (LSQ; original ref. not cited)	Carlsson et al. (1999) concluded that the LSQ, developed for women with breast cancer, has acceptable validity and reliability. Respondents indicate their degree of satisfaction with their finances, health, religious experience and social relationships. Neither users, Powers & Wisocki (1997) or Carlsson, present information on validity or reliability.
Life Satisfaction Visual Analogue Scale	Apparently created by Dungan et al. (1996) for their own use. Unfortunately they give no details of content, scoring or use.
Locus of Control Scale (Levenson, 1974).	A 24-item self-report measure of internal orientation and two types of external orientation – belief in chance and belief in control by others. The scale shows adequate reliability and validity, and has been used effectively with older adults (Shewchuk et al. 1990).
Loneliness Scale (De Jong et al., 1999).	The self-report scale has 5 positive and 6 negative items assessing sense of belonging and discrepancies in desired relationships. Item scores range from 0 (not lonely) to 11 (extremely lonely); scores above 3 show loneliness, while those above 9 show extreme loneliness. The scale has been used in several surveys and is reliable and valid with older persons.
Memorial University of Newfoundland Scale of Happiness (MUNSH; Kozma & Stones, 1980)	Self-report measure with 4 subscales – positive affect (PA), negative affect (NA), positive experiences (PE) and negative experiences (NE). There are 5 binary items for PA, 5 for NA, 7 for PE, and 7 for NE. The total score comes from the formula (PA-NA) + (PE-NE). As this can give a negative score, Elavsky et al. (2005) added 24 to the total score, thus giving a range from 0 to 48. The scale has been validated in several settings. Elavsky et al. (2005) reported internal consistency as more than 0.75.
Mental Health Index (MHI-5; McCabe et al., 1996)	This is derived adapted from five mental health items of the SF-36, leading to a summary score between from 0 & 25. Each item asks respondents about mood over the last 4 weeks and how long they have felt that way (from none of the time to all). There is no information on reliability or validity in the study which employed this measure (Clark et al., 2003).

Morale & Life Satisfaction Scale (Clark & Anderson, 1967)	A 45-item measure covering 8 dimensions: depression & satisfaction; equanimity; negative aspects of age; physical condition; positive aspects of age; social accessibility; social alienation; & will to live. Evidence of validity & reliability reported by Clark & Anderson (1967) and Pierce & Clark (1973).
Perceived Control of Life Situations (Eizenman et al., 1997).	Self-report measure with eight items answered on a 4-point scale from 1 = agree strongly to 4 = disagree strongly. Scores therefore range from 8-32 with higher scores indicating greater control. White et al. (2002) reported adequate reliability and validity.
Perceived well-being scale (Reker & Wong, 1984)	The scale includes 14 items about physical and mental function. No information available about validation.
Perceived well being scale – revised (PWB-R)	Measures perceived physical & emotional well being, using 16 items – 8 psychological & 8 physical, each measured on 7-point Likert scale. Items are randomly ordered and varied between positive and negative to control for response set bias. Cronbach's alpha reported between 0.79 & 0.85.
Philadelphia Geriatric Center (PGC) Morale Scale (Lawton, 1975)	This 17-item revised scale is a measure of general well-being and positive future outlook. There are two widely used alternative versions – interview, or self-report questionnaire. High morale responses are scored as a 1; low morale responses are scored as 0, so scores range from 0-17, 0 being low morale, 17 being high morale. Three factors represent morale – agitation, attitude toward own aging, and lonely dissatisfaction. Valid and reliable
Philadelphia Geriatric Center (PGC) Morale Scale -- modified	A measure of life satisfaction. The resulting scale ranges from 0 to 11, with higher values indicating greater satisfaction with life. The alpha reliability of this measure is reported as 0.78 by Rabiner et al. (2003).
Positive and negative affect schedule (PANAS; Watson et al., 1988).	A 20 item self-report measure with 10 items reflecting positive affect (PA), and 10 negative affect (NA). High PA reflects high energy, full concentration and pleasurable engagement. Low PA is characterised by sadness and lethargy. Low NA reveals a state of calmness and serenity. Internal consistency is 0.80 for PA and 0.84 for NA (Martina & Stevens, 2006).

Profile of Mood States (POMS; McNair et al., 1971)	Self-report measure assessing general mood over "the past week including today". Respondents, provided with a list of 65 adjectives describing mood states, indicate their agreement with adjectives on a 5-point Likert scale ranging from 0 (not at all) to 4 (extremely). This yields 6 sub-scores – anger & hostility, confusion & bewilderment, depression & dejection, fatigue & inertia, tension & anxiety and vigour & activity (with differing numbers of items per subscale). Total mood disturbance is calculated by summing the 5 negative scores and subtracting the one positive score – vigour & activity (McLafferty et al., 2004). Most respondents complete the POMS in 3 to 5 minutes (McNair et al., 1992). Across subscales internal consistency ranges from 0.87 to 0.95, and test-retest coefficients from 0.65 to 0.74. Concurrent validity has also been shown.
POMS short form (POMS-SF; Curren et al., 1995)	Curren et al (1995) reduced POMS (previous row) to 30 items – 5 for each of the 6 subscales. All internal consistencies remained above 0.7.
POMS modified version (Jette et al., 1996)	Another shortened version of the POMS, similar to the 30-item POMS-SF, but with the 5 items with highest item-total correlations (as reported in the test manual) selected for 5 of the subscales and 8 such 'best' items for the depression subscale. Though methodologically slightly better than POMS-SF, internal consistencies were similar, i.e. greater than 0.7.
POMS – Korean version (Shin, 1999)	The POMS was adapted for Korean elders through cultural verification and psychometric evaluation (Shin, 1996). The new instrument comprises 3 factors – anxiety & depression (21 items), vigour (8 items) and anger (5 items). Cronbach's alpha has risen to 0.95 – very high if not too high.
Psychological well-being (Becker, 1989)	Self-report personality questionnaire with 3 subscales – meaning of life, self-attentiveness & preoccupation (having self-centred thoughts and anxiety & concern about self & the future), and 'complaintlessness'. Validity and reliability are not reported.
Psychological well-being (Ryff,1989)	Scale to measure perceived psychological well-being comprising 14 items about personal growth, 3 items about positive relations with others, and 3 about self-acceptance. High internal & test-re-test reliability, and convergent & discriminant validity reported for various age-groups including older adults.

Quality of Life Profile: Senior Version (Renwick et al., 1996)	According to Raphael (1996) the QOLP-SV rates highly in importance, satisfaction and quality of life scores. Nevertheless neither paper reports on content or scoring!
Rosenberg Self-Esteem Scale (Rosenberg, 1965)	This scale comprises 10 self-reported 4-point Likert scales labelled "strongly disagree", "agree", "disagree" & "strongly disagree". High total score shows high self-esteem. Reliability over time > 0.83; reproducibility coefficient = 0.92; scaleability coefficient = 0.72. Validity shown by significant negative correlations between self-esteem and clinical ratings of depression.
Satisfaction With Life Scale (SWLS) (Deiner et al., 1985)	Measure with 1 item for each of 5 life domains, all rated on a 7-point Likert scale from 1 (strongly disagree) to 7 (strongly agree). Higher scores show greater level of satisfaction with life. Reports of good psychometric properties including internal consistency include those from Diener et al. (1985) [also test-retest reliability], McAuley et al. (2000), Macfarlane et al. 2005 [also test-re-test reliability and construct validity] and Martina & Stevens (2006).
Self-rated mood (Tamake et al., 1999)	Single-item self-report measure of mood. No other details are provided.
Sense of Mastery Scale (Pearlin & Schooler, 1978)	Self–report scale with 7 items assessing sense of control over one's life. (Froelicher et al. 2004) reported construct validity and internal consistency in a sample of women with mean age = 61.
Short Form-12 (SF-12) (Ware et al., 1995)	Widely used self-report measure comprising 12 items from SF-36 (following row) yielding separate scores for physical and mental health. One can also derive a single health utility score for economic analysis. Scores are transformed to lie between 0 & 100, with higher scores showing better health. Reliability & validity have been established in numerous studies, including Ware et al. (1996) and Ware et al. (1998).

Short Form 36 (SF-36; Ware & Sherbourne,1992), previously known as Medical Outcomes Study Instrument (MOSI)	Ubiquitous, multi-purpose, self-reported, generic health survey, comprising 8 scales derived from 36 questions about functional health & well-being, and yielding psychometrically-based physical & mental health summary scores and a preference-based health utility index. Scores are transformed to lie between 0 & 100 with high scores showing better health. There are many reports of good psychometric properties including internal consistency, test-retest reliability, and all types of validity (content, concurrent, predictive, criterion & construct) including Ware et al. (1993), McHorney et al. (1994), Ware et al. (1994), Tsai et al. (1997) and Schechtman & Ory (2001). 
Social Production Function Index Level Scale (SPF-IL; Nieboer et al, 2005)	Self-report measure of affective and cognitive components of well-being. There are 15 items – 3 in each of the 5 subscales representing the dimensions of well-being from the SPF theory – affection, behavioural confirmation, comfort, status and stimulation. Nieboer et al. (2005) report that validity and reliability have been extensively tested and that SPF-IL has other good psychometric properties.

Sources of Meaning Profile-Modified (Reker, 1988)	Measures the sources and degree of personal meaning in one's life, by using 16 7-point items to tap activities, commitments, and pursuits.
Subjective Quality of Life Profile (Gerin et al. 1992).	This tool has 4 categories from which investigators can choose items relevant to their study, including functional life (concerning both physical and mental health), social life (relationships, social roles & interest in the exterior world) & spiritual life (capacity to have aesthetic or religious experiences, to meditate and to reflect). The last category asks subjects to evaluate their personal programme, to rate its relevance, and describe how it was experienced. Item scores lie between -2 (perceived to be highly unsatisfactory) and a maximum of 2 (perceived to be highly satisfactory). Importance attached to items has 3 levels – without importance (0) up to very important (2).
Subjective Satisfaction & Refreshment Scale (Hirawayka et al., 2001)	Responses to variations on the question "To what extent do you feel satisfied and refreshed in daily life?" are assessed on a 4 point scale (3 = strongly; 2 = moderately; 1= slightly, 0 = not at all). Details of validity & reliability are not available.
Subjective scale of well-being for older persons (SSWO; Tempelman, 1987)	Paw et al. (2007) used the Dutch version of the SSWO, a self-report measure with 30 items divided into 5 subscales – health (5 items), contacts (5 items), morale (6 items), optimism (7 items) and self-respect (7 items). The total score measures general well-being. The test-retest reliability coefficient was 0.85.
UCLA Loneliness Scale (Russell et al., 1980).	This scale is often considered the gold standard of loneliness scales. It assesses subjective feelings of loneliness or social isolation on a 20-item scale with scores ranging from 20 to 80. McAuley et al. (2000) & Brown et al. (2004) both reported internal consistency of more than 0.9, & Brown added test-retest reliability of 0.73. Russell (1996) established convergent validity through highly significant correlations of 0.65 with the NYU Loneliness Scale and 0.72 with the Differential Loneliness Scale

World Health Organisation Quality of Life questionnaire (WHO-QOL-100; WHO Quality of Life Group, 1998)	The WHO-QOL-100 was developed simultaneously in 15 international centres through item creation, focus groups, pilot tests and field tests. An initial pool of 1000 questions was reduced to 100 items, grouped into one generic facet on quality of life & health perceptions, and 24 specific facets, originally grouped into 6 domains – Environment, Independence, Physical, Psychological, Social Relationships & Spirituality. Most participants complete the survey themselves, though a small number with literacy problems get a structured interview. Items are scored on a 5-point Likert scale specifying only anchor points (e.g. 'never' & 'always'). Test-retest reliability & internal consistency are both good, especially in Britain (Skevington, 1999). The WHO-QOL-100 is adept at identifying facets of quality of life which are cross-culturally important (WHO-QOL Group 1998, Power et al. 1999). Confirmatory factor analysis showed that the 6-domain model was not as good a fit as a 4-domain model combining Independence with Physical, and Spirituality with Psychological. Scores discriminate well between sick and well people and concur with reported health status.
Worry Questionnaire (Wisocki, 1988)	Respondents estimate the amount of worry & associated physical feelings they experience in the domains of health, finance and social relationships. Wisocki (1988) presents no information on reliability or validity.

Abbreviations

ABS	Affect Balance Scale
ADL	Activity of Daily Living
ANOVA	Analysis of Variance
ANCOVA	Analysis of Covariance
AOR	Adjusted Odds Ratio
APV	Analysis of Partial Variance
C	Control
CBAS	Controlled Before-&-After study (two groups of participants)
CI	Confidence Interval
CFI	Comparative Fit Index
GHQ	General Health Questionnaire
HUI	Health Utilities Index
HR	Heart rate
HRQoL	Health-related quality of life
I or IV	Intervention
LSI	Life Satisfaction Index
M	Mean
MA	Meta-analysis
MANOVA	Multivariate Analysis of Variance
MANCOVA	Multivariate Analysis of Covariance
MCS	Mental Component Score (of SF-36)
MI	Mental Illness
MM	Mixed methods
MMSE	Mini Mental Status Examination
MOS	Medical Outcome Study
N	Number
NCT	Non-randomised controlled trial
NSF	National Service Framework
OR	Odds Ratio
OT	Occupational Therapy ***or*** Occupational Therapist
PAQ	Physical Activity Questionnaire

PCS	Physical Component Score (of SF-36)
POMS	Profile of Mood States
PWB	Psychological Well-being
Q	Qualitative study
QALY	Quality adjusted life year
QoL	Quality of life
QWB	Quality of Well-being
RCT	Randomised controlled trial
RPE	Rating of Perceived Exertion
SD	Standard Deviation
SE	Standard Error
SEIQoL	Schedule for Evaluating Individual Quality of Life (validated questionnaire)
SES	Socio-economic Status
SWB	Subjective Well-being
T	Time
UBAS	Uncontrolled Before-&-After Study (single group of participants)
wk	Week

1 Introduction

1.1 Aims of the review

1.1.1 To identify and review all relevant evidence about public health interventions to promote mental well-being in older people aged 65 and over.

1.1.2 To identify and review data on the costs and cost-effectiveness of public health interventions to promote mental well-being in older people aged 65 and over.

1.1.3 To highlight gaps in the evidence base and make recommendations for further research.

1.2 Target audience and structure

The work is aimed at professionals and practitioners working in the NHS, other public sector organisations, the private sector and the voluntary and community sectors. It is also relevant to carers and family members who have direct or indirect responsibility for the care and support of older people.

The report therefore adopts the traditional structure – introduction (Chapter 1), methods (Chapter 2), results and discussion (Chapter 6). To expedite the work we divided it into three tasks – review of effectiveness (Chapter 3), review of cost-effectiveness (Chapter 4) and economic modelling of cost-effectiveness (Chapter 5). As Chapter 3 reports on 95 included papers, while Chapter 4 reports on only two, Chapter 5 shows how studies of effectiveness can be extended after publication to throw light on cost-effectiveness, provided they are reported assiduously.

1.3 Research Questions

The review addresses one main research question:
What are the most effective and cost-effective ways for primary and residential care services to promote the mental wellbeing of older people?

Table 1 below uses the Population Intervention Comparison Outcome (PICO) format to show how the main research question addresses specific issues within the types of interventions – for all older people and for sub-groups of this population.
Within the main question the review addresses six sub-questions:

1.3.1 What is the frequency and duration of an effective intervention?

1.3.2 What are the significant features of an effective intervener?

1.3.3 Are interventions that engage older people in their design and delivery more effective than those that do not?

1.3.4 Are interventions that engage immediate family members or carers more effective than those without such engagement?

1.3.5 Does the intervention lead to any adverse or unintended effects?

1.3.6 What are the barriers to and facilitators of effective implementation?

1.4 Background – setting the context

Population ageing is emerging as a worldwide trend, reflecting economic development, improvements in education and health care, increases in life expectancy and falls in fertility. The oldest old (80 and over) are the fastest growing group in many nations (Kinsella & Velcoff, 2001). This demographic change emerged first in Europe (Scharf et al., 2003). By 2021 this oldest age group will constitute almost 5 per cent of the population of the UK (Office for National Statistics, 1999). The UK is one of the world's 25 oldest countries, with 20.4% being aged 60 or over (Kinsella & Velcoff, 2001). By 2025 the number of people over 65 in the UK will exceed the number under 16 by 1.6 million (Office for National Statistics, 2003).

Population ageing presents many challenges for government policies and the health and social services, particularly the perceived increasing burden of pensions, and health and social care provision. Ageing can be accompanied by biological changes that increase the risk of illness, disability and death (Office for National Statistics, 1999). Although life expectancy has increased and mortality decreased, it is not clear whether there have been concomitant improvements in morbidity in older age (Office for National Statistics, 1999).

Table 1 Which interventions are effective and cost effective for promoting mental well-being?

POPULATION	INTERVENTION	COMPARISON	OUTCOME
All people > 65 years	o Self-care interventions (e.g. health promotion, health education, exercise & physical activity, dietary advice, leisure activities, e.g. hobbies, gardening, arts). o Psychological interventions (e.g. cognitive training, relaxation techniques) o Social interventions (e.g. peer support, volunteering, group activity or participation, befriending, provision of advice & information, social support) o Environmental interventions (e.g. housing adaptations, low-level support, technology, transport)	Interventions compared with each other & with no intervention	• Mental well-being • Utility • Cost / QALY
Population sub-groups eg: o Age groups o Gender o Ethnicity o Frail elders o Older carers o Physically restricted o Rural or urban o Sexual orientation	o Self-care interventions (e.g. health promotion, health education, exercise & physical activity, dietary advice, leisure activities, e.g. hobbies, gardening, arts). o Psychological interventions (e.g. cognitive training, relaxation techniques) o Social interventions (e.g. peer support, volunteering, group activity or participation, befriending, provision of advice & information, social support) o Environmental interventions (e.g. housing adaptations, low-level support, technology, transport)	Interventions compared with each other & with no intervention	• Mental well-being • Utility • Cost / QALY

In addition to health, there are many other life events and changing circumstances that threaten the well-being of older people. Bereavements, changes in financial situation and social relationships are all important aspects of older age. Hence there is concern within public policy and society as to how quality of life can be maintained and enhanced in older age. This requires a wide ranging approach, moving away from current perspectives that tend to portray older age as a problem and a burden, and challenging discrimination and negative stereotypes of ageing. Providing opportunities for improving well-being can help to ensure that older people lead healthy and fulfilling lives for as long as possible.

Why examine the promotion of mental well-being?

The examination of mental health has traditionally focused on mental illness in clinical populations. A review of the literature found that research publications address negative psychological states rather than positive in a ratio of seventeen to one (Diener et al, 1999). This focus has also been reflected in government policy. However, more recent policy initiatives (e.g. National Service Frameworks, The Healthy Ageing Action Plan for Wales) stress the importance of promoting good mental health for all in addition to treating and supporting those with mental illness.

The 2006 UK Inquiry into Mental Health and Well-Being in Later Life (Age Concern England, 2006) stresses the importance of promoting positive mental health in older age. It acknowledges that given the growing numbers of older people, the promotion of mental well-being will be of benefit to the whole of society and potentially reduce the costs of care (p.10). The Inquiry also suggests that the promotion of mental well-being has traditionally been neglected in favour of promoting physical health.

Hence there is a need to examine how mental well-being might be promoted in the general older population, rather than clinical samples. A review of reviews examining public health interventions to promote positive mental health and prevent mental health disorders among adults (Taylor et al, 2007) found little level evidence on universal interventions for the general population. They state:

> "systematic reviews of mental health promotion interventions for all adults, not just those who are already using mental health services, are vital to develop a

> credible evidence base that will support the implementation of broad policy goals to improve the mental health of populations".

This is reinforced by 'Living Well In Later Life', the review by the National Service Frameworks (NSF) in England which stresses the need to improve all aspects of mental health services. Thus the development of services to promote mental health in later life is a priority. Furthermore Wanless (2004) highlighted the need to examine the effectiveness of public health interventions, so as to provide a cost-effective service that reduced health inequalities. Thus it is timely to examine the effectiveness and cost-effectiveness of interventions that promote mental well-being in older age.

What is mental well-being?

Well-being is an elusive concept, and the term is often used interchangeably with that of quality of life. A recent literature review concluded that there is no accepted definition of either well-being or quality of life (Scottish Executive, 2005). Philosophical writings show that the search for a life of well-being is ancient, yet it is only recently that it has been systematically measured and studied (Diener et al., 1997). Therefore attempts to investigate well-being scientifically are relatively new (Diener & Suh, 2000).

In defining mental well-being, this review acknowledges that well-being is primarily a subjective, individual experience. Keyes (2003) proposes that an individual's sense of well-being encompasses a syndrome of positive symptoms that engender a state of mental health, just as syndromes of negative symptoms can trigger mental illnesses such as depression. Thus the focus is to understand the individual's positive levels of psychological functioning, not just on undesirable clinical states such as depression. So this review will examine a range of measures that are indicators of positive psychological functioning.

The determinants of well-being

A substantial amount of research has examined the determinants of well-being in older age. In her review of the quality of life literature, Bowling (2004) classified the many approaches taken towards examining the concept of well-being and quality of

life. She focussed on literature driven by theory or other research, rather than derived from older people themselves. She focused on the results of her previous electronic and manual searches. The inclusion criterion was that the literature should have made a contribution to the conceptual development and definition of quality of life. Whilst acknowledging the multi-dimensionality of quality of life, the review examined the concept of subjective well-being. The most frequently reported associations with well being or quality of life in older age are good health and functional ability, a feeling of usefulness or adequacy, social participation, friends and social support and level of income or other indicator of socio-economic status (Bowling, 2004, p.8).

A parallel systematic review focused on the factors nominated by older people themselves so as to draw comparisons with those identified by Bowling (Brown & Flynn, 2004). This review focussed on articles written in English which had either explicitly used individualised quality of life measures (such as SEIQoL), survey methods or qualitative methods to explore what was important for quality of life in older age. Forty three papers met the inclusion criteria, 22 from the UK. The review found that the main factors regarded as important for quality of life were: relationships with family and others; happiness; religion or spirituality; independence, mobility & autonomy; health; social and leisure activities; and finance & standard of living (Brown & Flynn, 2004).

Though these two reviews do not provide an exhaustive list of associations of well-being in later life, they do provide a general consensus of the key factors associated with well-being. The UK Inquiry into Mental Health in Later Life reports that participation in meaningful activity, relationships and physical health all engender positive mental health in older age, while discrimination and poverty reduce it. So they highlight areas where interventions might usefully promote mental well-being. Examining the literature for evidence about such interventions is the aim of this work.

2 Methods

2.1 Literature Search

A systematic search of the literature was conducted by the NHS Centre for Reviews and Dissemination (CRD) at the University of York using the search strategy in Appendix A. The following electronic databases and websites were searched from January 1993 to February 2007 for literature published in English.

2.2 Electronic Databases

Age Info, Ageline, AMED, ASSIA, British Nursing Index, CINAHL, Cochrane Database of Systematic Reviews (CDSR), Cochrane Central Register of Controlled Trials, Database of Abstracts of Reviews of Effectiveness (DARE), ECONLIT, EmBase, HEED, HMIC, Medline, National Electronic Library for Health (NELH) – specifically the Specialist Libraries for Later Life & Mental Health, NHS EED, National Research Register, Current Controlled Trials, PsycInfo, Research Findings Register, SIGLE, Social Care Online, Social Science Citation Index & Sociological Abstracts.

2.3 Websites

- Age Concern England http://www.ageconcern.org.uk/
- Centre for Policy on Ageing, http://www.cpa.org.uk/index.html
- Department for Work and Pensions http://www.dwp.gov.uk/
- Help the Aged http://www.helptheaged.org.uk/en-gb
- Joseph Rowntree Foundation http://www.jrf.org.uk/
- Mental Health Foundation http://www.mentalhealth.org.uk/
- National Institute for Health & Clinical excellence (NICE) http://www.nice.org.uk/ (including past work by the Health Development Agency, searched separately within the site at http://www.nice.org.uk/page.aspx?o=hda.publications)
- Sainsbury Centre for Mental Health http://www.scmh.org.uk/80256FBD004F6342/vWeb/wpKHAL6S2HVE
- Scottish Executive – research section of website http://www.scotland.gov.uk/Topics/Research/Research
- UK Independent Inquiry into Mental Health http://www.mhilli.org/index.aspx
- Welsh Assembly government – health and social care section http://new.wales.gov.uk/topics/health/?lang=en

2.4 Inclusion and exclusion criteria

The reviews focus on interventions that promote mental well-being in older people, defined as those over 65. Studies were considered if they included older people, for example studies of 50-70 year olds, but only if they subdivided results by age groups.

2.4.1 Inclusion Criteria

Population:

The target population was people aged 65 and over living at home, in the community, in supported housing or in residential care homes.

Interventions:

Interventions and activities that promote or sustain mental well-being in older people, provided by their carers, families, peers, practitioners, professionals or volunteers. The wide range of interventions considered included:

- Self-care interventions (e.g. health promotion, education, advice and information, exercise and physical activity, dietary advice)
- Psychological interventions (e.g. cognitive training, relaxation techniques)
- Social interventions (e.g. peer/social support, volunteering, group activity or participation, befriending, leisure activities e.g. hobbies, gardening, arts, crafts)
- Environmental interventions (e.g. housing adaptations, low-level support, technology, transport)

Study designs:

There is general consensus that randomised controlled trials (RCTs) and meta-analyses of RCTs provide the best evidence of effectiveness. In public health, however, a lot of evidence about interventions comes from studies lower in the hierarchy of evidence (e.g. non-randomised controlled trials – NCTs). To provide a comprehensive picture, all study designs were included, and their limitations noted.

Outcomes:

The term ‘mental well-being’ covers a broad spectrum of possible outcome measures. For the purpose of these reviews interventions were included that seek to promote, improve, enhance, sustain and benefit validated measures and self-

reported indicators of: acceptance, affect, autonomy, competence, control, efficacy, happiness, life satisfaction, mastery, mental well-being or wellness, optimism, personal growth, positive mental states, psychological well-being, purpose in life, quality of life, resilience, self-esteem, and subjective well-being.

For the cost-effectiveness review, literature was identified that had addressed economic evaluation or analysis, including cost allocation, cost benefit analysis, cost containment, cost effectiveness analysis, cost minimisation analysis, cost utility analysis, health care costs, health care finance, health economics & social economics. Additional outcome measures include disability adjusted life years (DALYs), quality adjusted life years (QALYs), value of life & extra health status indicators including equivalent health utility, EuroQol (EQ5D), HUI, quality of wellbeing, SF6, SF12 & SF36.

2.4.2 Exclusion Criteria

- Studies that included people under the aged of 65, except where they both included older age groups and subdivided results by age group.
- Studies that included older people undergoing treatment for a clinically diagnosed physical illness (e.g. cancer) or mental illness (e.g. dementia).
- Assessments for long-term continuing care.
- Community interventions to improve the physical and social environment that are not directly targeted at people over 65 or their carers.
- Interventions tailored to those in acute or palliative care.
- Medical or surgical interventions.
- Pre-retirement financial planning schemes.
- Specific therapeutic interventions (e.g. reminiscence therapy) covered by NICE clinical guidelines.

2.5 Data management

The searches retrieved more titles than expected. This was aggravated by some sources not being accessed by software to identify and remove duplication. Two files (Age Info & Social Care Online) containing many hits were not incorporated into the original EndNote file and screened for duplicates. When these two files eventually

arrived, abstract screening had begun and it was too late to remove duplicates with the original file. To keep to schedule, it was necessary to prioritise screening. As a result the following data bases have not been screened:

File Name(s)	Reason
Age Info Social Care Online (5500 references)	The internal NICE team conducted a limited, rapid screening of this database. This process identified 191 intervention studies, of which 23 possible inclusions were double checked by a second screener. As the external review team had already excluded or included all of these, few if any references have been missed.
NEHL – Laterlife 1 to 5 & Mental Health 1	Downloaded as a screen-saved file – not possible to access details; the file did not scroll through the abstracts; and there was nowhere to record decisions electronically. Fortunately these files appear clinical and therefore not relevant.
NEHL Later life 6	Ditto, except that the file contained material from the Joseph Rowntree Foundation (JRF). Fortunately the search of the JRF site gives confidence that no relevant files have been missed.
Undated Endnote file (699 references from NRR)	As there were no dates with these references, it was not possible to apply all the exclusion criteria.

Notwithstanding these weaknesses, the research team is confident that they have identified virtually all key studies. In total 15,388 citation titles and abstracts were screened for inclusion (Section 2.4) by one reviewer. To reduce selection bias and variance, a random sample of abstracts was considered by two reviewers, and all identified studies were checked by the second reviewer (NICE, 2006, p.25). If both did not agree, the abstract was referred to a third reviewer. By this process 248 articles were identified – 218 for the effectiveness review and 30 for the cost effectiveness review. Hard copies of almost all were retrieved, mostly through the British Medical Association Library & the University of Wales Bangor Library.

2.6 Selection of Studies

Each retrieved article was assessed for final inclusion independently by two reviewers, again applying the criteria in Section 2.4. A total of 96 papers were included – 94 in the effectiveness review and two in the cost effectiveness review.

2.7 Quality assessment

The methodological quality of included studies was assessed using the NICE methodology checklists, including one containing economic and qualitative criteria (NICE, 2006, pp. 65-110). Two reviewers initially piloted this process independently on a sample of six studies to develop consistency between them. Thereafter the primary reviewer assessed quality, and this was checked by the second reviewer. Differences were resolved by discussion or referral to a third reviewer.

Whilst accepting the hierarchy of study designs in evaluating effectiveness, the review team considered all study designs for inclusion in this new review. During the review NICE changed its procedures for reporting study designs. So this report specifies each study design, for example as RCT, NCT, CBAS, UBAS or MA (Glossary), rather than a number summarising the rigour of the study design. When included papers did not state the study design, the reviewers judged the most likely design. Table 2 specifies the grades of methodological quality used in this review.

Table 2 Grading of studies and reviews

++	Review or study fulfils all or most of the NICE criteria; where criteria have not been fulfilled, the conclusions are still thought very unlikely to alter if the study were replicated.
+	Review or study fulfils several of the NICE criteria; those criteria not fulfilled or adequately described are thought unlikely to alter the conclusions if the study were replicated.
–	Review or study fulfils few if any of the NICE criteria; the conclusions are thought likely or very likely to alter if the study were replicated.

2.8 Synthesis of effectiveness studies

Data were extracted by the primary reviewer into the standard format recommended by NICE (2006, pp.122-130) and checked by the second reviewer. For each study the main results of quality assessment and data extraction are presented in the relevant tables and narrative summary. Often the diversity of the interventions, the settings in which they were delivered, and the outcomes measured means that

pooled estimates of effect are not appropriate. The number of possible comparisons was further reduced by the absence of key statistics from several reports. Where possible average effects have been estimated from differences between groups in changes between scores before and after the intervention. Comparing change scores is better than comparing post-treatment means because many of the studies are small, and differences in baseline scores could induce bias. In calculating the standard deviations of change scores, the conservative assumption was made that baseline and post-treatment scores were uncorrelated. A fixed-effects model was used to combine effect sizes, except when the index of heterogeneity between studies was significant, necessitating a random effects model.

2.9 Cost-effectiveness review

To identify potentially eligible papers, two independent reviewers screened titles and abstracts for inclusion. Where there was disagreement the article was referred to a third reviewer. All papers that appeared to meet the inclusion criteria were retrieved for critical appraisal by a validated checklist (Drummond & Jefferson, 1996), updated to include economic modelling as recommended by NICE (2006). One researcher assessed quality and extracted data, and a second checked both.

2.10 Synthesis of cost-effectiveness studies

The heterogeneity of published economic evaluations meant that meta-analyses were not possible. Therefore a narrative summary is presented. Economic modelling was undertaken to establish the cost-effectiveness of interventions found to be effective in promoting the mental well-being of older people. A decision-analytic model was constructed to assess cost-effectiveness, following the 'reference case' set out in the *Guide to Methods of Technology Appraisal* (NICE, 2004) as far as possible. Deviations from the reference case were justified and agreed with NICE.

The costs and consequences (expressed in QALYs where possible) of alternative interventions to promote mental well-being in older people were compared. Estimates of public sector resource use were derived from nationally available data sources (Curtis & Netten, 2006; Department of Health, 2006 & 2007). The resource use of each intervention was estimated in terms of staff time, travel and consumables for each intervention, generating costs per intervention and per individual.

If enough data had been available to characterise uncertainty, the results of the economic model would have been presented as cost-utility acceptability curves. Instead we conducted univariate sensitivity and threshold analyses to test the model for robustness.

2.11 Currency conversion

For ease of comparison we have used an on-line historical currency converter (www.oanda.com/convert/fxhistory) to convert local currencies used in reviewed studies into pounds sterling when those studies quoted years for their costs.

2.12 Assessing applicability

Whether the included studies were directly applicable to the target population(s) and setting(s) was assessed using the framework in *Methods for Development of NICE Public Health Guidance* (NICE, 2006, p.30). The reviewers gave weight to whether the study was conducted in the UK and allocated studies to one of four categories:

- Likely to be applicable across a broad range of populations or settings
- Likely to be applicable across a broad range of populations or settings, assuming it is appropriately adapted
- Likely to be applicable only to populations or settings included in the studies – the success of broader application is uncertain
- Applicable only to settings or populations included in the studies.

3 Review of published evaluations of effectiveness

Of 248 papers identified for possible inclusion in this review, 96 were included. As each paper is summarised in the Evidence Tables, the narrative is brief. The largest number of papers (42) addressed exercise and physical activity – comprising 17 for mixed exercise, nine for strength & resistance exercise, five for aerobic exercise, five for walking interventions, four for Tai Chi (4) and two others.

3.1 Mixed exercise

3.1.1 Mixed exercise studies – quality assessment

Fifteen primary studies and two meta-analyses were identified that had evaluated a combination of exercise types (e.g. aerobic and strength). Table 3 summarises the findings of the quality assessment of these 17 papers.

Table 3 Quality assessment of mixed exercise studies

Author(s)	Intervention(s)	Design & quality	Comments
Annesi (2004a)	Moderate exercise programme	CBAS–	Weak design; lacks methodological detail
Annesi (2004b)	Moderate exercise programme	UBAS–	Short article lacking methodological detail
Annesi et al. (2004a)	Resistance & cardiovascular exercise programme	UBAS–	Weak design; lacks methodological detail
Arent et al. (2000)	Cardiovascular exercise, resistance training or both	MA+	Generally sound but does not include quality assessment of included papers
Clark et al. (2003)	Group-based community exercise programme	UBAS–	Lacks detail about recruitment & follow up
Cochrane et al. (1998)	Community-based exercise programme	CBAS–	Poor follow-up rates (76% of exercisers & 56% of controls); measurement properties?
Elavsky et al. (2005)	Walking & toning exercise programme	UBAS–	Lacks methodological detail
Grant et al. (2004)	Exercise programme for overweight women	NCT–	Weak design; limited sampling frame
Hardcastle & Taylor (2001)	GP exercise referral	Q+	Generally sound

Helbostad et al. (2004)	Combined training and home training in frail older people	RCT+	Generally sound but no power analysis presented, & not clear how many people were initially contacted.
Hill et al. (1993)	Endurance exercise training programme	NCT–	Weak design; not clear whether assessors were blind.
King et al. (2000)	Comparing 2 physical activity programmes	RCT+	Generally sound but not clear whether assessors were blind.
Matsouka et al. (2005)	Exercise programme for groups in Public Care Institutes for the Elderly	NCT–	Very small sample for 3 group comparisons; not clear how comparable the groups were initially.
Netz et al.(2005)	Aerobic, callisthenics or resistance training	MA+	Generally sound but does not search widely or assess quality of studies
Paw et al. (2007)	Physical exercise versus micronutrient supplementation	RCT–	Generally sound but high attrition reduced power
Stiggelbout et al. (2007)	Group-based exercise programme	RCT–	Generally sound but high attrition reduced power
Williams & Lord (1997)	Community-based group exercise	RCT–	Weak design; high attrition

3.1.2 Mixed exercise studies – findings

Two good meta-analyses report a range of effects of exercise on mental well-being, but neither assesses the quality of included studies. Arent et al. (2000; MA+) undertook a meta-analysis of 32 studies evaluating the effects of exercise (cardiovascular exercise, resistance training or a combination of both) on mood (positive and negative affect) in older adults [Glossary & Table 2 explain abbreviations like MA+]. They found that:

- The mean of 61 effect sizes on mood was 0.34 ($p<.05$) showing that, compared with controls, exercise improves mood. The mean effect size when participants exercised less than 4 days a week (0.69) was significantly more than the mean effect size (0.28) when they exercised more than 3 days a week.
- The largest effect size (0.86) was associated with exercise that was self-selected and variable in duration. Exercise that lasted more than 45 minutes was also significant (effect size = 0.36).

- All levels of duration of participation were associated with significant effect sizes (1-6 weeks = 0.48; 7-12 weeks = 0.45; > 12 weeks = 0.19).

Rigorous before-and-after comparisons found that:

- The mean effect size for mood scores was 0.36 in exercisers compared with 0.06 in controls.
- Resistance training achieved greater effect sizes (0.80) than all other types of activity (mixed cardiovascular and resistance = 0.37; cardiovascular exercise = 0.26; yoga & flexibility = 0.12; motivational control = 0.12);
- The effect sizes of high (0.29), medium (0.38) and low (0.34) intensities were all significant ($p<.01$) but did not significantly differ from each other.
- Exercise was associated with improved mood across all levels of initial health status, not just those in poorest health.
- Studies that reported an increase in cardiovascular fitness were associated with significantly larger effect sizes on mood (0.48) than those with no reporting of fitness increase (0.16), suggesting that physiological improvement might enhance mood improvement.

Netz et al. (2005; MA**+**) undertook a meta-analysis of 36 studies with 3118 participants evaluating the effects of organised physical activity (aerobic, callisthenics and resistance training) on psychological well-being (bodily well-being, emotions, self-perceptions and global well-being). They reported that:

- Aerobic exercise and resistance training had the greatest effects on psychological well-being, but did not differ significantly.
- Moderate exercise benefitted older adults' psychological well-being the most, and light exercise the least.
- In control groups the mean change for participants who had light callisthenics was similar to that for those with no exercise, but both were significantly greater than zero.
- There were significant differences between intervention and control groups for anxiety, general well-being, self-efficacy, and view of self; exercise affected physical symptoms most and life satisfaction least.

- The combination of aerobic exercise and resistance training achieved the smallest mean change in outcome; however nearly half the studies measured life satisfaction (a global measure) rather than specific outcomes.
- Longer exercise programmes (i.e. more weeks of exercise) were either less effective or ineffective.
- More sessions per week yielded inconsistent benefits, significantly positive only for anxiety and self-efficacy; longer exercise sessions (in minutes) also improved anxiety.
- The oldest participants (over 75) achieved the smallest benefits while those aged 54-64 years achieved the largest benefits.

Of the identified studies of mixed exercise for older people not reviewed by Arent et al. (2000) or Netz et al. (2005), only five were randomised trials. Williams & Lord (1997; RCT**–**) examined the effects of a free community-based exercise programme lasting 42 weeks on women over 60 living in the community in Australia. Five self-reported subjective measures of well-being (general fitness, general health, mood, outlook & sociability) were used only after the programme, by asking participants to compare their feelings then with their feelings before the trial. Though all five measures showed significant differences in favour of the intervention group, the absence of prospective baseline data makes interpretation difficult.

King et al. (2000; RCT**+**) compared two six-month physical activity programmes in the USA, one consisting of aerobic strength training and muscle toning exercises ('fit & firm') and the other consisting of stretching and flexibility exercises ('flex & stretch'). After 12 months the former improved SF-36 vitality scores while the latter improved SF-36 emotional well-being scores. There were no further significant changes at 12 months for either group.

Helbostad et al. (2004; RCT**+**) compared two types of exercise interventions for frail older people living in the community in Norway – home training (HT – home exercises plus monthly group meetings) and combined training (CT – the same home exercises plus twice-weekly group classes). The SF-36 mental health summary score

improved significantly more in the CT group than in the HT group over three months. At nine months, however, there were no differences between groups.

Both of the two negative individual studies in Table 3 were sound randomised trials with high attrition. Paw et al. (2007, RCT**–**) examined the effects of a 17-week physical exercise and micronutrient supplementation programme on the subjective well-being of frail but independent older people in the Netherlands. Participants were allocated to supervised group exercise, enriched food products, both or neither. No significant differences were found between any of the groups. Stiggelbout et al. (2007, RCT**–**) compared the effects of a 10-week exercise programme with health education in 65 to 80-year-old independent people, also in the Netherlands. The findings were not significant for the mental health component of the SF-36.

These five trials increase the number of participants meta-analysed by Arent et al. (2000) and Netz et al. (2005) by some 10 per cent. While it would be useful to update searching for, and analysis of, the original review, we judge that, in the meantime, it is acceptable to use the findings of the two original reviews. Fortunately the conclusions of these reviews about exercise for older people are both robust and positive. Preliminary calculations suggest that adding the four identified rigorous trials is unlikely to change these conclusions.

The quality of the other five identified studies that reported significant improvements relative to controls (Hill et al., 1993, NCT**–**; Cochrane et al. 1998; CBAS**–**; Annesi 2004a, CBAS**–**; Grant et al.,2004, NCT**–**; Matsouka et al. 2005, NCT**–**) is generally poor. So findings are unlikely to be robust. For example Hill et al. (1993; NCT**–**) delivered a 12-month endurance training programme for improving cardiovascular function to healthy but relatively inactive males and females over 60 in the USA. While exercise apparently improved morale, sample selection and assessment could both be biased.

Cochrane et al. (1998; CBAS**–**) delivered a 10-week community-based intervention of regular physical activities to predominantly sedentary older people in the UK and compared them with inactive controls. Though the mental health component of the

SF-36 improved in the exercise group but not in the control group, both design and differential follow-up could have introduced bias.

Annesi (2004a; CBAS**–**) compared the effects of a moderately intensive 10-week exercise programme on younger and older samples of female members of a community wellness centre in the USA. While both younger and older exercise groups improved significantly more than the no-exercise control group on four POMS scales (depression, fatigue, tension & vigour), the paper does not say enough about methods to give confidence in these findings.

Grant et al. (2004; NCT**–**) compared the effects of a 12-week exercise programme for overweight women aged 55-70 registered with a single general practice in Scotland with inactive controls. Though the exercise group reported improved life satisfaction significantly compared with controls, both design and sample selection are open to criticism. Despite methodological shortcomings the outcome measures used by Hill et al. (1993) and Grant et al. (2004) are conceptually similar. Nevertheless the resulting meta-analysis does not yield an improvement from exercise that is significant at the 5% level (the criterion used throughout this review).

Matsouka et al. (2005; NCT**–**) evaluated a free 12-week exercise programme comprising leisure activities and callisthenic exercises to improve coordination, flexibility and general strength, undertaken with varying intensity by women in Greece recruited through an advertisement. The programme had a significant effect on self-reported mood state for the two groups who exercised two or three times weekly, but not for the group who exercised once a week or the control group. However this difference may have been due to self-selection and baseline imbalance.

Another four studies (Clark et al., 2003, UBAS**–**; Annesi, 2004b, UBAS**–**; Annesi et al., 2004a, UBAS**–**; Elavsky et al., 2005, UBAS**–**) had neither control groups nor any redeeming feature to permit unbiased analysis.

Finally one insightful qualitative study (Hardcastle & Taylor, 2001; Q**+**) explores the perceptions of 15 women aged 50 to 80, newly referred by their GPs to an exercise

programme in the UK. Almost all appeared to have initiated the referral, suggesting that they had already thought about changing their activity levels. The authors describe how getting older and associated health perceptions and life events had caused their participants to resume sufficient physical activity to enhance quality of life. The older women desired a sense of belonging and usefulness. Respondents also highlighted the importance of practical support through good supervision in the gym. The authors suggest that the gym environment at a leisure centre could be seen as a social outlet that enhances sense of purpose and social inclusion. In contrast some women had experienced negative feelings through the effect of ageist social norms, like “people should not become active in later life”.

3.1.3 Mixed exercise – evidence statement 1

Two meta-analyses (Arent et al., 2000, MA+; Netz et al., 2005, MA+), together comprising 68 controlled trials from many developed countries, since augmented by four other rigorous trials in the Netherlands (2), Norway and the US, together provide strong evidence that mixed exercise programmes generally have small-to-moderate effects on mental well-being. As the reported exercise programmes cover a range of types, settings and countries, firm conclusions about the duration of programmes and the frequency of sessions are difficult. It is clear, however, that exercise of moderate intensity (not well defined in the meta-analyses) has beneficial effects on physical symptoms and psychological well-being.

The programmes evaluated were generally community-based, well organised and run by trained instructors. The findings apply to similar populations (relatively healthy and independent, and motivated to take exercise) in similar community settings in the UK. The sole qualitative study (Hardcastle & Taylor, 2001; Q+) highlights the importance of appropriate facilities and good supervision.

3.2 Strength & resistance exercise

3.2.1 Strength & resistance exercise studies – quality assessment

Nine studies were identified that had estimated the effects of strength and resistance exercise. Table 4 summarises the variable quality of these studies.

Table 4 Quality assessment for strength & resistance exercise studies

Author(s)	Intervention	Design & quality	Comments
Annesi et al. (2004b)	Resistance training	UBAS–	Lacks methodological detail
Damush & Damush (1999)	Strength training	NCT–	Lacks methodological detail
Engels et al. (1998)	Exercise with weights	NCT–	Lacks methodological detail
Jette et al. (1996)	Strength training	NCT–	Intervention weak on supervision; lacks detail about allocation
McLafferty et al (2004)	Resistance training	UBAS–	Weak design; lacks detail
Perrig-Chiello et al. (1998)	Resistance training	NCT–	Lacks methodological detail
Schechtman & Ory (2001)	Various exercises for frail older people	MA+	Generally sound but lacks quality assessment of included studies
Tsutsumi et al. (1997)	Strength training	NCT–	Lacks methodological detail
Tsutsumi et al. (1998)	Strength training	NCT–	Lacks methodological detail

3.2.2 Strength & resistance exercise studies – findings

Schechtman & Ory (2001; MA+) undertook a meta-analysis of four trials (not covered elsewhere in this review) in four different sites (with a total sample of 1733) of interventions specifically developed for frail older people living in the community. Within 'Frailty and Injuries: Cooperative Studies of Intervention Techniques' (FICSIT), the trials used four SF-36 scales to evaluate interventions to enhance endurance, resistance, endurance and resistance, balance, and resistance & balance, together with Tai Chi. Methods of delivery included training and supervision, both individual and in groups, and unsupervised exercise. In summary:

- The mean effect over all interventions including both endurance exercise programmes and flexibility programmes of up to six months was a significant small-to-moderate improvement in the SF-36 emotional health subscale.
- Neither exercise intensity or gait speed was a significant factor in changing the SF-36 emotional health score.
- Exercise programmes had no significant effects on the SF-36 pain score.

As these well-powered trials were not designed to be compared across sites, they provide limited information about the type of exercise programme that is most likely to improve well-being. Thus the main contribution of the FICSIT trials is the evidence they provide about exercise for frail older people, thus complementing the bulk of exercise research focusing on younger, healthier populations.

Five weaker and smaller studies report that strength training improved measures of mental well-being in comparison with controls. Jette et al. (1996; NCT**–**) estimated the effects of a 12 to 15-week home-based strength-training programme on able-bodied, older people in the community. Results are reported for the POMS dimensions of anger, confusion, depression, fatigue, tension and vigour. While the intervention increased vigour and reduced anger among males, it increased confusion (in the POMS sense of confused mood) among women.

Tsutsumi et al. (1997; NCT**–**) found significant improvements in the POMS dimensions of tension and vigour following strength training of high or moderate intensity in sedentary older adults. Using a very similar design Tsutsumi et al. (1998; NCT**–**) showed that the psychological benefits of highly or moderately intensive strength-training exercise significantly exceeded those of no exercise in a group of healthy but sedentary women over 60. Engels et al. (1998; NCT**–**) found that the effects on the POMS vigour scale of a 10-week supervised exercise class with and without the use of light wrist weights were significantly greater than those of the control group in older adults living independently. Perrig-Chiello et al. (1998; NCT**–**) compared the short-term (8 weeks) and long-term (1 year) effects of resistance training in normally active older people with a control group on the waiting list for training, and found only short-term benefits in ‘self-forgetfulness’.

In contrast Damush & Damush (1999; NCT**–**) found no significant improvements from an 8-week resistance training (using elastic bands) course in women recruited from two retirement communities in Southern California. Including two other studies without controls (Annesi, et al., 2004, UBAS**–**; McLafferty et al., 2004; UBAS**–**) none of these eight studies is rigorous enough for the results to be conclusive.

3.2.3 Strength & resistance exercise – evidence statement 2

Meta-analysis of four US trials that included a total of 1733 independent frail older people and used four of the SF-36 scales to evaluate similar resistance exercise interventions found a significant small-to-moderate improvement in emotional health (Schechtman & Ory, 2001; MA+). The findings are likely to be broadly applicable to frail older people in a range of settings in the UK.

Of six smaller controlled studies evaluating the benefit of resistance exercise for older people in general, five reported significant positive effects, mostly on the POMS measure (a self-reported measure of general mood over the past week). As all six were of poor quality, this finding should not be considered robust.

3.3 Aerobic exercise

3.3.1 Aerobic exercise studies – quality assessment

Five studies were identified that had examined the effects of aerobic interventions on mental well-being. Four were of low quality (Table 5) and report conflicting results.

Table 5 Quality assessment of aerobic exercise studies

Author(s)	Intervention	Design & quality	Comments
Fabre et al. (1999)	Aerobic or memory training or both	NCT–	Small sample; lacks methodological detail.
Moore & Bracegirdle (1994)	Low-intensity exercise programme	NCT–	Small sample; lacks methodological detail.
McAuley et al. (2000)	Aerobic walking versus stretching & toning	RCT+ & UBAS+	Lacks some methodological detail.
Topp & Stevenson (1994)	Aerobic exercise	NCT–	Small sample; lacks methodological detail.
Watanabe et al. (2001)	Aqua-aerobics	NCT–	Small sample; lacks methodological detail.

3.3.2 Aerobic exercise studies – findings

McAuley et al. (2000; RCT+) designed their study as an RCT to test whether aerobic brisk walking improves mental well-being in sedentary adults aged 60-75 years more

over 12 months than the comparison intervention of toning and stretching. Finding no significant differences between interventions, they combined the two groups and turned the RCT into an UBAS of physical activity. They used latent growth curve analysis to model the trajectory of participants' MUNSH and SWLS scores over 12 months. This showed significant growth in happiness and satisfaction over the exercise period of six months, followed by a significant decrease at 12 months.

The other four studies suffered from small sample sizes and methodological weaknesses. Moore & Bracegirdle (1994; NCT**–**) measured mental well-being in older women attending a six-week aerobic exercise programme at a UK day centre or engaging in 'usual activities'. They reported significantly improved happiness in the exercise group, but not the controls, but did not compare the groups directly, for example in a two-sample t test.

Topp & Stevenson (1994; NCT**–**) report no significant change in LSI-A scores in US adults over 60 who volunteered for mild or moderate exercise over 9 months. In contrast Watanabe et al. (2001; NCT**–**) found that Japanese volunteers over 60 who reported one of three different levels of energy expenditure during an aqua-aerobic intervention showed significant differences over time; POMS depression decreased among the moderate and high energy groups but not in the low energy group.

In France Fabre et al. (1999; NCT**–**) compared 8 weeks of aerobic training, training to improve memory, or both in a small group of sedentary older people. Mental well-being significantly improved in those receiving aerobic and mixed training compared with those receiving only memory training. Interesting though these four small studies are, small opportunistic samples undermine confidence in their findings.

3.3.3 Aerobic exercise – evidence statement 3

A medium-sized RCT in the US showed that both interventions – supervised aerobic brisk walking and 'toning & stretching' – generated similar trajectories of MUNSH and SWLS scores over 12 months in sedentary adults aged 60 to 75; these trajectories showed significant growth in happiness and satisfaction over the six-month exercise period, followed by a significant decrease at 12 months (McAuley et al., 2000,

RCT+). The findings are likely to be broadly applicable to similar populations in the UK.

3.4 Walking interventions

3.4.1 Walking intervention studies – quality assessment

Five studies of varying quality were identified that had evaluated walking exercise interventions. Only one study is considered methodologically sound (Table 6).

Table 6 Quality assessment of walking intervention studies

Author(s)	Intervention	Design & quality	Comments
Dubbert et al. (2002)	Nurse counselling on walking	NCT–	Lacks details about concealment & allocation
Fisher & Li (2004)	Community-based walking programme	Cluster RCT+	Generally sound, but suffered high attrition
Kjos & Etnier (2006)	Qigong versus walking	UBAS–	Weak design with small sample; lacks methodological detail
Shin (1999)	Outdoor walking programme	CBAS–	Small sample; lacks methodological detail
Starkweather (2007)	Walking at individual established pace	CBAS–	Small sample; lacks methodological detail

3.4.2 Walking intervention studies – findings

One robust trial reports improved mental well-being for older people in a walking programme. Fisher & Li (2004; Cluster RCT+) examined the effects of a 6-month, thrice-weekly, community-based walking programme in 56 heterogeneous neighbourhoods in Portland, Oregon. Using multi-level latent curve analysis the authors report that randomly-allocated intervention neighbourhoods significantly increased their SF-12 mental health and SWLS scores relative to the residual control neighbourhoods. Differential compliance with the walking intervention had no effect on these scores, either for neighbourhood groups or for individuals.

The other four studies were of low quality, lacked methodological detail and reported different results, two significant but two not. Shin (1999; CBAS–) compared the effects of an outdoor walking programme and no-treatment controls on cardio-

respiratory function, emotional state and flexibility of sedentary Korean women aged 60 to 75. Without reporting means and standard deviations the author claimed that the exercise group had better POMS anxiety-depression, mood disturbance and vigour scores than the control group after the intervention.

Dubbert et al. (2002; NCT**–**) evaluated the effects over one year of three levels of telephone follow-up on adherence to a walking programme in older adults who had initially received individual nurse counselling at a clinic. They reported no changes in SF-36 mental health component scores in any of the three groups of participants recruited through co-operating primary care providers in the USA.

Kjos & Etnier (2006; UBAS**–**) recruited 15 female volunteers between 55 and 80 from existing Qigong classes in the US to complete Qigong and walking exercises (each lasting about an hour?) three days apart. Their PANAS scores showed that positive, but not negative, affect increased immediately after exercise and then gradually decreased during the recovery period, regardless of the type of exercise.

Starkweather (2007; CBAS**–**) reported on a 10-week intervention in which student nurses taught 10 of the 20 volunteering older residents of an assisted living facility in the US to walk at 60% of their maximum heart rate and to exercise independently over 10 weeks. The POMS mood disturbance and SF-36 mental health summary scores improved significantly in the exercise group but not the control group.

3.4.3 Walking interventions – evidence statement 4

A walking programme delivered to older people in 28 heterogeneous neighbourhoods in Portland, Oregon by trained leaders three times a week over six months improved SF-12 mental health and SWLS life satisfaction scores relative to control neighbourhoods (Fisher & Li, 2004, Cluster RCT+). This cluster randomised trial recruited 279 people to the intervention group (of whom 156 completed the intervention) and compared them with 303 controls who received education only. Though recruitment and retention of participants is important for such programmes, the results are likely to be broadly applicable to similar populations in the UK.

3.5 Tai Chi

3.5.1 Tai Chi studies – quality assessment

Four studies were identified that examined the effects of Tai Chi on mental wellbeing. Three of these are methodologically sound (Table 7).

Table 7 Quality assessment of Tai Chi studies

Author(s)	Intervention	Design & quality	Comments
Kutner et al. (1997)	Tai Chi versus balance training	NCT+	Generally sound but poor response to final assessment.
Li et al. (2002)	Tai Chi programme	RCT+	Generally sound but lacks detail on randomisation
Li et al. (2004)	Tai Chi versus mild exercise	RCT+	Generally sound but high attrition
Macfarlane et al. (2005)	Tai Chi programme	CBAS–	Lacks methodological detail

3.5.2 Tai Chi studies – findings

There is conflicting evidence from the three competent studies of the effects of Tai Chi. Most positively Li et al. (2002; RCT+), reporting a six-month trial in sedentary older people, found a significant increase over time in self-esteem within the Tai Chi group compared with controls who had been put on a waiting list for Tai Chi. However the other two studies report improvements in mental well-being in both the intervention groups and the comparator groups receiving mild exercise or balance training (Kutner et al., NCT+; 1997; Li et al., 2004; RCT+).

In US adults over 70 living in the community Kutner et al. (1997; NCT+) undertook a three-armed trial comparing a 15-week group intervention of two types of exercise – Tai Chi and balance training – with an education-only control group. The only significant improvement was in a single question assessing 'sense of confidence' among people receiving Tai Chi or balance training, compared with the educational control group. However there were no significant improvements in SF-36 mental health summary scores or self-esteem scores within or between the three groups.

Building on their previous trial using waiting-list controls, Li et al. (2004) compared a community-based 24-week Tai Chi programme for inactive but healthy people over 60 in the US with a mild exercise programme – essentially Tai Chi without meditation. There were significant improvements in SF-12 mental health scores in both groups. Pooling either the active controls of Kutner et al. (1997) and Li et al. (2004) or the inactive controls of Kutner et al. (1997) and Li et al. (2002) did not lead to significant differences between Tai Chi and controls.

Less helpfully McFarlane et al. (2005; CBAS**–**) compared 15 older women living in the community in Hong Kong who accepted a 3-month Tai Chi programme with 23 who declined! Though the Tai Chi group reported a significant improvement in 'perceived well being' compared with the controls, the design is flawed and the scale does not appear to have been validated.

3.5.3 Tai Chi – evidence statement 5

Two out of three rigorous evaluations in the US showed that 3 to 6-month community-based Tai Chi programmes delivered by professionals improve differing mental health measures in older people (Li et al., 2002 & 2004; RCT+ but not Kutner et al. 1997; NCT+). However there was little difference between Tai Chi and less specific exercise programmes. Hence there is no evidence that the distinctive element of Tai Chi confers any benefit.

3.6 Other exercise

3.6.1 Other exercise studies – quality assessment

Two other exercise-based interventions did not fit into the previous categories. Both studies are of very poor quality (Table 8).

Table 8 Quality assessment of other exercise studies

Author(s)	Intervention	Design & quality	Comments
Tanaka et al. (2002)	Exercise & sleep	UBAS**–**	Very small sample; lacks methodological detail
Williams et al. (2000)	Balance training	CBAS**–**	Very small sample; lacks methodological detail

3.6.2 Other exercise studies -- findings

Williams et al (2000; CBAS**–**) undertook a pilot study to explore the effect of home-based exercises in improving balance. There were no significant findings in the achieved sample of 13 women over 73 from two senior residencies in North Carolina.

Tanaka et al. (2002; UBAS**–**) tested whether a four-week programme of exercise and short naps improved mental health in 11 Japanese people over 68. Though they report a significant improvement (decrease) in GHQ scores following the intervention, they supply no means or standard deviations and little information on participants or methods.

3.6.3 Other exercise – evidence statement 6

A US pilot study (Williams et al CBAS-) found that home-based balance-training for 13 older females had no effect on the self-reported SWLS. Another very small study (Tanaka et al., 2002, UBAS-) tested a four-week programme of exercise and short naps on 11 older people in Japan. As only these weak studies were found in this category, the conclusion is that there is no robust evidence that these forms of exercise improve mental well-being.

3.7 Health promotion

Of the 96 papers the second largest number (20) addressed health promotion & education – comprising group-based interventions (11 papers about 10 studies), mixed interventions (5) and individually-targeted interventions (4).

3.7.1 Group-based health promotion – quality assessment

Ten studies were identified that examined the effects of group-based health promotion on mental wellbeing. Only one was well conducted (Table 9).

3.7.2 Group-based health promotion & education – findings

Clark et al. (1997 & 2001; RCT**++**) undertook a 9-month trial to evaluate the effectiveness of ‘preventative occupational therapy’. The intervention comprised weekly sessions led by occupational therapists to promote positive changes in lifestyle in participants recruited from two federally subsidised apartment complexes

for older adults in the US. The intervention group was compared with two control groups, one of which undertook extra social activity, the other no extra activity. As the trial was well conducted, the marked benefits on the SF-36 mental health component score and the significant benefits in life satisfaction (LSI-Z), both relative to both control groups, can be attributed to the occupational therapy intervention. Reporting on follow-up 6 months after the end of the intervention Clark et al (2001) found that, though life satisfaction was no longer significant, the significant between-group differences for the SF36 mental health factor had been maintained. The authors warn that the findings may not generalise to older adults in other circumstances.

Table 9 Quality assessment of group-based health promotion studies

Author(s)	Intervention	Design & quality	Comments
Clark et al. (1997)	Preventive occupational therapy	RCT++	Well conducted
Clark et al. (2001)	Preventative occupational therapy	RCT++	Follows up Clark et al. (1997) six months after the end of the OT
Collins & Benedict (2006)	Community-based health promotion programme	UBAS–	Lacks methodological detail
Janssen (2004)	Leisure education	NCT–	Very small sample; lacks methodological detail
Kocken et al. (1998)	Peer-led health education	CBAS–	Lacks methodological detail, e.g. about follow up
Kremers et al. (2006)	Self management group intervention	RCT–	High levels of attrition
Lucchetti & Cerasa (2002)	Health education	UBAS–	Weak design; lacks methodological detail
Martina & Stevens (2006)	Friendship enrichment programme	CBAS–	Substantial baseline differences create potential for bias
Matuska et al. (2003)	Lifestyle education classes	UBAS–	Lacks methodological detail
Searle et al. (1998)	Leisure education	RCT–	Very small sample; lacks methodological detail
Stewart et al. (1997)	Promotion of existing community exercise classes	CBAS–	Lacks methodological detail

The quality of the remaining studies is poor and needs scepticism in interpreting results, especially as these are conflicting. For example Stewart et al. (1997; CBAS**–**) reported that a broad 6-month intervention to promote participation in moderate-intensity physical activity classes already available in the community improved self esteem, but not sense of mastery or life satisfaction, relative to the analogous control group. However they did not report means, standard deviations or effect sizes.

Kocken & Voorham (1998; CBAS**–**) compared four weeks of peer-led group education to encourage participation in healthy activities with an inactive control group. Though they reported that the programme did not affect general well-being or self-efficacy, the high attrition could have masked a true effect.

Searle et al. (1998; RCT**–**), examined the effects of leisure education in older adults who had withdrawn from a leisure activity. Though they reported that 12 experimental subjects significantly improved their locus of control relative to 10 controls, the combination of delay between randomisation & intervention and baseline differences in past education arouses suspicion. The attempt by Janssen et al. (2004; NCT**–**) to conduct a similar trial was even less well described. They found no significant differences in relevant domains of the Quality of Life Profile Senior Version.

In the Netherlands Kremers et al. (2006; RCT**–**) evaluated a 6-week group intervention to improve self-management skills in single women between 55 and 80. Though regression analysis suggests that the intervention group gained in wellbeing (SPF-IL) relative to no-treatment controls, differential attrition may have caused this.

In the Netherlands Martina and Stevens (2006; CBAS**–**) examined the effects over a year of the group-based ‘friendship enrichment programme’ on women between 53 and 86. Although they found significant improvements in self-esteem and positive & negative affect among the intervention group relative to controls, the substantial baseline differences in these variables is a major threat to their evaluation.

The 10 studies are completed by three without control groups (Collins & Benedict, 2006, UBAS**–**; Lucchetti & Cerasa, 2002, UBAS**–**; Matuska et al., 2003, UBAS**–**) or any means of drawing robust inferences.

3.7.3 Group-based health promotion – evidence statement 7

There is evidence from one well-designed longitudinal trial [Clark et al., 1997, RCT**++**; Clark et al., 2001, RCT**(++)**] that weekly educational sessions led by occupational therapists promoted and maintained positive changes in the SF-36 mental health summary score in participants recruited from two federally-subsidised apartment complexes for older adults in the US. Though the findings are likely to be broadly applicable to a similar population in the UK, the findings may not generalise to those in other circumstances (e.g. owner-occupiers & nursing home residents).

3.8 Mixed health promotion programmes

3.8.1 Mixed health promotion studies – quality assessment

Five identified studies examined the effects of mixed programmes of health promotion on mental wellbeing: four qualitative primary studies were generally sound, but the quantitative study was flawed, mainly by the lack of control group (Table 10).

Table 10 – Quality assessment of mixed health promotion studies

Author(s)	Intervention	Design & quality	Comments
Buijs et al. (2003)	Mixed programme of health promotion	Q**+**	Generally sound
Dungan et al (1996)	Mixed programme of health maintenance	UBAS**–**	Weak design without critical discussion
Greaves & Farbus (2006)	Upstream Healthy Living Centre	Mixed methods**+** (observation, interviews, focus group)	Generally sound
Wilcock (2006a)	Holistic occupation service	Q**+**	Generally sound
Wilcock (2006b)	Homeless older people project	Q**+**	Generally sound

3.8.2 Mixed health promotion studies – findings

Buijs et al. (2003; Q**+**) used qualitative interviews and focus groups to evaluate Active Living In Vulnerable Elders (ALIVE), a 10-month health promotion programme for low-income residents of 'senior apartment buildings' in Canada. The programme, delivered in the apartment buildings, included exercise classes, health information

sessions and newsletters. The most frequently reported effects of the programme were improvements in concentration and self-esteem. Some participants reported that they could undertake more ADLs. Staff also mentioned their perceptions of increased happiness among participants. The social interaction in the intervention was felt to alleviate boredom and isolation. However the disadvantages included the formation of 'cliques' within the resident population.

Greaves & Farbus (2006; MM**+**) evaluated the Upstream Healthy Living Centre, an outreach service in England for socially isolated older people in which mentors work with participants in creative, cultural and physical activities with stress on social interaction. After 6 months, but not 12, there was a marked and significant increase in the mean SF-12 mental health score. Complementing this, the qualitative analysis reported social, psychological and physical benefits.

Wilcock (2006a; Q**+**) used qualitative interviews and focus groups to evaluate 'Live Choices', an occupation service in London, England for people between 50 and 75 who have experienced homelessness and isolation. The service provides a holistic response to help older people return to occupation, gain skills and access other support. It provides a gateway into housing and other services, a range of group activities and one-to-one emotional support. The 36 client respondents (more than 50% of those eligible) reported improved psychological health. These benefits were independently observed by project staff who described clients as less withdrawn, more motivated, and increasingly sociable and confident in new skills like writing, spelling and artwork.

Wilcock (2006b; Q**+**) used qualitative interviews and quantitative questionnaires to evaluate two projects for older homeless people in London, England. The first is a day-centre-based programme of group activities to improve independence and confidence, including computing, discussion groups, German classes, talks, music group, social group, swimming, and walking. The second provided training for independent living, including budgeting, communication, coping with the challenges of moving, home making and maintenance and accessing leisure activities. Clients

and project staff reported informally that clients had benefitted cognitively, physically, psychologically & socially, and in their skills and sense of mastery.

3.8.3 Mixed health promotion – evidence statement 8

There is no quantitative evidence on the effectiveness of mixed health promotion in improving mental well-being. Four qualitative studies suggest that comprehensive health promotion programmes delivered by professionals to homeless, poor or socially isolated older people are acceptable to users and perceived to improve mental wellbeing markedly (Buijs et al., 2003, Q+; Greaves & Farbus, 2006, MM+; Wilcock, 2006a & 2006b, Q+). As three of these studies are British and the fourth Canadian, they are applicable to the UK.

3.9 Individually targeted health promotion

3.9.1 Individually targeted health promotion studies – quality assessment

Four studies were identified that examined the effects on mental wellbeing of health promotion delivered by health professionals to individual participants. All four were generally sound randomised trials (Table 11).

Table 11 Assessment of individually targeted health promotion studies

Author(s)	Intervention	Design & quality	Comments
Frieswijk et al.(2006)	Bibliotherapy correspondence course	RCT--	Generally sound trial, but the evidence for bibliotherapy is weak
Halbert et al. (2000)	One-off advice from exercise specialist	RCT+	Generally sound
Kerse et al. 2005	GP-initiated physical activity counselling	Cluster RCT+	Generally sound
Markle-Reid et al. (2006)	Health promotion through home care	RCT++	Well conducted

3.9.2 Individually targeted health promotion studies – quality assessment

In Australia Halbert et al. (2000; RCT+) evaluated the provision of 20 minutes of advice on physical activity by an exercise specialist to older patients in two general practices. Unfortunately this intervention significantly reduced mental well-being in two SF-36 dimensions – vitality (in common with the no-treatment controls) and emotional limitations on role (by a significantly greater margin than controls).

In primary care in New Zealand Kerse (2005; RCT+) estimated the effect on the mental health of older people of the Green Prescription physical activity counselling programme, in which patients receive monthly phone calls from exercise specialists. Although the SF-36 vitality subscale improved, there was no effect on the mental health subscale or summary scores (measures of mental well-being on p.24).

In the Netherlands Frieswijk et al.(2006; RCT+) evaluated a five-part bibliotherapy correspondence course in 193 volunteer older adults, who were somewhat frail. The 96 controls were put on a waiting list to receive the course later – perhaps at the risk of 'resentful demoralisation'. The authors reported improvement on the Sense of Mastery Scale in the short term, but not at six months. Furthermore there were no significant differences on the five dimensions of the SPF-IL.

In Canada Markle-Reid et al. (2006; RCT++) evaluated the addition of monthly home visits by health promotion nurses in addition to usual home care for people over 75.

There was a significant improvement over time in the SF-36 mental health summary score in the nursing group compared with the usual care group. There was also a significantly lower cost of prescription drugs in the health promotion group, apparently sufficient to offset the costs of the scheme.

Though these four studies appear inconsistent, there are two possible explanations. First the interventions are genuinely heterogeneous: not only is there little evidence for bibliotherapy in general, but the other three interventions were delivered by different combinations of professions in different ways in different countries. Secondly the findings appear consistent with the hypothesis that the response to individual health promotion is proportional to the time in contact with the professional responsible. Whatever the explanation, more rigorous research is needed.

3.9.3 Individually targeted health promotion – evidence statement 9

There is conflicting evidence from four randomised trials (Halbert et al., 2000, RCT+; Kerse, 2005, RCT+; Frieswijk et al., 2006, RCT–; Markle-Reid et al., 2006, RCT++) on the effects on mental well being.of differing health promotion interventions delivered to individuals by professionals

In Canada Markle-Reid et al. found that monthly home visits of 1 hour by health promotion nurses significantly improved SF-36 mental health summary scores in the intervention group compared with usual care. The intervention also significantly reduced the costs of prescription drugs in the health promotion group, by enough to offset the costs of the scheme.

In the Netherlands Frieswijk et al.. found that a five-part bibliotherapy correspondence course to aid self management in slightly to moderately frail older volunteers living at home, resulted in significant improvements on the Sense of Mastery Scale in the short term, but not at six months..

In New Zealand Kerse found that a primary care intervention in which independent sedentary older patients received monthly phone calls from exercise specialists improved SF-36 vitality subscale scores, but had no effect on the mental health scores.

In contrast in Australia Halbert et al found that both the provision of 20 minutes of advice on physical activity by an exercise specialist to older patients, and no treatment for controls, in two general practices significantly reduced mental well-being in two SF-36 dimensions – vitality and emotional limitations on role.

3.10 Psychological Interventions

3.10.1 Psychological intervention studies – quality assessment

Eight studies were identified that examined the effects of psychological interventions on mental well-being. Only the meta-analysis appears sound (Table 7).

3.10.2 Psychological intervention studies – findings

Pinquart & Sörensen (2001; MA**+**) undertook meta-analysis of the effectiveness of psychotherapeutic and psycho-social interventions for older adults. They included 84 studies with 3718 participants (mean age ~ 55) and analysed self-rated measures of well-being including affect, life satisfaction, morale and self-esteem. Relevant to this review are the findings for six types of psycho-social interventions:

- Activity promotion
- Cognitive training
- Control-enhancing interventions (e.g. increasing personal control over daily life)
- Psycho-education (e.g. through health promotion – cf. Sections 3.7 to 3.9)
- Relaxation
- Supportive interventions (e.g. supporting individuals towards goal attainment)

Table 12 Quality assessment of psychological intervention studies

Author(s)	**Intervention**	**Design & quality**	**Comments**
Cusack et al. (2003)	Mental fitness training	UBAS**–**	Weak design; small sample of volunteers
Funkhouser et al. (2000)	Dream telling	NCT**–**	Small sample of interested volunteers; lacks methodological detail
Houston et al. (2000)	Memory tapping (not reminiscence therapy)	UBAS**–**	Weak design but well reported
Pinquart & Sörensen (2001)	Psychosocial interventions	MA**+**	Generally sound, though inclusion criteria could have been stricter
Wilkstrom et al. (1993)	Visual stimulation	NCT**–**	Lacks methodological detail, eg uses unvalidated outcome measure

Wolinksy et al. (2006a)	Advanced Cognitive Training for Independent & Vital Elderly (ACTIVE)	RCT–	This large trial scores poorly on the CONSORT checklist, eg by analysing 2-year data by 'treatment received' – known to be at risk of bias
Wolinsky et al. (2006b)	ACTIVE	RCT–	Analyses 5-year data of Wolinsky (2006a) by 'intention to treat' – more rigorous than 'treatment received'
Zauszneiwski (1997)	Resourcefulness training	NCT–	Small sample recruited by advert; lacks methodological detail

Summary of findings:

- Cognitive training, control-enhancing interventions, psycho-education, relaxation and supportive interventions all significantly increased the subjective well-being (SWB) of participants.
- Activity promotion did not have a significant effect on SWB.
- Interventions addressed to individuals yielded larger changes in SWB than interventions in groups.
- Control-enhancing interventions were usually more effective in changing SWB.
- Psychosocial interventions in nursing homes yielded larger changes in SWB than in the community.
- Analysis of immediate and delayed follow-up of interventions showed that improvements in SWB generally remained stable and significant.
- Professional qualifications of the therapist influenced effect size: for psychosocial interventions the greatest improvement in SWB was found when the therapist had both a postgraduate degree and professional experience or training in working with older adults.
- There was no significant relationship between age and change in SWB.
- Quality of research was correlated with the effectiveness of the intervention: papers that were methodologically weak or provided little information on the psychosocial intervention were less likely to report improved SWB. (It is arguable that the meta-analysis should have addressed this weakness by adopting stricter inclusion criteria.)

Of the identified studies of psychological treatments for older people not reviewed by Pinquart & Sörensen (2001), Wolinsky et al. (2006a & b; RCT–) is the only truly randomised trial and by far the largest, with an initial sample of 2800. This trial compared a no-treatment control group of older people with three groups each

receiving a different component of Advanced Cognitive Training for Independent and Vital Elderly (ACTIVE): (a) training focusing on inductive ***reasoning*** (the ability to solve problems that follow a serial pattern); (b) training focusing on verbal episodic ***memory*** (which retrieves information acquired in a particular place and time); and (c) training focusing on the ***speed*** of identifying and locating visual information. At the 24-month follow-up the speed-training group performed better than the other three groups in minimising the number of participants suffering 'clinically important declines' on the eight SF-36 scales. However this large trial scored poorly on the CONSORT checklist (Moher et al., 2001). The principal criticism is that it analysed data by 'treatment received' rather than 'intention to treat', thus increasing the risk of bias. Another important criticism is that the trial focused on the binary outcome 'clinically important declines' rather than SF-36 scores themselves. Though this approach distinguishes between 'cases' and 'non-cases' (of depressed SF-36 scores!) it sacrifices statistical power.

In reporting the five-year follow up of this trial, completed by 1800 (64%) of the 2800 older people originally recruited, Wolinsky et al. (2006b: RCT**–**) reverted to analysing by 'intention to treat', the recommended approach (Moher et al., 2001). They reported that the speed-training group continued to perform better than the other three groups in minimising the number of 'clinically important declines' on the eight SF-36 scales. Nevertheless there is still concern over missing methodological information and the power of the statistical analysis.

As Wolinsky et al. (2006a & b) alone increases the number of participants meta-analysed by Pinquart & Sörensen (2001) by more than 50 per cent, it is important to update searching for, and analysis of, the original review. In the meantime it is prudent to use the findings of the original review. This conclusion is based on the judgement that, as published, Wolinsky et al. (2006a & b) does not fulfil the inclusion criteria of the original review. Fortunately the conclusions of Pinquart & Sörensen (2001) about cognitive training are both robust and positive. Preliminary calculations suggest that rigorous reanalysis of Wolinsky et al. (2006a & b) is likely to reinforce these conclusions rather weaken them.

The other five identified studies of psychological treatments for older people not reviewed by Pinquart & Sörensen (2001) are all small and weak. Wilkstrom et al. (1993; NCT**–**) evaluated 'visual stimulation' in which the intervention group viewed, selected, discussed and criticised four different sets of art over 16 weekly meetings. The control group discussed current news and their own hobbies. Though this comparison generated significant differences on the 'Frame of Mind Test', this is not referenced and may not have been validated. Furthermore the paper is vague about other methodological issues and difficult to follow.

Zauszneiwski (1997; NCT**–**) compared the effects of six weekly group sessions of Learned Resourcefulness Training (LRT) with diversional activities in the control group. LRT is a cognitive behavioural repertoire of skills to control the effects of disturbing thoughts and sensations on daily task performance. Though there was a significant increase in life satisfaction scores (LSI) for the intervention groups relative to the placebo groups, the use of volunteers reduces the value of this study.

Funkhouser et al. (2000; NCT**–**) evaluated the effects of volunteers already interested in dreams reporting their dreams to trained telephone interviewers. The World Health Organisation Quality of Life questionnaire (WHO-QOL-100) showed no significant differences in subjective well-being between the dream-telling group and two control groups. Again the use of volunteers reduces the value of this study.

Houston et al. (2000; UBAS**+**) evaluated a five-week intervention to interview 43 older people about their war-time experiences and 'publish' a book summarising these. Their psychological well-being, as measured by the GHQ, had improved significantly by publication, with the greatest benefit to participants low in 'attributional generality'. However the lack of a control group prevents general conclusions.

Cusack et al. (2003, UBAS**–**) delivered an eight-week Mental Fitness for Life Programme covering "goal setting, critical thinking, creativity, positive mental attitude, learning and memory" to 22 people over 50 who wanted to improve their mental abilities. There were no significant changes in the Rosenberg Self Esteem Scale. The study's use of volunteers and lack of control group undermines its value.

3.10.3 Psychological interventions – evidence statement 10

A meta-analysis (Pinquart & Sörensen, 2001; MA**+**) covering a total of 84 studies from many developed countries provides strong evidence for the effectiveness of cognitive training, control-enhancing interventions, psycho-education, relaxation and supportive interventions in improving the subjective well-being of older people. Their findings apply to similar populations – relatively healthy and independent – in similar community settings in the UK.

There is little robust evidence on the effectiveness of more specific psychological interventions – dream telling, memory tapping, mental fitness training, resourcefulness training and visual stimulation.

3.11 Computer use

3.11.1 Computer use studies – quality assessment

Four studies were identified that had examined the effect of computer training or use. All were of poor quality and require caution in interpreting findings (Table 13).

3.11.2 Computer use studies – findings

In Israel Sherer (1996; NCT**–**) evaluated group-based training in computer use within a long-term care facility in Israel. Outside the three assisted computer sessions per week participants were able to use the computer at will. The study reported a small positive effect on self-esteem (Morale & Life Satisfaction Instrument). However the measures used in the study were translated into Hebrew and not subsequently validated. Furthermore the change in outcome may be due to social interaction rather than the intervention, as the control group did not interact.

Table 13 Quality assessment of computer use studies

Authors	Intervention	Design & quality	Comments
Billipp (2001)	Computer training	NCT**–**	Very small sample; recruitment appears biased
Sherer (1996)	Group-based computer training & assisted use	NCT**–**	Lacks methodological detail & control for effect of social interaction.
White et al. (1999)	Group based computer training & support	NCT**–**	Lacks methodological detail & control for effect of social interaction

White et al. (2002)	Group based internet training	RCT–	Lacks methodological detail; intervention used different hardware in different settings

In the US Billipp (2001; NCT-) evaluated home-based computer training and assisted use for older people over 3 months. Participants were assigned to one of three groups – weekly training for three months by registered nurse, individual introductory training by registered nurse, or introductory group training. The small sample and the high risk of bias weakens confidence in the results.

In the US White et al. (1999; RCT–) conducted an evaluation of 24-hour computer access with nine hours of group-based instruction by a computer consultant delivered over six months to older people in a retirement community. Though there was no change in the Affect Balance Scale for either intervention or control group, no group comparisons are reported. Furthermore there was no control for social interaction.

Still in the US White et al. (2002; RCT–) compared two weeks of group training in computer and internet use for older people with waiting-list controls in six communal housing settings showing marked differences in hardware. This may have contributed to the lack of significant differences in Perceived Control of Life Situations or life satisfaction between intervention and control groups.

3.11.3 Computer use – evidence statement 11

Four trials examined the effect of computer training or use (Shrerer et al., 1996, NCT–; White et al., 1999, NCT–; Billipp et al., 2001, NCT-; White et al, 2002, RCT–). As all were of poor quality, there is no robust evidence on the effectiveness of computer use in improving mental well-being.

3.12 Gardening interventions

3.12.1 Gardening intervention studies – quality assessment

Three studies were identified that had examined the role of gardening in the mental wellbeing of older people (Table 14). Though two merit discussion, the third was flawed by lack of both control group and unbiased participants (Heliker et al., 2000).

Table 14 Quality assessment of gardening intervention studies

Authors	Intervention	Design & quality	Comments
Barnicle & Midden (2003)	Indoor gardening	CBAS–	Rigorously controlled before-&-after study, marred by errors in reporting
Heliker et al. (2000)	Horticulture therapy	UBAS–	Flawed design – uncontrolled study of enthusiastic volunteers
Milligan et al. (2004)	Communal gardening	Mixed methods– (inc. focus group, interviews, diaries)	Single scheme; high attrition

3.12.2 Gardening intervention studies – findings

In the US Barnicle & Midden (2003; CBAS–) estimated the effects on well-being, as measured by the Affect Balance Scale (ABS), of a 7-week indoor gardening class for residents of a long-term care facility, compared with waiting list controls from an identical facility. The authors reported that mean ABS scores had increased for the intervention group and decreased for the control group. Though they claimed that the resulting test was significant, this cannot be verified from the paper.

Milligan et al. (2004; MM–) undertook an ethnographic study of a communal gardening scheme in which older people supported by a gardener worked together to manage and maintain shared allotments. Perceptions from 19 of the original 30 participants suggested that being in and part of a garden environment in an urban area of Northern England was perceived to be therapeutic in that people felt peaceful and tranquil. Participants acknowledged that, despite keen interest, declining physical fitness was a worry, making them unable to undertake the heavier tasks.

3.12.3 Gardening interventions – evidence statement

Three studies examined the role of gardening in the mental wellbeing of older people (Barnicle & Midden 2003, CBAS-, Milligan et al, 2004, MM-, Heliker et al., 2000; UBAS-). As there were critical flaws in each study, there is no robust evidence on the effectiveness of gardening interventions in improving mental well-being.

3.13 Support groups

3.13.1 Support group studies – findings

All three studies identified under the heading of support group interventions are methodologically weak (Table 15).

Table 15 Quality assessment of support group studies

Author(s)	Intervention	Design & quality	Comments
Barnes & Bennett (1998)	Older people's forum	Q–	Lacks methodological detail about ethical approval & method of analysis
Powers & Wisocki (1997)	Single focus group	UBAS–	Very small sample; lacks methodological detail
Stewart et al. (2001)	Support groups for older widowers	UBAS–	Lacks methodological detail & coherent analysis plan.

3.13.2 Support group studies – findings

In Scotland Barnes & Bennett (1998; Q+) evaluated the Users Panel Project, which enabled house-bound older people (average age = 82) to influence service planning and provision through discussions in one of seven panels, which ran for seven years. Discussion focused on their experiences of growing older and of using health & social care services. The analysis highlights three key themes: personal development and enhanced self-esteem, empowerment of panel members, and reduced social isolation. Participants reported that they had more courage to 'voice their opinion', but were less certain of their effect on services.

In Canada Stewart et al. (2001; UBAS–) sought to evaluate a 20-week support group intervention for widowers over 55 without control group or comprehensible analysis plan. In the US Powers & Wisocki (2006; UBAS–) sought to evaluate the effect of a single 90-minute focus group discussion of worrying with only six groups totalling 21 older self-designated worriers, of whom only 12 responded.

3.13.3 Support groups – evidence statement 13

Three studies reported the effect of support groups on mental well-being (Barnes & Bennett; 1998, Q-; Stewart et al., 2001, UBAS–; Powers & Wisocki, 2006, UBAS–). As each was of poor quality, there is no robust evidence that support groups improve mental well-being.

3.14 Volunteering interventions

3.14.1 Volunteering intervention studies – quality assessment

Three studies were identified that had examined the role of volunteering. All of these were of low quality (Table 15).

Table 16 Quality assessment of volunteering intervention studies

Authors	Intervention	Design & quality	Comments
Butler (2006)	Senior Companion Programme	MM–	Lacks methodological detail
Rabiner et al. (2003)	Senior Companion Programme	CBAS–	Lacks methodological detail; very high attrition
Wheeler et al. (1998)	Volunteering	MA–	Lacks methodological detail needed to assess quality.

3.14.2 Volunteering interventions – findings

Wheeler et al (1998; MA–) undertook meta-analysis of 37 American studies of undefined origin and variable quality to estimate the effectiveness of volunteering in achieving beneficial change among older volunteers and their clients. They did not effectively address the critical task of disentangling cause and effect in the reported correlation between volunteering and improved quality of life.

In the US Senior Companions are poor volunteers over 60 who receive a small tax-free stipend, health insurance and other benefits. They help frail, housebound older people with tasks of daily living and through regular contact. Rabiner et al. (2003; CBAS–) examined the effect of 50 randomly selected Senior Companion Programmes (SCPs) on the quality of life of current clients using potential clients on the waiting list and clients of other agencies as controls. Current and other clients

had significantly higher life satisfaction than potential clients 3 months after baseline interviews, but not 6 months after that, when the response rate had fallen to 20%.

Butler (2006; MM**–**) also tried to evaluate SCPs – by interviewing a small sample of volunteers and their clients in one US rural county exactly once. This design cannot generate robust evidence, either quantitative or qualitative.

3.14.3 Volunteering interventions – evidence statement 14

Three studies reported the effect of volunteering interventions on older people (Wheeler et al., 1998, MA–; Rabiner et al., 2003, CBAS-; Butler, 2006, MM–). As all were of poor quality, there is no robust evidence on the effectiveness of volunteering in improving the mental well-being of older volunteers or older clients.

3.15 Other interventions

3.15.1 Other intervention studies – quality assessment

Eleven interventions were identified that did not fit any of the other categories in this review. All were of poor quality (Table 17).

Table 17 Quality assessment of other intervention studies

Authors	Intervention	Design & quality	Comments
Bower & Greene (1995)	Occupational therapy – 3 modes	NCT**–**	Weak design; very small sample
Columbo et al. (2006)	Pet therapy	NCT**–**	Lacks methodological detail, e.g. about validity of outcome measure.
Doric-Henry (1997)	Art therapy – pottery	CBAS**–**	Therapist is sole author; small sample; lacks methodological detail
Goldstein et al. (1997)	Video games	NCT**–**	Weak design; very small sample; lacks methodological detail
Hirakawa et al. (2003)	Home massage	NCT**–**	Weak design, e.g. assessors were not blind & participants could receive other services (e.g. acupuncture)
Hoch et al. (2001)	Sleep management	CBAS**–**	Weak design e.g. different protocols for cases & controls; very small sample
Mathey et al. (2001)	Catering redesign	CBAS**–**	Lacks methodological detail; high attrition
Noice et al	Art therapy – drama or	NCT**–**	Biased allocation of interventions; lacks

(2004)	art criticism		methodological detail
Richeson & McCullough (2003)	Pet therapy	NCT–	Weak design; very small sample; no comparative analysis!
Trefler et al. (2004)	Wheelchair modification	CBAS–	Small sample; high attrition; confusing methodological detail
Yuen (2002)	Altruistic activity	CBAS–	Weak design; very small sample

3.15.2 Other intervention studies – findings

In the US Bower & Green (1995; NCT–) compared the effects of three different types of weekly occupational therapy: (a) making holiday baskets for families with special needs (described as 'altruistic activity') (b) non-altruistic activity such as playing cards (c) discussion with occupational therapists. Though significant findings emerged, they were not consistent in direction. The very small sample size and methodological weaknesses mean that this study does not contribute useful evidence.

In the US Doric-Henry (1997; CBAS–) reported positive effects on self-esteem (Coopersmith Inventory) of weekly pottery classes over two months relative to a no-treatment control group from the same long-term care facility. However the paper lacks rigour, the study is small, and the author is both therapist and evaluator!

In the Netherlands Goldstein et al. (1997; NCT–) compared the effects on their own non-validated scale of emotional well-being of a very small group of volunteers playing video games relative to a no-treatment control group. Though they found a significant benefit to the intervention group, the study is compromised by methodological limitations.

In the US Hoch et al. (2001; CBAS–) reported a pilot project examining the effects of two sleep management interventions (one restricting time in bed, the other providing training in sleep hygiene) in older people in good mental and physical health, relative to a no-intervention group. Methodological weaknesses permit no valid inferences.

In the Netherlands Mathey et al. (2001: CBAS–) made meal-times within two wards of a nursing home more client-friendly and compared this with two similar wards in status quo. The changes focused on physical characteristics of the dining room and nursing activities and occurred three times daily for 12 months. Unfortunately

participants fell from 38 to 22 over this time. Though there were no differences on the Philadelphia Geriatric Center Morale Scale, the power of the study is very low.

In the US Yuen (2002; CBAS**–**) compared 10 residents in four long-term care facilities undertaking altruistic activity by mentoring English Second Language Students with 10 supposed controls. Though life satisfaction scores (LSI-A) for the mentoring group improved significantly, there were inherent biases in this study.

In the US Richeson & McCullough (2003; NCT**–**) examined the effects of one hour of pet therapy spread over four visits on the subjective well-being of older adults in 3 long-term care facilities. Of many methodological flaws, it was the failure to undertake comparative analysis that was fatal.

In the USA Noice et al. (2004; NCT**–**) compared two short-term art therapies – focusing on drama and visual arts – with a no-treatment control. The drama group experienced significantly greater psychological well-being for four months after the intervention than either the visual arts group or the controls. However the recruitment process ensured that only motivated individuals could join the theatre group.

In the US Trefler et al. (2004; CBA**–**) evaluated the personalised assessment and provision of wheelchair needs among older people in long-term care facilities. The social functioning scale of the SF-36 was the only scale to show significant changes between groups; this positive effect died away over several months. Methodological confusion, small sample and high attrition all undermine the value of the study

In Japan Hirakawa et al. (2005, NCT**–**) compared comprehensive home massage, performed by qualified practitioners twice or thrice a week for 12 weeks, with treatment as usual for chair-bound older people. No significant differences between groups were reported in the Subjective Satisfaction & Refreshment Scale or the Apathy Scale. However methodological limitations prevent useful inferences.

In Italy Colombo et al. (2006; NCT**–**) compared an intervention group of nursing home residents given a canary to look after with a comparator group given a pot plant and a no-intervention control group. The results suggest significant

improvement for the intervention group in life satisfaction (LEIPAD-SV). However the paper lacks methodological detail, including the validity of this measure.

3.15.3 Other interventions – evidence statement 15

A range of other interventions have been evaluated in poor quality studies. It is concluded that there is no robust evidence on the effectiveness of altruistic activity, art therapy, catering redesign in long-term care, home massage, occupational therapy, pet therapy, sleep management, video games and wheelchair modification.

4 Review of published evaluations of cost-effectiveness

Two published studies were identified for inclusion in the cost-effectiveness review; both were randomised control trials (RCT) based in the community. One was a high quality cluster RCT that focuses on mixed physical activities in the UK (Munro et al 2004; RCT**++**). The second paper was a health education intervention. It was a good quality RCT of preventative occupational therapy based in the US (Hay et al 2002; RCT**+**); the economic evaluation will be less applicable in the UK setting.

4.1 UK-based mixed exercise programme delivered in the community

Munro et al (2004; RCT**++**) conducted a cluster randomised controlled trial to establish the cost-effectiveness of a community based exercise programme for older adults (>65). Primary care practices were recruited to the study and randomised into four intervention practices and eight control practices. Older adults registered with the intervention group practices, who were not in the most active fifth, were invited to attend locally held twice-weekly exercise classes for up to two years. The classes were led by a qualified exercise instructor and lasted for 75 minutes, 45 minutes of which was physical activity. Older adults registered in the control practices received no such invitation. The main outcomes measures were SF-36 (administered at baseline, 1 and 2 years), all cause mortality, specific cause mortality (CHD, stoke, hip fracture, diabetes or mental disorder) and hospitalisation.

4.1.1 Effectiveness outcome

After adjusting for baseline differences participants in the intervention practices were estimated to have had less decline in health status over 2 years, than those in the control practices. The adjusted mean difference for the SF-36 mental health component showed intervention participants scored more highly than the control, although the difference was not significant (Table 18).

Table 18 Munro et al. (2004) – differences in SF-36 scores at 2 years

SF-36 dimension	Adjusted difference	95% CI	P Value	Estimated effect in ever exercised
Physical functioning	1.01	-0.98 to 3.0	0.36	3.9
Social functioning	1.73	-0.23 to 3.69	0.10	6.7
Physical role	3.52	-0.62 to 7.66	0.10	13.5
Emotional role	1.57	-3.24 to 6.39	0.55	6.0
Mental Health	0.98	-0.76 to 2.72	0.29	3.8
Energy	2.12	0.47 to 3.77	0.01	8.2
Pain	0.38	-1.81 to 2.57	0.80	1.5
General health perception	1.67	-0.00 to 3.34	0.06	6.4
Extended physical functioning	0.91	-1.02 to 2.83	0.41	3.5
Composite indices				
Mental health	2.65	-0.13 to 5.42	0.06	10.2
Physical health	2.95	0.17 to 5.74	0.04	11.3
Single index	0.01	0.001 to 0.02	0.03	4.08

There was no difference in all cause mortality between groups (around 12%) and little difference in specific cause mortality (5.4% intervention and 4.7% control, p value 0.25) or hospitalisation (37% intervention and 36% control, p value 0.13).

Only 26% (n = 590 of 2283) of those invited to exercise classes took up at least one session. Of those taking up at least one session, 50% attended at least 28 sessions and 30% attended at least 60 sessions during the 2 years of the intervention period. Attendance was more likely among women than men (29% versus 20%, p = <0.001), younger than older (29%, < 75 versus 23%, >75, p = 0.00). No further breakdown was given by the authors of the results either by age or gender.

4.1.2 Resource utilisation and cost data

The cost of the two-year exercise programme was reported as €267,033 (£188,205) with a mean annual reported cost of €128,302 (£90,427) [Table 19]. We have converted Euros, the currency used in the paper, into the nearest UK£ for the mid point of 2003-04, the cost year used, when the Euro was worth £0.7048.The mean cost per session was reported as €125.78 (£88) with a mean cost per participant of

€9.06 (£6). The programme costs were annuitised over five years. The costs included a baseline activity survey to identify possible participants, staff costs (facilitators, coordinators and exercise leaders), staff accommodation, hire of halls, travel and refreshments.

Table 19 Munro et al. (2004) – costs of 2-year exercise programme

Resource	Resources Used	Valuation	Central cost estimate(€)[1]	Cost estimate £[1]
Recruitment	Initial survey of activity (n-3520) with two reminders and a second survey with an invitation and leaflet (n-2283)	Commercial quotation	10,725	7,559
Facilitators	4 x 0.5 who university technicians grade D	€22969 pro rata plus ongoing costs	113,928	80,296
Coordinator start up	0.3 who RII for 6 months	€43898 pro rata plus ongoing costs	8,165	5,755
Ongoing	0.2 RII per year		21,773	15,346
Accommodation	Office space for three work stations	Standard rental	19,637	13,840
Hire of Halls	Hire of halls for 2040 sessions	Mean of €16.0 per session across 13 venues	32,645	23,008
Exercise Leaders	1337 sessions (excluding sessions undertaken by facilitators)	€31.24 per session	41,769	29,439
Travel	Average 3.5 miles per session	€0.54 per mile (actual paid)	3,824	2,695
Refreshments	Tea, coffee and biscuits per session	€8.03 per session	14,566	10,266
TOTAL COST (2 year programme)			**267,033**	**188,205**

Note 1 Rounded to nearest UK£

As no difference in use of health services between the groups was identified, the authors did not cost service use.

4.1.3 Cost effectiveness analysis

The authors describe a cost utility analysis using incremental cost effectiveness ratios. The incremental average QALY gain of 0.011 per person in the intervention population resulted in an incremental cost per QALY gain of €17,172 (95% CI = €8,300 *(£5850)* to €87,115 *(£61,399)*) or *(£12,103 (*95% CI = *£5850* to *£61,399)*)

In the sensitivity analysis the authors reported that changing assumptions concerning number of participants or varying the grade of professionals involved in recruitment, co-ordination or delivery resulted in a cost-effectiveness of the exercise programme varying between €4,739 *(£3,340)* and €32,533 *(£22,929)* per QALY. The trial does appear to be cost effective.

4.1.4 Comment

This is a well-designed cluster RCT, the recruitment process was clearly explained and the rationale was clear. However there are no details of control group, losses to follow-up or participating practices; and the cost analysis was based on the exercise programme costs only. Such missing information, particularly the costs relating to hospital admissions, affects the incremental cost effectiveness ratio.

There was a low level of adherence to the exercise programme. Although age and gender differences are mentioned by the authors, no other details were reported. The trial reported an improvement in mental health for exercisers and is therefore directly applicable to this review. However the QALY calculation was based on the SF-36 component scores, not the specific measure of mental health.

4.2 US based health education programme of preventative occupational therapy delivered in the community

The second study reviewed was cost utility analysis based on a RCT. The trial aimed to evaluate the cost-effectiveness of a 9-month preventive occupational therapy (OT) program in Well-Elderly Study (Hay et al 2002). The study population were a sub-set of a larger study reported in the effectiveness section (Clark et al 1997, 2001).

The trial was conducted in an urban area of the USA. The 218 participants in the study population were all >60, described as healthy independent people resident in

subsidised housing for older people. The poorly described intervention was led by an occupational therapist and focused on health through occupation, regularly performed activities and "how to select or perform activities to achieve a healthy and satisfying lifestyle". Group sessions were held weekly for 8 – 10 participants. In addition each participant in the control group had 9 individual sessions with the occupational therapist. The intervention was compared with an active and a passive control group. The active control received a 2.25 hour of generalised social activity led by non-professionals and the passive control no therapy. The authors reported no difference between control groups

The outcomes measures reported were the health related quality of life index, SF-36, healthcare resource use and programme costs.

4.2.1 Effectiveness outcome

The authors reported that, after the treatment, the analysis showed a statistically significant improvement in quality of life as measured by the SF36, favouring the occupational therapy group. The scores were not reported. The authors converted SF36 domain scores into health utility index (HUI) using regression-based algorithm. There were no significant differences in scores at baseline between intervention and combined controls ($p = 0.13$). After treatment change scores showed better health status in the occupational therapy group than in the controls. The average change in HUI-adjusted score was -0.2 +/- 1.1 for the occupational therapy group and -4.7 +/- 0.7 for the combined control group, yielding a difference of 4.5 ($p<0.01$).

4.2.2 Resource utilisation and cost data

The carer giver and healthcare utilisation was collected by telephone interview, costed and reported. There were no differences between groups in either so the cost analysis was based on the programme costs only. The programme costs were based on salary costs only, these are detailed in Table 20. The cost per participant of the intervention programmes was US$548, US$144 in the active control group, and US$0 in the passive control group (US$68 in the combined control group). No clear cost year was given so we have been unable to convert these costs into UK£.

Table 20 Hay et al (2002) – staff costs of intervention & control groups

Professional	Resource use	Hourly rate (US$)	Cost estimate (US$)[1]
Intervention (N=51)			
Occupational therapist	915.5 contact hours	23	21033
	300 preparation and travel time	23	6900
Intervention programme cost			**27933**
Cost per participant (programme / 51)			**548**
Active control (N=53)			
	623.5 contact time	10	6235
	140.5 meeting time	10	1405
Active control cost			**7640**
Cost per participant (programme / 53)			**144**
Passive control			
No intervention			0
Combined control (n=112)			
Active control cost only			**7640**
Cost per participant (combined control total / 112)			**68**

Note 1 Rounded to nearest US$

4.2.3 Cost-effectiveness analysis

An incremental cost-utility analysis was used to calculate costs and benefits of the interventions based on the programme costs. The incremental QALY gained in the whole intervention group over the combined control group, based on the average HUI-adjusted score was 4.5 ($p<0.01$). The incremental cost per QALY gained of occupational therapy was US$10,666 (95% CI: US$6,747 - US$25,430) over the combined controls. The incremental cost per QALY gained of occupational therapy was US$13,784 (95% CI: US$7,724 - US$57,879) over the passive control, and US$7,820 (95% CI: US$4,993 - US$18,025) over the active control.

4.2.4 Comment

Although the measure of mental health has not been reported clearly in the economic evaluation paper, it is reported in Clark et al. (2001). The generalisability of the study results to other settings was not addressed by the authors. The external validity is limited and applicability in the UK setting is unclear; not all programme costs were included and the cost year was not stated. There was no sensitivity analyses nor analysis by age band, gender or ethnicity.

4.3.1 Summary

Two published economic evaluations based on RCTs were identified for inclusion in the review. Fortunately for comparability both economic analyses were based on the SF-36 component scores, not the specific SF-36 mental health score.

One study was an RCT of a community-based mixed exercise programme for older people conducted in the UK (Munro et al 2004; RCT**+**). The intervention consisted of twice-weekly exercise classes of 75 minutes duration delivered by an exercise professional for 2 years. The mean net QALY gain of 0.011 per person in the intervention group population yielded an incremental cost per QALY of €17,172 (95% CI = €8,300 to €87,115) i.e. £12,103 (95% CI = £5850 to £61,399). The intervention was effective and probably cost effective despite low levels of adherence.

The other study was an RCT of a health education programme conducted in the US. The trial evaluated the cost-effectiveness of a 9-month preventive occupational therapy program in Well-Elderly Study (Hay et al 2002; RCT**+**). The intervention was delivered by an occupational therapist. The incremental QALY gained in the whole intervention group over the combined control group, based on the average HUI-adjusted score was 4.5 ($p<0.01$). The incremental cost per QALY gained with occupational therapy was US$10,666 (95% CI: US$6,747 - US$25,430) over the combined controls. The program was found to be cost-effective in the US. As no cost year was given, it has not been possible accurately to convert the costs into UK£.

4.3.2 Published studies of cost-effectiveness – evidence statement 16

Only two studies provided good evidence about the cost-effectiveness of interventions to improve the mental well-being of older people. First Hay et al. (2002; RCT**+**) showed that a two-hour group session of preventive advice from an occupational therapist per week is cost-effective in the USA with an incremental cost per QALY of $10,700 (95% CI $6,700 to $25,400). Secondly Munro et al (2004; RCT**+**) showed that twice-weekly 75-minute exercise classes led by qualified instructors is probably cost-effective in the UK with an incremental cost per QALY of £12,100 (95% CI = £5,800 to £61,400). While both studies are sound, one cannot be entirely confident that such sparse findings will apply to similar populations (relatively healthy, living independently, and motivated to take advice and exercise) in similar community-based settings in the UK.

5. Economic modelling of cost-effectiveness

5.1 Summary

Interventions identified in the effectiveness review which were without supporting economic evidence, were selected for inclusion in an economic model that was developed to assess their benefits (expressed in quality-adjusted life-years, QALYs) relative to their respective costs. The criteria for selection included whether the interventions had plausible effectiveness, and whether appropriate data were available. Four studies that described three interventions were considered; counselling programmes to promote physical activity (Halbert et al., 2000; Kerse et al., 2005), a community-based walking scheme (Fisher et al., 2004), and a proactive nursing health promotion intervention (Markle-Reid et al., 2006). The 6-month cost of each intervention ranged from £9.50 per participant, in the case of provision of advice on physical activity from an exercise specialist (Halbert et al., 2000), to £220 per participant in the case of the community-based walking programme (Fisher et al., 2004). Relative to their respective control groups, QALY gains ranged from 3.0 per 1000 individuals over 6-months (physical activity counselling intervention, Kerse et al., 2005) to 28.3 QALYs per 1000 individuals over 6-months (community-based walking programme, Fisher et al., 2004). A nursing health promotion intervention (Markle-Reid et al., 2006) improved health outcome (compared with control) by 4.3 QALYs per 1000 participants over 6-months.

At willingness to pay thresholds of £20,000 and £30,000 per QALY gained, the provision of advice from exercise specialists was not considered cost-effective, but this was based on the results of only one study (Halbert et al., 2000), and under the restrictions of the modelling assumptions. The provision of activity counselling or "Green Prescription", by primary care practice nurses, together with follow-up telephone support –although representing a similar intervention- was associated with a 6-month cost per QALY gained of £26,177 (Kerse et al., 2005). The provision of health promotion information by community nurses was associated with a cost per QALY gained of £45,593 over 6-months. Compared with the control group, a community-based walking intervention appeared to be most cost-effective at £7,372 and £4,915 per QALY gained (at 6 and 12-months, respectively).

5.2 Cost-effectiveness analysis

5.2.1 Introduction

Estimation of cost-effectiveness is essential for supporting decisions on resource allocation. With the exception of the two cost-effectiveness analyses detailed in section 4 of the evidence review, however, the evidence presented in the effectiveness review was largely devoid of any economic data. A significant problem in determining an economic case for interventions to promote wellbeing in older people is the lack of available data; and of the data that does exist, even less is reported in a form that lends itself to an economic model. Despite these limitations, an attempt was made to estimate the cost utility of a range of potential interventions.

5.2.2 Rationale for the model

The aim was to develop a model to generate cost per QALY estimates from data available from published sources. Given time constraints, it was not possible to model the cost utility of all the potential interventions identified in the evidence review. Rather, a systematic approach was adopted to identify the most relevant studies. Section 3.1 describes the criteria for selecting and including studies from which data were extracted to populate the model.

5.3 Development of the model

5.3.1 Selection of competing alternatives

From the interventions identified in the evidence review (see section 3), a decision on which studies were to be included in the economic model, was made by the project team on the basis of the following criteria:

1. The quality of evidence supporting each of the categories of intervention. This was guided by the evidence as summarised in the evidence statements. Interventions for which the evidence was not deemed to be supportive or was inconsistent were excluded.
2. There was sufficient data and relevant outcome measures that could be utilised by the cost-effectiveness analysis.
3. From the interventions that met criteria 1 and 2, studies that presented data on health state utility, SF-36, SF-12 or SF-6D were considered for inclusion. Authors of studies were contacted for further information, if necessary, depending on the level of detail presented in each study.

4. There were some instances of studies that were uncontrolled. These were excluded on the basis that an estimate of incremental cost-effectiveness could not be calculated without appropriate comparator group.
5. Two trials had previously been part of published economic evaluations (see section 4 of the evidence review) and were therefore not included in the model.

Four studies were identified for inclusion. Three of the studies are social interventions whereby participants receive advice and information from health or exercise professionals. Markle-Reid et al., (2005) [RCT**++**] conducted a randomised controlled trial evaluating the comparative effects of a proactive nursing health promotion intervention in addition to usual home care for frail older people compared with usual home care services alone.

Kerse et al., (2005) [cluster RCT**+**] looked at the effectiveness of the 'Green Prescription' physical activity counselling programme in increasing physical activity and quality of life in older community-dwelling people compared with usual practice care. This was a cluster randomised trial delivered mainly by GPs but also by practice nurses. Halbert et al. (2000) [RCT**+**] compared the effectiveness of advice about physical activity and a 3-month exercise plan delivered by an exercise professional compared with nutrition advice.

The remaining study was a self-care intervention. Fisher et al. (2004) [cluster RCT**+**] compared the effectiveness of a community walking programme with the provision of information and education only.

Additional details of these studies are provided in the evidence review, section 3.

5.3.2 Time horizon and perspective

Three studies reported outcomes at 6-months (Markle-Reid et al., 2006; Halbert et al., 2000; Fisher et al., 2004), while two reported outcomes at 12-month follow-up (Kerse et al., 2005; Halbert et al., 2000). For convenience of comparison, 6-month

data were modelled for one study (Kerse et al., 2005), and 12-month data were modelled for two studies (Markle-Reid et al., 2006; Fisher et al., 2004).

Results are presented for both periods, by assuming (for studies with 6-month follow-up) that the health gain achieved at 6-months would be sustained between months 6 and 12. For studies reporting only at 12-months, mean change in utility scores at 6-months was assumed to be the same as that reported for 12-months. When calculating costs, consideration was given to those costs that would be incurred independently of the time horizon of analysis (i.e. fixed costs). These were calculated separately from the variable costs when arriving at figures for 6- and 12-months.

The perspective of the Public Sector was adopted, in line with NICE's guide for the methods of public health guidance (NICE, 2006), to include the wider costs outside the NHS that are associated with many public health interventions.

5.3.3 Healthcare resource utilisation and cost analysis

Given that none of the studies presented details of resource utilisation, the cost analysis was based on identifying the key components of the interventions that may incur costs. This was accomplished by careful assessment of the methodology sections of each study, to calculate the number of units of resources consumed. These included capital, staff and consumables. These were multiplied, in turn, by unit costs which were derived from standard sources e.g. Curtis et al., (2006). Individual costs per participant were calculated for each intervention, and for calculation of the incremental cost-effectiveness ratio (ICER).

5.3.4 Health outcomes

Derivation of utility scores

None of the studies included in the economic model reported a preference-based measure of health to derive health state values used to calculate QALYs. Rather, the studies presented a profile of scores covering physical and emotional health to demonstrate the full impact of change. These were most often measured by means of the SF-36 or SF-12 questionnaires. The SF-6D was developed from the SF-36 as a single-index preference-based measure for health for use in economic evaluation (Brazier et al., 2002). The preference weightings were derived from a sample of 611

members of the UK general adult population, in line with NICE's guidance that the valuation of the impact of interventions on health-related quality of life should reflect the preferences of the general population (NICE, 2006).

The calculation of the SF-6D requires access to patient level data which was not available for the current analysis. Recently, however, Ara & Brazier (2007) developed an algorithm for estimating mean cohort preference-based SF-6D scores from the eight SF-36 dimension scores (measures of mental well-being on p.24). The ordinary least squares regression that forms the basis of the algorithm was based on a sample of 6,890 observations from a variety of sources covering a wide range of health conditions. It explained over 83% of the variance in the individual SF-6D scores with a mean absolute error of 0.040:

$$SF-6D = .343814 + .000994 * PF + .000215 * RP + .001083 * BP + .000140 * GH + .000479 * VT + .001011 * SF + .000394 * RE + .001269 * MH$$

[Roberta Ara of the School of Health and Related Research, University of Sheffield has given permission to include this regression model in the report for the benefit of PHIAC, but not for publication on the internet or other publicly accessible source, until the article describing the model has been published.]

Hanmer & Fryback (2007) also developed a regression model that calculates the SF-6D index from summary SF-36 scores:

$$SF-6D = \text{-}.01904 - .0022 * \text{Female} + .00014 * \text{Age} + .0081 * \text{MCS} + .0081 * \text{PCS}$$

where Female is the proportion of females in the cohort; Age their mean age; MCS their Mental Component Score; and PCS their Physical Component Score.

This model was developed and evaluated on a sample exceeding 20,000 adults who participated in the US Medical Expenditures Panel Survey between 2001 and 2003 (MEPS website). It was shown to explain 88% of the variance in the individual SF-6D scores.

These algorithms were applied to derive SF-6D health state utility indices to enable the calculation of cost utility estimates. However, the assumption of generalisability across populations is a limitation of conducting such an analysis.

Baseline adjustment

For each group (intervention and control), 6- and 12-month utility estimates were adjusted for differences in baseline values. This was accomplished by multiplying the mean baseline SF-6D score (across intervention and control groups) with the ratio of unadjusted 6-month to baseline SF-6D scores in each case.

Incremental cost-effectiveness ratios

In order to summarise the trade off in costs and effects between interventions, incremental cost-effectiveness (utility) ratios were calculated. In the formula presented below, if C and E represent the costs and effects, and the subscripts $_A$ and $_B$ represent two options, then:

$$ICER = \frac{\sum C_A - \sum C_B}{\sum E_A - \sum E_B} = \frac{\Delta C}{\Delta E}$$

Ideally, each intervention would be compared against one another, ranked accordingly, and tested for dominance and extended dominance, to arrive at the most cost-effective intervention(s). However, the comparator groups and baseline characteristics of each study were considered too heterogeneous for what would in effect represent an indirect comparison, and therefore ICERs were calculated for each intervention against the comparator group in each respective study.

Uncertainty and sensitivity analysis

A range of potential sources of uncertainty are likely to affect the results of the analysis, and the estimation of the ICERs. In particular, the costs were based exclusively on the costs of delivering the intervention. No additional costs (or cost offsets e.g. reduced use of primary care services) could be considered as these were not evaluated in four identified studies. The application of the algorithms, described in section 3.4.1, is another source of uncertainty – in particular the assumption that populations in the studies used to construct the algorithms (Ara et al., 2007; Hanmer & Fryback, 2007) were comparable to those included in the economic analysis. In the case of translating SF-12 to SF-6D scores, the outcome at 12-month in the trial by Fisher et al., (2004) was censored at 1.0, as the calculated value exceeded this value (1.008). The assumptions made in relation to extrapolating costs and benefits from 6-months to 12-months (and *vice versa*) are also likely to affect the ICERs.

A probabilistic sensitivity analysis was not feasible given the data limitations. Alternative scenarios around the estimation of QALYs from SF-36 or SF-12 scores could not be tested in any meaningful way, and alternative scenarios for the extrapolation of costs and benefits would require more information (which was unavailable) on the sustainability (or otherwise) of effect. The sensitivity analysis presented herein is limited to a series of threshold analyses, conducted to ascertain what cost the intervention would need to be, and what change in effectiveness is required, to produce ICERs of £20,000 and £30,000 per QALY, respectively.

5.4 Results

Tables 21 to 28 present the detailed costs and outcomes associated with each of the four interventions. Data are presented only for the time horizons reported in the original publications. Tables 29 and 30 present the ICERs for each intervention, assuming a 6- and 12-month time horizon, respectively. These are also summarised in figure 1. Table 31 concludes by reporting the results of the threshold analysis.

5.4.1 Costs

Mean (6-month) cost per person participating in each intervention ranged from £9.50 (Halbert et al., 2000) to £220 (Fisher et al., 2004). The majority of the costs comprised of salary costs associated with the training of those delivering the intervention, and the intervention itself (Tables 22, 24, 26 and 28). Consumable costs comprised the majority of the remaining costs, and included (typically), the costs of telephone or newsletter follow-up, and in the case of the community walking programme, the costs of mobile phones.

5.4.2 Outcomes

The results of the studies in terms of effectiveness data (modelled QALY gains) are presented in Tables 21, 23, 25 and 27.

Absolute increases in QALY scores from baseline to study endpoint were evident for all but one intervention (Halbert et al., 2000). Relative to their respective control groups, increases in QALY scores ranged from 3.0 QALYs (physical activity counselling intervention, Kerse et al., 2005) to 28.3 QALYs (community-based

walking programme, Fisher et al., 2004) per 1000 individuals over 6-months. A nursing health promotion intervention (Markle-Reid et al., 2006) improved health outcome (compared with control) by 4.3 QALYs (per 1000 participants).

A similar trend was apparent at the 12-month time horizon. The community-based walking programme was associated with an increase of 84.8 QALY per 1000 individuals vs. control; whereas physical activity counselling intervention generated 6.1 additional QALYs per 1000 participants, compared with its control (according to data from Kerse et al., 2005), but was inferior to control in the study by Halbert et al., (2000). The nursing health promotion intervention described by Markle-Reid et al., (2006) generated 12.8 additional QALYs per 1000 individuals over control, and over a 12-month period.

5.4.3 Cost-effectiveness

Compared with their respective control groups, the incremental cost-utility ratio at 6-months ranged from £7,372 for the community-based walking intervention (Fisher et al., 2004) to the intervention being dominated (in the case of advice from an exercise specialist, Halbert et al., 2000). These results are summarised in Tables 29 and 30. A cost-effectiveness plane, illustrating the relative placement of each intervention, is presented in Figure 1.

Each ICER represents an incremental cost-effectiveness ratio of the intervention compared with its respective control. In two studies, the comparator group consisted of an education and information intervention (Fisher et al., 2004; Halbert et al., 2000), whereas in the studies by Markle-Reid et al., 2006; and Kerse et al., 2005, the control group consisted of routine care. Comparisons among the four interventions would require an assumption of homogeneity of populations, which is unlikely to be the case.

5.4.4 Threshold analysis

Table 31 presents a series of threshold analyses to identify the total, per participant costs at which each intervention is no longer cost-effective, at thresholds of £20,000 and £30,000 per QALY gained. At 6-months, the cost of the walking intervention described by Fisher et al. (2004) must increase to £577 (from £220) for it no longer to

be cost-effective at the £20,000 per QALY threshold. The corresponding cost assuming a 12-month time horizon is £1,720. For other interventions that were associated with base-case ICERs greater than £20,000 per QALY, (e.g. nursing health promotion intervention described by Markle-Reid et al., 2000) the cost of the intervention must decrease to £85, or £255 for it to be cost effective at this threshold over 6- or 12-months, respectively.

With regards to effectiveness, the health outcomes associated with the walking intervention must decrease by 18 QALYs (per 1000 population) for it no longer to be cost-effective at 6-months at the £20,000 per QALY threshold. For interventions that are not deemed cost-effective at either the £20,000 or £30,000 per QALY gained thresholds, their effectiveness must increase to become cost-effective. This may be by as much as 6.6 QALYs (per 1000 population) in the case of the physical activity programme described by Halbert et al., (2000).

Overall, and as would be expected, the results of the threshold analysis suggest that for interventions with ICERs close to £20,000 or £30,000 per QALY, relatively small changes in either cost or effectiveness makes them cost-effective at these thresholds.

5.5.1 Conclusion & discussion

At a willingness to pay threshold of £20,000 to £30,000 per QALY gained, two of the four interventions were deemed cost-effective. The provision of advice from exercise specialist was not considered cost-effective, but this was based on the results of only one study (Halbert et al., 2000), and under the restrictions of the modelling assumptions. The provision of activity counselling or “Green Prescription”, by primary care practice nurses, together with follow-up telephone support –although representing a similar intervention- was associated with a 6-month ICER of £26,177 per QALY gained (Kerse et al., 2005). The provision of health promotion information by community nurses was associated with ICERs of £45,593 and £30,396 QALYs gained at 6 and 12-months, respectively.

Compared with the control group, a community-based walking intervention, appeared to be most cost-effective, at £7,372 and £4,915 per QALY gained (at 6 and 12-months, respectively).

The walking intervention compared favourably with published health economic analyses of other interventions that were identified to promote wellbeing in older adults. Hay et al. (2002) [RCT**+**] calculated the incremental cost per QALY gained for occupational therapists delivering a health education programme to be US$10,666 (versus control). Munro et al., (2004) [RCT**++**] calculated the incremental cost per QALY gained for a community-based mixed exercise programme for the over 65s to be €17,172 (£12,103) (see section 4 of the evidence review).

5.5.2 Economic modelling of cost-effectiveness – evidence statement 17

There are only two published economic analyses of interventions to improve the mental well-being of older people (evidence statement 16). To complement these sparse data needs economic modelling based on the integration of existing studies of effectiveness and existing sources of data about patient utilities and resource costs. The most cost-effective intervention was a thrice-weekly community-based walking programme, delivered to sedentary older people who are able to walk without assistance (Fisher & Li, 2004; Cluster RCT**+**). Modelling yielded an incremental cost per QALY of £7,400 after six months, which is comparable with the two published economic analyses.

Modelling was also used to enhance three RCTs of advice about physical activity. Such advice had an estimated incremental cost per QALY of £26,200 when modelled from Kerse et al. (2005; NCT**+**), who estimated the effects of the primary care 'green prescription' counselling programme in New Zealand. The estimated incremental cost per QALY rose to £45,600 when modelled from Markle-Reid et al. (2006; RCT**++**), who evaluated proactive health promotion by nurses in Canada in addition to usual home care for people over 75. However Halbert et al. (2000; RCT**+**) reported decreased mental well-being in response to 20 minutes of individual advice on physical activity by an exercise specialist in general practice in Australia. Thus the advice was dominated by the control group to whom no advice was given.

	Baseline mean score	Baseline utility	6-month mean score	6-month utility	Baseline-Adjusted
Intervention					
Physical Function score	19.57	0.61165908	31.65	0.67126305	0.677218008
Role physical score	27.19		56.04		
Bodily Pain	54.03		56.7		
General Health Perception	56.85		57.19		
Vitality score	28.75		39.41		
Social functioning score	50.94		71.88		
Role emotional score	73.19		83.26		
Mental Health score	64.59		71.06		
Control					
Physical Function score	19.24	0.62251147	25.92	0.66600331	0.660198013
Role physical score	27.15		49.74		
Bodily Pain	51.63		58.95		
General Health Perception	56.46		58.68		
Vitality score	33.76		40.03		
Social functioning score	45.59		67.52		
Role emotional score	83.27		82.86		
Mental Health score	74.74		73.75		

Table 21. SF-36 domain scores at baseline and at 6-months for a proactive nursing health promotion intervention versus usual care (Markle-Reid et al., 2006). These are converted to SF-6D utility scores, and adjusted for baseline differences.

Resource item	Quantity	Unit cost (£)	Total Cost (£)	Total per participant (£)	Reference	Comments
Intervention Group (n=144)						
Home visits, (median of 5 home visits per participant)	720	36 per hour	25920		Curtis et al., (2006) (p.140)	Salary cost of practice nurse per hour of home visit is £36
Telephone calls, (median of 1 x 30 minute phone call per participant).	144	0.05 per minute	216			Average BT landline call per minute
Salary cost of practice nurse making one 30 minutes phone call to 144 participants, £25 per hour*		12.50	1800		Curtis et al., (2006) (p.140)	Salary cost of practice nurse per hour is £25. Therefore, £12.50 salary cost for practice nurse making a 30min phone call
Total Intervention Cost			**27936**	**194**		
Control Group (n=144)						
Total Control Cost			**0**	**0**		

Table 22. Resource use and cost estimates for delivering a proactive nursing health promotion intervention versus usual care (Markle-Reid et al., 2006).

	Baseline mean score	Baseline utility	12-month mean score	12-month utility	Baseline-Adjusted
Intervention					
Physical Function score	59.38	0.718019667	63.67117117	0.731722401	0.732989935
Role physical score	46.6		59.59821429		
Bodily Pain	60.76612903		62.90909091		
General Health Perception	59.50403226		66.90540541		
Vitality score	52.99212598		56.37387387		
Social functioning score	77.67857143		78.29545455		
Role emotional score	74.19354839		73.15628319		
Mental Health score	77.12698413		78.27027027		
Control					
Physical Function score	60.88235294	0.720507261	60.97916667	0.722077083	0.720830576
Role physical score	47.59259259		51.05042017		
Bodily Pain	61.22058824		60.73109244		
General Health Perception	64.79166667		65.00423729		
Vitality score	55.90073529		55.79166667		
Social functioning score	78.50378788		79.55508475		
Role emotional score	68.18166667		73.72889831		
Mental Health score	76.88235294		75.33333333		

Table 23. SF-36 domain scores at baseline and at 12-months for a physical activity counselling intervention versus usual care (Kerse et al., 2005). These are converted to SF-6D utility scores, and adjusted for baseline differences.

Resource item	Quantity	Unit cost (£)	Total Cost (£)	Total per participant (£)	Reference (cost)	Comments
Intervention Group (n=130)						
Consultation with primary care doctors	111	25	2774		Curtis et al., (2006) (p. 143)	The proportion of participants seen by primary care doctor (GP) and practice nurse (PN) is counted using the ratio of 385:66 (GP:PN), extracted from Elley et al. (2004) who conducted a similar study Number of GP: 385/451 x 130 = 111 Number of PN: 66/451 x 130 = 19
Consultation with practice nurse	19	10	190		Curtis et al., (2006)	
Fax materials and billings	130	1.50	195			This unit cost is based on salary, telephone charges & cost of consumables
Telephone support by trained exercise specialists	195 hr	17.64 per hour	3439.80		Munro et al., (2004)	Based on calls made on 3 occasions over 3 months. Salary cost of exercise specialist per hour - £17.64.
Telephone costs	11700 min	0.05	585		Average British Telecommunications landline call per minute	Assuming 30min per telephone call
Newsletters	520	1.00	520			Sent out quarterly over the 12 month period
Number of participants seen by primary care doctors (follow up)	191	25	4779		Curtis et al., (2006) (p. 143)	The proportion of participants seen by primary care doctor (GP) and practice nurse (PN) is counted using the ratio of 64:14 (GP:PN), extracted from Elley et al. (2004) who conducted a similar study Number of GP: 64/78 x 233 = 191 Number of PN: 14/78 x 233 = 42
Number of participants seen by practice nurse (follow up)	42	10	418			
Total Intervention Cost			**12902**	**82**		
Control Group (n=140)						
Total Control Cost			**0**	**0**		

Table 24. Resource use and cost estimates for delivering a physical activity counselling intervention versus usual care (Kerse et al., 2005).

	Baseline mean score	Baseline utility	6-month mean score	6-month utility	Baseline-adjusted utility (6 months)	12-month mean score	12-month utility	Baseline-adjusted utility (12 months)
Intervention								
Physical Function score	84.0473	0.810365558	76.6107	0.785047998	0.782803706	75.0738	0.780470567	0.778239362
Role physical score	85.1351		70.1724			67.7365		
Bodily Pain	79.7635		76.0479			74.8456		
General Health Perception	83.8784		75.9128			74.723		
Vitality score	65.5405		64.0068			65.5168		
Social functioning score	94.9595		90.0608			89.1622		
Role emotional score	92.7973		83.5753			82.8851		
Mental Health score	80.8649		80.6689			80.1959		
Control								
Physical Function score	84.4667	0.805732219	75.2733	0.786602528	0.788864195	75.1867	0.788979441	0.791247942
Role physical score	83.6711		77.0134			77.1133		
Bodily Pain	77.6309		73.52			74.64		
General Health Perception	84.68		76.8			75.88		
Vitality score	66.7667		65.2685			66.3533		
Social functioning score	93.396		90.953			92.3267		
Role emotional score	87.9262		85.6846			86.24		
Mental Health score	81.16		82			81.3933		

Table 25. SF-36 domain scores at baseline, at 6-months, and at 12-months for a physical activity counselling intervention versus usual care (Halbert et al., 2000). These are converted to SF-6D utility scores, and adjusted for baseline differences.

Resource item	Quantity	Unit cost (£)	Total Cost (£)	Total per participant (£)	Reference	Comments
Intervention Group (n=149)						
Sessions with an exercise specialist at GP surgeries	99.33 hours	17.64	1752.24		Munro et al., (2004)	
Pamphlet	149	1.00	149			
Follow-up (n=299)						
Questionnaires	897	1.00	897		Estimated as £1 based on salary & cost of consumables	On 3 occasions (at 3, 6 and 12 months). Average unit cost per questionnaire is based on salary, printing and cost of consumables
Questionnaire mailing	598	1.00	598		Average £1 postage per large size letter (A4 size)	
Postage paid envelope	150	1.00	150			
Telephone costs	4470	0.05	223.5		British Telecommunications landline call per minute	Assuming 30min each telephone call
Total Intervention Cost			**3769.74**	**19.01**		
Control Group (n=150)						
Pamphlet	150	1.00	150			
Consultation with exercise specialist about promoting good nutrition for older adults	100 hours	17.64	1764		Munro et al., (2004)	20 minute discussions
Total Control Cost			**1914**	**12.76**		

Table 26. 12-month resource use and cost estimates for delivering a physical activity counselling intervention versus usual care (Halbert et al., 2000).

	Baseline mean score	Baseline utility	6-month mean score	6-month utility	Baseline-adjusted utility (6 months)
Intervention					
Mental component score	64.88	0.8971216	72.46	1.0*	1.0*
Physical component score	66.16		72.32		
% Female	74		74		
Age	74.03		74.03		
Control					
Mental component score	65.42	0.8969816	66.99	0.8879906	0.888059898
Physical component score	65.58		62.9		
% Female	64		64		
Age	73.94		73.94		

Table 27. SF-12 mental, and physical component scores at baseline, and at 6-months for a community-based walking intervention versus usual care (Fisher et al., 2004). These are converted to SF-6D utility scores, and adjusted for baseline differences.
*These values were censored at 1.0.

Resource item	Quantity	Unit cost (£)	Total Cost (£)	£ / participant	Reference	Comments
Intervention Group (n=279)						
# neighbourhoods with walking group	28					
# walks / week / neighbourhood	3					
Length of intervention (weeks)	26					
Hours / walking session	2					
# leaders / walking group	2					
∑ walking hours with 2 leaders / group / neighbourhood	8736	£6.42 / hour	56085		HM Revenue and Customs (2006)	£6.42 per hour = £5.35 (UK minimum wage) x 1.2 for employer's expenses
# training sessions / experienced leader	5					
# experienced leaders / 5-hour training session	5					Assume 56 leaders taught in 5 classes of <12 (each with 1 experienced leader)
Total training hours provided by 5 experienced leaders	25	£17.64 / hour	441		Munro et al. (2004) Curtis et al. (2006). Inflation Indices. p.197	Inflated to 2005-6 cost year & converted from € to £ using currency converter www.oanda.com/convert/fxhistory
Handbook for leaders			200			Assume cost ~ £200
Mobile phones for 56 leaders	56	25	1400		Nokia (2006)	Cheapest mobile phone cost 2006
Monthly newsletters for 279 participants for 6 months	1674	1	1674			Estimate cost ~ £1 (salary & consumables)
Mailing newsletters	1674	1	1674		Royal Mail (2005)	Average postage = £1 / A4 letter
Total Intervention Cost			**61474**	**220.34**		
Control Group (n=303)						
Monthly newsletters for 303 participants for 6 months	1818	1	1818			
Mailing newsletters	1818	1	1818			
Total Control Cost			**3636**	**12**		

Table 28. Resource use and cost estimates for delivering a community-based walking intervention versus usual care (Fisher et al., 2004).

	Intervention		Control				
	Costs	**Effects**	**Costs**	**Effects**	**ΔCost**	**ΔEffects**	**ICER**
Fisher et al., (2004)	£220.34	0.4742629	£12.00	0.446002853	£208.34	0.028260047	£7,372.15
Kerse et al., (2005)	£79.57	0.36306335	£0.00	0.36002351	£79.57	0.00303984	£26,176.57
Markle-Reid et al., (2006)	£194.00	0.323575821	£0.00	0.319320822	£194.00	0.004254999	£45,593.43
Halbert et al., (2000)	£9.50	0.397713149	£6.38	0.399228271	£3.12	-0.001515122	Dominated

Table 29. 6-month costs, benefits and ICER, calculated as cost per QALY gained versus control group.

	Intervention		Control				
	Costs	**Effects**	**Costs**	**Effects**	**ΔCost**	**ΔEffects**	**ICER**
Fisher et al., (2004)	£440.67	0.9742629	£24.00	0.889482759	£416.67	0.084780141	£4,914.77
Kerse et al., (2005)	£81.57	0.726126699	£0.00	0.72004702	£81.57	0.006079679	£13,417.25
Markle-Reid et al., (2006)	£388.00	0.662184825	£0.00	0.649419829	£388.00	0.012764996	£30,395.62
Halbert et al., (2000)	£19.01	0.787973916	£12.76	0.794256305	£6.25	-0.006282389	Dominated

Table 30. 12-month costs, benefits and ICER, calculated as cost per QALY gained versus control group.

Study	No longer cost-effective at £20,000 per QALY threshold		No longer cost-effective at £30,000 per QALY threshold	
	6-months	12-months	6-months	12-months
1. Cost of intervention				
Fisher et al. (2004)	£577	£1,720	£860	£2,567
Kerse et al. (2005)	£61	£122	£91	£182
Markle-Reid et al. (2006)	£85	£255	£128	£383
Halbert et al. (2000)	-	-	-	-
2. Effectiveness of intervention				
Fisher et al. (2004)	- 17.8	- 63.9	- 21.3	- 70.9
Kerse et al. (2005)	+ 0.9	- 2.0	- 0.4	- 3.4
Markle-Reid et al. (2006)	+ 5.4	+ 6.6	+ 2.2	+ 0.2
Halbert et al. (2000)	+ 1.7	+ 6.6	+ 1.6	+ 6.5

Table 31. Threshold analyses conducted by varying the 6- and 12-month (1) costs of the interventions and (2) the effectiveness of the interventions, to ascertain when thresholds of £20,000 or £30,000 per QALY are exceeded. For effectiveness, the figures represent the absolute increase (decrease) in the effectiveness of the intervention, expressed as QALYs per 1000 population.

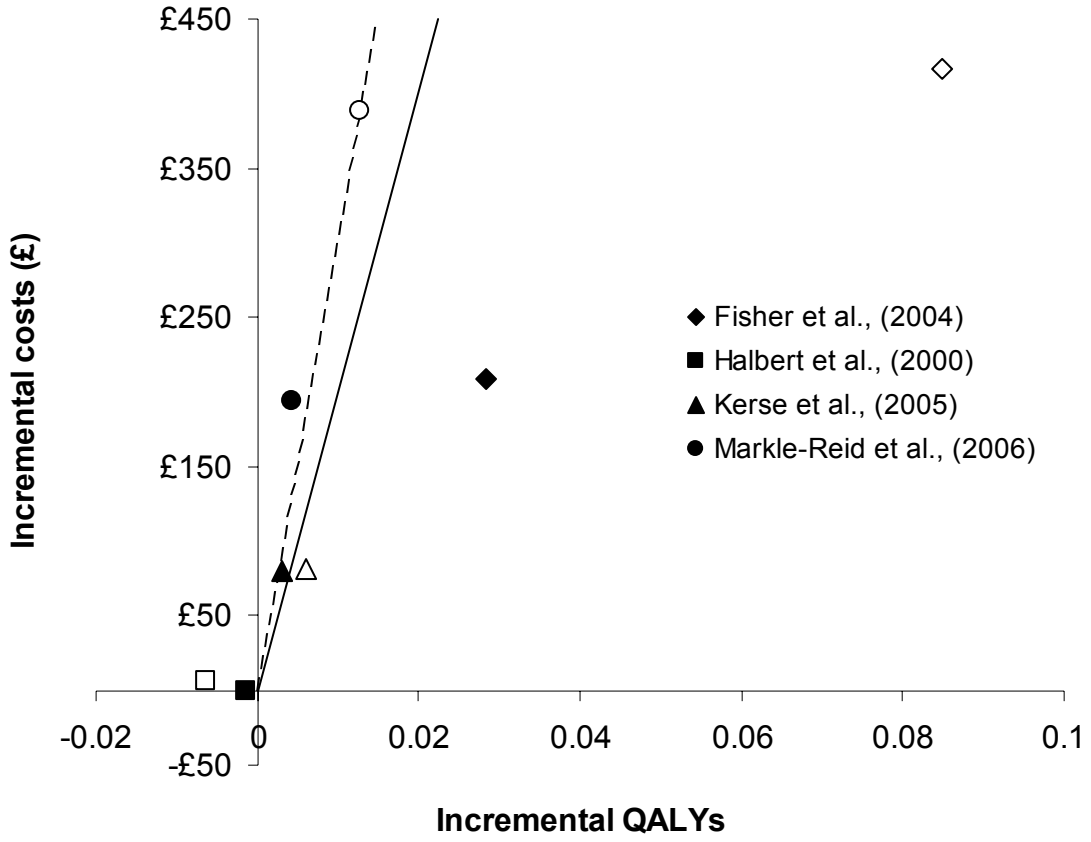


Figure 1. Cost-effectiveness plane, illustrating the relative costs and benefits (QALYs gained) for each of the four interventions at 6-months (closed symbols) and 12-months (open symbols). The solid and dashed lines represent £20,000 and £30,000 per QALY, respectively.

6 Discussion

6.1 The quality of the evidence

A systematic search of 21 data bases and 11 websites sought evidence, published between January 1993 and February 2007, of the effectiveness or cost-effectiveness of interventions to promote mental well-being in later life. So the search sought measures of positive psychological functioning. However many identified papers used the term 'well-being' while reporting only indicators of psychological dysfunction like anxiety or depression. Though such papers might provide evidence relevant to our goal, they were excluded, not least to avoid overlap with NICE clinical guidance.

In principle all study designs were considered for inclusion. In total 15,388 citation titles and abstracts were screened for relevance. By this process 248 articles were identified for further appraisal for inclusion in either review – 218 for effectiveness and 30 for cost-effectiveness. Application of inclusion criteria selected 96 papers for the review – 94 for effectiveness and two for cost-effectiveness. The 94 effectiveness papers described four meta-analyses, 13 trials of good quality (one of which generated two papers), 68 quantitative studies of poor quality (one of which generated two papers) and seven qualitative papers (including five of good quality).

Thus the review identified a broad range of interventions and reviewed evidence rarely found in traditional systematic reviews, notably qualitative research. However most included studies were of poor quality. Many used small samples that may not represent the population. Many recruited participants through advertisements, probably recruiting more motivated individuals, and again making findings less representative. The frequent use of self selection means that males and ethnic minorities are under-represented in samples. Very few included studies focused on people over 80. Few interventions were targeted at alleviating poverty, and none at gay, lesbian, bisexual or trans-sexual older people. Many studies did not include information needed to answer sub-questions, for example who delivered the intervention, or where it took place.

We divided the 95 effectiveness studies into 15 categories – six concerned with different types of exercise, three with different types of health promotion, one each

with psychological interventions, computer use, gardening, support groups and volunteering, and one residual category. Three of these categories generated useful evidence statements from meta-analyses, and another three from single rigorous trials. Unfortunately two categories, each with three rigorous trials, did not generate a useful evidence statement because the evidence from these trials was conflicting. Another six categories generated no rigorous evidence at all. While the four good qualitative studies in the final category – mixed health promotion – are helpful, they cannot estimate strength of evidence. Hence the review has generated six robust evidence statements – nos. 1 to 4 relating to exercise, no. 7 relating to health promotion and no. 10 relating to psychological interventions (pp. 9 to 11). Of the two identified cost-effectiveness papers, one added to the evidence on exercise and the other to the evidence on health promotion.

6.2 Overview of the evidence

Within the main research question, this review addresses six sub-questions:

1.3.1 *When?* – what is the frequency and duration of an effective intervention?

1.3.2 *Who?* – what are the characteristics of an effective intervener?

1.3.3 *With whom?* – are interventions that engage older people in their design and delivery more effective than those that do not?

1.3.4 *Alongside whom?* – are interventions that engage immediate family members or carers more effective than those without such engagement?

1.3.5 *Why not?* – does the intervention lead to adverse or unintended effects?

1.3.6 *How?* – what are barriers to, and facilitators of, effective implementation?

In principle this section addresses these questions for each of the six categories of intervention with robust evidence of effectiveness, and 'mixed health promotion' where the evidence is essentially qualitative. However many studies did not provide the information needed. In particular no evidence was identified to answer questions 1.3.3 and 1.3.4. Only one study provided evidence to answer question 1.3.5: in the US Jette et al. (1996; NCT**–**) reported that a three-month home-based strength-training programme had increased confusion (in the POMS sense of confused mood) among older women, but recommended stronger supervision in future (Section 3.2.2). Again only one study provided evidence to answer Question 1.3.6: in the UK

Hardcastle & Taylor (2001; Q**+**) reported that many women referred to an exercise programme felt it provided an opportunity to socialise. They also highlighted the importance of good supervision in the gym. However some women had experienced negative feelings through the impact of ageist social norms such as people should not become active in later life (Section 3.1.2).

6.2.1 Mixed exercise – when & who?

Effective? Meta-analyses (Arent et al., 2000, MA**+**; Netz et al., 2005, MA**+**) of 68 trials provide strong evidence that mixed exercise programmes generally have small-to-moderate effects on the mental well-being of older people.

When? As the reported programmes cover a range of types, settings and countries, firm conclusions about duration and frequency are difficult. However exercise of moderate intensity (not well defined in the meta-analyses) has beneficial effects on physical symptoms and psychological well-being.

Who? Though the evaluated programmes were generally run by trained leaders, this observation is suggestive rather than conclusive.

6.2.2 Strength & resistance Interventions – when & who?

Effective? Meta-analysis of four large RCTs (Schetman & Ory, 2001; MA**+**) shows that supervised strength & resistance exercise groups for frail older people yield a small-to-moderate improvement in the SF-36 emotional health subscale.

When? Though the successful programmes lasted between three and six months in the US, this observation is suggestive rather than conclusive.

Who? Though the successful programmes were supervised by trained leaders, this observation is suggestive rather than conclusive.

6.2.3 Aerobic interventions – when & who?

Effective? Robust trial (McAuley et al., 2000, RCT**+**) showed that supervised aerobic walking and ‘toning & stretching’ by sedentary older adults in the US generated significant growth in happiness and satisfaction over the exercise period.

When? Though the successful programme lasted six months, this observation is suggestive rather than conclusive.

Who? Though the successful programme was supervised by trained leaders, this observation is suggestive rather than conclusive.

6.2.4 Walking Interventions – when & who?

Effective? Robust trial randomised by neighbourhood (Fisher & Li, 2004, Cluster RCT**+**) showed that a walking programme for older people in the US improved SF-12 mental health and SWLS life satisfaction scores.

When? Though the successful programme ran thrice weekly for six months, this observation is suggestive rather than conclusive.

Who? Though the successful programme was supervised by trained leaders, this observation is in principle suggestive rather than conclusive; however the authors reinforced this recommendation from their own experience.

6.2.7 Group-based heath promotion – when & who?

Effective? Robust longitudinal trial (Clark et al., 1997, RCT**++**; Clark et al., 2001, RCT**++**) showed regular group-based health promotion initiated and maintained positive changes in the SF-36 mental health score of older adults in the US.

When? Though the successful programme ran weekly for nine months, this observation is suggestive rather than conclusive.

Who? Though the successful programme was supervised by occupational therapists, this observation is in principle suggestive rather than conclusive.

6.2.8 Mixed programme evaluations – when & who?

Effective? Four sound qualitative studies (Buijs et al., 2003, Q**+**; Greaves & Farbus, 2006, MM**+**; Wilcock, 2006a & 2006b, Q**+**) suggest that comprehensive health promotion programmes for homeless, poor or socially isolated older people are acceptable to users and perceived to improve mental wellbeing.

When? The successful programmes differ in duration (5 weeks, 10 months or unlimited), type and setting, though hardly in country because three are in the UK and one in Canada. Thus firm conclusions are difficult.

Who? Though the successful programmes were supervised by professionals, mostly in health and social care, this observation is in principle suggestive

rather than conclusive; however authors reinforced this recommendation from their own experience.

6.2.10 Psycho-social interventions – when & who?

Effective? Meta-analysis (Pinquart & Sörensen, 2001; MA**+**) of 84 studies in developed countries provides strong evidence for the effectiveness of cognitive training, control-enhancing interventions, psycho-education, relaxation and supportive interventions in improving the subjective well-being of older people.

When? Not addressed in this, otherwise comprehensive, meta-analysis.

Who? The greatest improvement in subjective well-being was found when therapists had both advanced degrees and professional experience or special training in working with older adults rather than when therapists had advanced degrees but no gerontological or geriatric experience or no advanced degree.

6.2.16

In summary we have uncovered no evidence relevant to the question *'with or alongside whom?'* and very little relevant to the questions *'why not?'* or *'how?'* Not surprisingly evidence relating to *'when?'* depends on the nature of the intervention. While evidence relating to *'who?'* also depends on the intervention, there is consensus across many categories of intervention that training is essential if the potential to achieve effectiveness and cost-effectiveness is to be exploited to the full.

6.3 Limitations of the evidence

The 95 papers included in the effectiveness review were divided into 15 categories (Chapter 3). However only six of these categories had generated robust quantitative evidence over the past 14 years. A seventh category – mixed health promotion – had contributed useful qualitative evidence. Hence the evidence in this field is variable in distribution, type and quality. So it is helpful to assess the evidence topic by topic.

Exercise and physical activity accounts for 42 of the 95 effectiveness papers. It also contributes six of the 15 categories: mixed exercise generated 17 papers including two good meta-analyses (Arent et al., 2000, MA**+**; Netz et al., 2005, MA**+**), four robust trials consistent with these meta-analyses, and one insightful qualitative paper

(Hardcastle & Taylor, 2001; Q**+**); strength & resistance exercise generated nine papers including one good meta-analysis (Schechtman & Ory, 2001; MA**+**); aerobic exercise generated five papers including one robust trial (McAuley et al., 2000, RCT**+**); and walking interventions generated five papers including one robust trial (Fisher & Li, 2004, Cluster RCT**+**). Tai Chi generated four papers including three robust trials; though these three trials show that this form of exercise is effective in absolute terms, they do not show it superior to other forms of exercise (Li et al., 2002 & 2004, RCT**+**; Kutner et al., 1997, NCT**+**). Taken together, however, the portfolio of 13 rigorous papers establishes beyond doubt that exercise and physical activity enhances mental well-being in older people.

To this evidence about the effectiveness of exercise one can add positive evidence about cost-effectiveness stemming from one primary cost-effectiveness study (Munro et al, 2004, RCT**+**) and inferences about cost-effectiveness derived from one of the 13 rigorous effectiveness papers (Fisher & Li, 2004; Cluster RCT**+**) by economic modelling (Chapter 5). In combination with information about the current provision of exercise for older people across the UK, this provides some basis for developing guidance and planning future provision.

Health promotion accounts for another 20 effectiveness papers. It also contributes three more categories – group-based health promotion (10 papers including one robust trial with two papers), mixed health promotion (six papers including four sound qualitative studies), and individually targeted health promotion (four papers including three robust trials). Unfortunately these three trials of health promotion delivered to individual older people by professionals yield evidence that is genuinely conflicting (Halbert et al., 2000, RCT**+**; Kerse, 2005, RCT**+**; Markle-Reid et al., 2006, RCT**++**). The hypothesis that the conflicting evidence arises from the different durations of these interventions, ranging from a one-off session to six home visits, merits further research. Thus the useful evidence about the effectiveness of health promotion is limited to one robust trial commending weekly educational sessions to older people by occupational therapists (the Well-Elderly study – Clark et al., 1997, RCT**++**; Clark et al., 2001, RCT**++**); and four sound qualitative studies confirming the acceptability,

and perceived benefit, of comprehensive programmes delivered by professionals to homeless, poor or socially isolated older people (Buijs et al., 2003, Q**+**; Greaves & Farbus, 2006, MM**+**; Wilcock, 2006a & 2006b, Q**+**).

To this equivocal evidence about the effectiveness of health promotion one can add evidence about cost-effectiveness, also stemming from the Well-Elderly study (Hay et al., 2002, RCT**+**) and other inferences about cost-effectiveness derived by economic modelling from each of the three rigorous trials about individual health promotion (Chapter 5). Not surprisingly the conflict of evidence about the effectiveness of individual health promotion is mirrored in the economic models derived from these trials. In summary the evidence about the value of health promotion in improving mental well-being among older people is essentially limited to one competent but single-centre trial.

Psychological interventions in well populations account for eight studies of differing qualities about differing therapies. The only robust study was a comprehensive meta-analysis showing five interventions to be effective in improving the subjective well-being of older people – cognitive training, control-enhancing interventions, psycho-education, relaxation and supportive interventions (Pinquart & Sörensen, 2001; MA**+**). Though it does not identify the optimal duration of interventions, it does identify other success criteria, including the qualifications and experience of the deliverer (the more, the better) and the age of the recipient (the younger, the better).

It is disappointing that the rigorous studies in this field are limited to exercise, health promotion and psychological interventions. In particular the search found only three papers about environmental interventions that had measured mental well-being. All three were about gardening, but none was very robust. Gardening is regularly described in the UK as important to older people (e.g. Burholt et al., 2006). Changes in later life that reduce gardening ability often trigger moves into long-term care. Thus there may be scope to develop and evaluate more environmental interventions.

Volunteering would also have yielded more papers but for failure to measure mental well-being. Voluntary work can maintain independence and provide meaningful

leisure (Wardell & Chesson, 1998). Voluntary work also benefits organisations and contributes to society. In Wales it is estimated that just under half a million people over fifty take part in voluntary activity, and the annual value of volunteering is just under £500 million (Meadows, 2004). Again there may be scope to evaluate more volunteering interventions.

The general category of social interventions yielded a wide range of studies, including four about computer use, three about support groups, two about art therapy, two about pet therapy, and others about altruistic activity, catering redesign in long-term care, home massage, occupational therapy, sleep management and video games. Unfortunately these studies tended to have weak designs and small samples, often self selected. In future research it will be important to draw rigorous conclusions, less about specific interventions, more about the intrinsic value of social interaction. As most interventions provide opportunities for socialising, there is a need to estimate the effects of getting out of the house and meeting other people.

Another disappointing feature of the included papers was the disproportionate number focusing on relatively healthy older people in the community, predominantly women. With a few exceptions (e.g. Schechtman & Ory, 2001; MA**+**) we know little about improving the mental well-being of older people in long term care facilities, notably those who are frail or over 80. Few interventions were targeted at alleviating poverty, and none at older people from ethnic or sexual minorities. Few studies answered sub-questions in full, including who delivered the intervention and where.

Most studies were of poor quality. Many used small samples, which may not represent the population of interest, and certainly lack statistical power. Many recruited participants through advertisements, probably recruiting more motivated individuals, and again making findings less representative.

With more time the external review team would have addressed the wide range of interventions and outcome measures by using meta-analysis to yield pooled estimates of generic effect sizes. When the data were not available in published papers, as often happened, they would have contacted authors for more information.

The final disappointment lay in the shortage of evidence for the cost-effectiveness review – exactly two published papers. To address the lack of economic papers, this review used four case-studies to show how economic modelling can extend studies concerned solely with effectiveness to throw light on cost-effectiveness. With more time the external review team would have extended this work to cover all the outstanding rigorous evaluations of effectiveness.

6.4 Conclusion

In summary there is a shortage of robust evidence for the effectiveness and cost-effectiveness of interventions to improve the mental well-being of older people. Better research is needed to estimate the value of most interventions. Research into cost-effectiveness is especially sparse, with little economic research even into programmes with evidence of effectiveness. Nevertheless this review has generated six positive robust evidence statements, relating to exercise and physical activity (four statements), group-based health promotion and non-clinical psychological interventions

7 References

Age Concern England (2006). *Promoting mental health and well-being in later life.* London: Age Concern.

Ara R & Brazier J (2007). Predicting the SF-6D preference-based index using mean SF-6D health dimension scores: approximating health related utilities when patient level data are not available. School of Health And Related Research, University of Sheffield (working paper).

Bowling, A. (2004). A taxonomy and overview of quality of life. In Brown, J., Bowling, A. & Flynn, T. *Models of quality of life: a taxonomy, overview and systematic review of the literature*. European Forum on Population Ageing Research.

Brazier JE, Roberts JF, Deverill MD. The estimation of a preference-based measure of health from the SF-36. J Health Econ 2002; 21:271-292.

Brown, J. & Flynn, T. (2004). The components of quality of life nominated by older people: a systematic review of the literature. In Brown, J., Bowling, A. & Flynn, T. *Models of quality of life: a taxonomy, overview and systematic review of the literature*. European Forum on Population Ageing Research.

Burholt, V., Windle, G., Naylor, D., Jones, B., Riley, B., Oades, J., Oades, R., Humphries, C.,Kennedy, D., Clifford-Jones, J. (2006). *Rural North Wales Intiative for the Development of Support for Older People (RuralWIDe): Report to Big Lottery Fund.* Age Concern Gwynedd a Mon, Caernarfon and Centre for Social Policy Research & Development, University of Wales Bangor.

Centre for Reviews & Dissemination. NHS EED review of Hay J, LaBree L, Luo R, Clark F, Carlson M, Mandel D et al (2002). Cost-effectiveness of preventive occupational therapy for independent-living older adults. *Journal of American Geriatrics Society, 50,* 1381-8. www.crd.york.ac.uk/CRDWeb/ShowRecord.asp?View=Full&ID=22002001406 accessed 5 May 2007.

Curtis L and Netten A. (2006) *Unit costs of health and social care*. Personal Social Services Research Unit, University of Kent Canterbury.: PSSRU. www.pssru.ac.uk/uc/uc2006contents.htm accessed 5 June 2007.

Department of Health (2007). *NHS Reference costs 2006-07.* www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_074072 accessed 5 June 2007.

Department of Health (2006). *Prescription Cost Analysis: England 2006.* www.ic.nhs.uk/statistics-and-data-collections/primary-care/prescriptions/prescription-cost-analysis-2006 accessed 05 June 2007

Diener, E., Suh, E., & Oishi, S. (1997). Recent findings on subjective well-being. *Indian Journal of Clinical Psychology, 24*, 25-41.

Diener, E., Suh, E. M., Lucas, R. E. and Smith, H. L. (1999). Subjective well-being; three decades of progress. *Psychological Bulletin, 2,* 276-302.

Diener, E., & Suh, E. M. (2000*). Culture and subjective well-being*. Boston MA: Massachusetts Institute of Technology.

Drummond & Jefferson (1996). Guidelines for authors and peer reviewers of economic submissions to the BMJ. BMJ 13:275-83

Elley R, Kerse N, Arroll B, Swinburn B, Ashton T, Robinson E (2004). Cost-effectiveness of physical activity counselling in general practice. NZ Med J 117 (1207):1216.

Fisher KJ, Li F (2004). A community-based walking trial to improve neighborhood quality of life in older adults: a multilevel analysis. Ann Behav Med 28(3):186-94.

Hanmer J, Fryback D (2007). A Bayesian estimation of an average SF-6D preference based score from commonly reported SF-12 statistics. In press.

Halbert JA, Silagy CA, Finucane PM, Withers RT, Hamdorf PA. Physical activity and cardiovascular risk factors: effect of advice from an exercise specialist in Australian general practice. Med J Aust. 2000; 173(2):84-7.

Hay, J., LaBree, L. et al. Cost-effectiveness of preventive occupational therapy for independent-living older adults. Journal of the American Geriatrics Society 2002; 50(8):1381-8. meps.ahrq.gov

Kerse N, Elley CR, Robinson E, Arroll B. Is physical activity counseling effective for older people? A cluster randomized, controlled trial in primary care. J Am Geriatr Soc. 2005; 53(11):1951-6.

Keyes, C. M. (2003). Complete mental health: An agenda for the 21st Century. In C. L. M. Keyes, & Haidt, J (Ed.), *Flourishing: Positive Psychology and the life well lived*. Washington: American Psychological Association.

Kinsella, K. & Velkoff, V. A. (2001). *An Aging World: 2001. International Population Reports.* USA: International Programs Center (IPC), Population Division of the U.S. Bureau of the Census.

Last, J.M. (2001). *A Dictionary of Epidemiology*, 4th ed. Oxford: OUP.

Markle-Reid M, Weir R, Browne G, Roberts J, Gafni A, Henderson S. Health promotion for frail older home care clients. J Adv Nurs. 2006; 54(3):381-95.

Meadows, P. (2004). *The Economic Contribution of Older People.* London: Age Concern.

Moher D, Schulz KF, Altman DG for the CONSORT group (2001). The CONSORT statement: revised recommendations for improving the quality of parallel-group randomised trials. *Lancet* 357: 1191-4.

National Institute for Health and Clinical Excellence (2004). *Guide to the Methods of Technology Appraisal.* www.nice.org.uk.

National Institute for Health and Clinical Excellence (2006). *Methods for Development of NICE Public Health Guidance.* www.nice.org.uk.

Munro JF, Nicholl JP, Brazier JE, Davey R, Cochrane T. Cost-effectiveness of a community based exercise programme in over 65 year olds: cluster randomised trial. J. Epidemiol. Community Health, 2004; 58:1004-1010.

NICE. Methods for development of NICE public health guidance. National Institute for Health and Clinical Excellence 2006.

Office for National Statistics (1999). *Social Focus on Older People*. London: The Stationery Office.

Office for National Statistics (2003). *Social Trends*. London: The Stationery Office.

Scharf, T., van der Meer, M., & Thissen, F. (2003). *Contextualising Adult Well-being in Europe: Report on Socio-cultural Differences in ESAW Nations.* Retrieved 2 February 2007 from www.bangor.ac.uk/esaw/summaries.htm

Scottish Executive (2005). *Well-Being and Quality of Life: Measuring the benefits of culture and sport: A literature review and think piece.* Scottish Executive Social Research: www.scotland.gov.uk/

Seymour, L. & Gale, E. (2004). Literature and policy review for the joint inquiry into mental health and well-being in later life.

Taylor, L., Taske, N., Swann, C. & Waller, S. (2007). *Public health interventions to promote positive mental health and prevent mental health disorders among older adults.* National Institute of Health and Clinical Excellence.

Wanless, D. (2004). *Securing good health for the whole population*. Norwich: HMSO. www.hm-treasury.gov.uk/media/D/3/Wanless04_summary.pdf

Wardell, F. and Chesson, R. (1998). Volunteering: a healthy choice? *Elderly Care, 10*(6), 12-13.

8 Papers included in this review

Annesi, J. J. (2004a). Mood states of formerly sedentary younger and older women at weeks 1 and 10 of a moderate exercise program. *Psychological Reports, 94*(3 Pt 2), 1337-42.

Annesi, J. J. (2004b). Relation of body fat to depression and overall mood changes over a 10-wk. course of moderate exercise in formerly sedentary older women. *Psychological Reports*, *95*(2), 447-8.

Annesi, J. J., Westcott, W. W., & Gann, S. (2004a). Preliminary evaluation of a 10-week resistance and cardiovascular exercise protocol on physiological and psychological measures for a sample of older women. *Perceptual & Motor Skills*, *98*(1), 163-70.

Annesi, J. J., Wescott, W. L., Loud, R. L. & Powers, B. S. (2004b). Effects of association and dissociation formats on resistance exercise-induced emotion change and physical self-concept in older women. *Journal of Mental Health & Aging*, *10*(2), 87-98.

Arent, S. M., Landers, D. M. & Etnier, J. L. (2000). The effects of exercise on mood in older adults: a meta-analytic review. *Journal of Aging & Physical Activity*, *8*(4), 407-430.

Barnes, M. & Bennett, G. (1998). Frail bodies, courageous voices: Older people influencing community care. *Health and Social Care in the Community 6*(2), 102-11.

Barnicle, T. & Middern, K.S. (2003). The effects of a horticulture activity program on the psychological well-being of older people in a long-term care facility. *Horttechnology*, *13*(1), 81-85.

Billipp, S. H. (2001). The psychosocial impact of interactive computer use within a vulnerable elderly population: a report on a randomized prospective trial in a home health care setting. *Public Health Nursing, 18*(2), 138-145.

Bower, M. & Greene, D. (1995). Altruistic activity and older adults in long term care facilities. *Physical and Occupational Therapy in Geriatrics*, *13*(3), 47-61.

Buijs, R. et al. (2003). Promoting participation: Evaluation of a health promotion program for low income seniors. *Journal of Community Health Nursing, 20*(2), 93-107.

Butler, S. S. (2006). Evaluating the Senior Companion Program: a mixed-method approach. *Journal of Gerontological Social Work, 47*(1-2), 45-70.

Clark, D. O., Stump, T. E. & Damush, T. M. (2003). Outcomes of an exercise program for older women recruited through primary care. *Journal of Aging & Health, 15*(3), 567-85.

Clark, F., Azen, S. P. Zemke, R., Jackson, J., Carlson, M., Mandel, D. et al. (1997). Occupational therapy for independent-living older adults: a randomized controlled trial. *JAMA, 278*(16), 1321-6.

Clark, F., Azen, S. P., Carlson, M., Mandel, D., LaBree, L., Hay, J., et al. (2001). Embedding health-promoting changes into the daily lives of independent-living older adults: Long-term follow-up of occupational therapy intervention. *Journals of Gerontology Series B-Psychological Sciences & Social Sciences*, *56*(1), 60-63.

Cochrane, T., Munro, J. et al. (1998). Exercise, physical function and health perceptions of older people. *Physiotherapy* (*London), 84*(12), 598-602.

Collins, C. C. & Benedict, J. (2006). Evaluation of a community-based health promotion program for the elderly: Lessons from Seniors CAN. *American Journal of Health Promotion, 21*(1), 45-8.

Colombo, G., Buono, M. D. et al. (2006). Pet therapy and institutionalized elderly: A study on 144 cognitively unimpaired subjects. *Archives of Gerontology & Geriatrics 42*(2), 207-216.

Cusack, S. A., Thompson, W. J. A., & Rogers, M. E. (2003). Mental fitness for life: Assessing the impact of an 8-week mental fitness program on healthy aging. *Educational Gerontology, 29*(5), 393-403.

Damush, T. M., & Damush, J. G. Jr. (1999). The effects of strength training on strength and health-related quality of life in older adult women. *Gerontologist*, *39*(6), 705-710.

Doric-Henry, L. (1997). Pottery as art therapy with elderly nursing home residents. *Art Therapy, 14*(3), 163-71.

Dubbert, P.M., et al. (2002). Effects of nurse counselling on walking for exercise in elderly primary care patients. *Journals of Gerontology Series A-Biological Sciences & Medical Sciences*, *57*(11), M733-40.

Dungan, J., Brown, A. V., et al. (1996). Health maintenance for the independent frail older adult: Can it improve physical and mental well-being? *Journal of Advanced Nursing, 23*(6), 1185-93.

Elavsky, S., McAuley, E. et al. (2005). Physical activity enhances long-term quality of life in older adults: Efficacy, esteem, and affective influences. *Annals of Behavioral Medicine, 30*(2), 138-45.

Engels, H. J., Drouin, J., et al. (1998). Effects of low-impact, moderate-intensity exercise training with and without wrist weights on functional capacities and mood states in older adults. *Gerontology 44*(4), 239-44.

Fabre, C., Masse-Biron, J., Chamari, K., Varray, A., Mucci, P., & Prefaut, C. (1999). *Archives of Gerontology & Geriatrics, 28*(1), 9-22.

Fisher, K. J. & Li, F. (2004). A community-based walking trial to improve neighborhood quality of life in older adults: A multilevel analysis. *Annals of Behavioral Medicine*, *28*(3), 186-94.

Frieswijk, N., Steverink, N., et al. (2006). The effectiveness of a bibliotherapy in increasing the self-management ability of slightly to moderately frail older people. *Patient Education & Counseling, 61*(2), 219-27.

Funkhouser, A. T., Cornu, C. M., Hirsbrunner, H. P., & Bahro, M. (2000). A preliminary study of dream-telling among mentally healthy elderly: No adverse effects on life or sleep quality. *International Journal of Geriatric Psychiatry*, *15*(10), 917-930.

Goldstein, J., Cajko, L., et al. (1997). Video games and the elderly. *Social Behavior and Personality, 25*(4), 345-52.

Grant, S., et al. (2004). The effects of a 12-week group exercise programme on physiological and psychological variables and function in overweight women. *Public Health*, *118*(1), 31-42.

Greaves, C. J. & Farbus, L. (2006). Effects of creative and social activity on the health and well-being of socially isolated older people: outcomes from a multi-method observational study. *Journal of the Royal Society of Health, 126*(3), 134-42.

Halbert, J. A., Silagy, C. A., Finucane, P. M., Withers, R. T., & Hamdorf, P. A. (2000). Physical activity and cardiovascular risk factors: Effect of advice from an exercise specialist in Australian general practice. *Medical Journal of Australia, 173*(2), 84-87.

Hardcastle, S. & Taylor, A. H. (2001). Looking for more than weight loss and fitness gain: Psychosocial dimensions among older women in a primary-care exercise-referral program. *Journal of Aging & Physical Activity*, *9*(3), 313-328.

Hay, J., LaBree, L. et al. (2002). Cost-effectiveness of preventive occupational therapy for independent-living older adults. *Journal of the American Geriatrics Society,* 50(8), 1381-8.

Helbostad, J.L., Sletvold, O., & Moe-Nilssen, R. (2004). Home training with and without additional group training in physically frail old people living at home: Effect on health-related quality of life and ambulation. *Clinical Rehabilitation*, 2004, *18*(5), 498-508.

Heliker, D., Chadwick, A., & O'Connell, T. (2000). The meaning of gardening and the effects on perceived well being of a gardening project on diverse populations of elders. *Activities, Adaptation and Aging*, *24*(3), 35-56.

Hill, R. D., Storandt, M., & Malley, M. (1993). The impact of long-term exercise training on psychological function in older adults. *Journal of Gerontology*, *48*(1), 12-17.

Hirakawa, Y., Masuda, Y. et al. (2005). Effects of home massage rehabilitation therapy for the bed-ridden elderly: A pilot trial with a three-month follow-up. *Clinical Rehabilitation, 19*(1), 20-7.

Hoch, C. C., Reynolds, C. F., III, Buysse, D. J., Monk, T. H., Nowell, P., Begley, A. E. et al. (2001). Protecting sleep quality in later life: A pilot study of bed restriction and sleep hygiene. *Journals of Gerontology*: *Series B: Psychological Sciences and Social Sciences*, *56B*(1), 52-59.

Houston, D. M., McKee, K. J., & Wilson, J. (2000). Attributional style, efficacy, and the enhancement of well-being among housebound older people. *Basic and Applied Social Psychology*, *22*(4), 309-317.

Janssen, M.A. (2004). The use of leisure education in assisted living facilities. *American Journal of Recreation Therapy*, *3*(4), 25-30.

Jette, A. M., Harris, B. A. et al. (1996). A home-based exercise program for nondisabled older adults. *Journal of the American Geriatrics Society, 44*(6), 644-9.

Kerse, N., Elley, C. et al. (2005). Is physical activity counselling effective for older people? A cluster randomized, controlled trial in primary care. *Journal of the American Geriatrics Society, 53*(11), 1951-1956.

King, A. C., Pruitt, L. A., Phillips, W., Oka, R., Rodenburg, A., & Haskell, W. L. (2000). Comparative effects of two physical activity programs on measured and perceived physical functioning and other health-related quality of life outcomes in older adults. *Journals of Gerontology Series A-Biological Sciences & Medical Sciences, 55*(2), M74-83.

Kjos, V. & Etnier, J. L. (2006). Pilot study comparing physical and psychological responses in medical qigong and walking. *Journal of Aging & Physical Activity, 14*(3), 241-53.

Kocken, P. L. &. Voorham, A. J. (1998). Effects of a peer-led senior health education program. *Patient Education & Counselling, 34*(1), 15-23.

Kremers, I. P. et al. (2006). Improved self-management ability and well-being in older women after a short group intervention. *Aging and Mental Health*, *10*(5), 476-484.

Kutner, N. G., Barnhart, H. et al. (1997). Self-report benefits of Tai Chi practice by older adults. *Journals of Gerontology Series B-Psychological Sciences & Social Sciences, 52*(5), 242-6.

Li, F. et al. (2002). Tai Chi as a means to enhance self-esteem: A randomized controlled trial. *Journal of Applied Gerontology*, *21*(1), 70-89.

Li, F. et al. (2004). Tai chi and self-rated quality of sleep and daytime sleepiness in older adults: A randomized controlled trial. *Journal of the American Geriatrics Society, 52*(6), 892-900.

Lucchetti, M., & Cerasa, F. (2002). Performance evaluation of health promotion in aging. *Educational Gerontology*, *28*(1), 1-13.

Macfarlane, D. J., Chou, K. et al. (2005). Effects of Tai Chi on the physical and psychological well-being of Chinese older women. *Journal of Exercise Science and Fitness 3*(2), 87-94.

Markle-Reid, M., Weir, R. et al. (2006). Health promotion for frail older home care Clients. *Journal of Advanced Nursing, 54*(3), 381-95.

Martina, C. & Stevens, N. (2006). Breaking the cycle of loneliness? Psychological effects of a friendship enrichment program for older women. *Aging & Mental Health, 10*(5), 467-475.

Mathey, M. F., Vanneste, V. G., de Graaf, C., de Groot, L. C., & van Staveren, W. A. (2001). Health effect of improved meal ambiance in a Dutch nursing home: A 1-year intervention study. *Preventive Medicine, 32*(5), 416-423.

Matsouka, O., Kabitsis, C. et al. (2005). Mood alterations following an indoor and outdoor exercise program in healthy elderly women. *Perceptual & Motor Skills, 100*(3 Pt 1), 707-15.

Matuska, K., Giles-Heinz, A., Flinn, N., Neibor, M., & Bass-Haugen, J. (2003). Outcomes of a pilot occupational therapy wellness program for older adults. *American Journal of Occupational Therapy, 57*(2), 220-224.

McAuley, E., Blissmer, B., Marquez, D. X., Jerome, G. J., Kramer, A. F., & Katula, J. (2000). Social relations, physical activity, and well-being in older adults. *Preventive Medicine*, *31*(5), 608-617.

McLafferty, C. L. Jr., Wetzstein, C. J., & Hunter, G. R. (2004). Resistance training is associated with improved mood in healthy older adults. *Perceptual & Motor Skills*, *98*(3 Pt 1), 947-57.

Milligan, C., Gatrell, A., & Bingley, A. (2004) Cultivating health: Therapeutic landscapes and older people in Northern England. *Social Science & Medicine*, *58*(9), 1781-93.

Moore, C., & Bracegirdle, H. (1994). The effects of a short-term, low-intensity exercise programme on the psychological wellbeing of community-dwelling elderly women. *British Journal of Occupational Therapy*, *57*(6), 213-216.

Munro, J. F., Nicholl, J. P. et al. (2004). Cost effectiveness of a community based exercise programme in over 65 year olds: cluster randomised trial. *Journal of Epidemiology & Community Health* 58(12): 1004-10.

Netz, Y., Wu, M. J., et al. (2005). Physical activity and psychological well-being in advanced age: a meta-analysis of intervention studies. *Psychology & Aging, 20*(2), 272-84.

Noice, H., Noice, T. & Staines, G. (2004). A short-term intervention to enhance cognitive and affective functioning in older adults. *Journal of Aging & Health*, *16*(4), 562-85.

Paw, M. J. M., de Jong, N., Schouten, E. G., van Staveren, W. A. & Kok, F. J. (2007). Physical exercise or micronutrient supplementation for the wellbeing of the frail elderly? A randomised controlled trial. *British Journal of Sports Medicine, 36,* 126-131.

Perrig-Chiello, P., et al. (1998). The effects of resistance training on well-being and memory in elderly volunteers. *Age & Ageing*, *27*(4), 469-75.

Pinquart, M., & Sorensen, S. (2001). How effective are psychotherapeutic and other psychosocial interventions with older adults? A meta-analysis. *Journal of Mental Health & Aging, 7*(2), 207-243.

Powers, C. B. & Wisocki, P. A. (1997) An examination of the therapeutic benefits of focus groups on elderly worriers. *International Journal of Aging & Human Development*, *45*(2), 159-167.

Rabiner, D. J., et al. (2003). The impact of the senior companion program on quality of life outcomes for frail older adults and their families. *Home Health Care Services Quarterly, 22*(4), 1-26.

Richeson, N. E. & McCullough, W. T. (2003). A therapeutic recreation intervention using animal-assisted therapy: Effects on the subjective well-being of older adults. *Annual in Therapeutic Recreation*, *64*(*20 ref).*

Schechtman, K. B., & Ory, M. G. (2001). The effects of exercise on the quality of life of frail older adults: a preplanned meta-analysis of the FICSIT trials. *Annals of Behavioral Medicine*, *23*(3), 186-97.

Searle, M. S. et al. (1998). Examining the long term effects of leisure education on a sense of independence and psychological well-being among the elderly. *Journal of Leisure Research*, *30*(3), 331-40.

Sherer, M. (1996). The impact of using personal computers on the lives of nursing home residents. *Physical & Occupational Therapy in Geriatrics*, *14*(2), 13-31.

Shin, Y. (1999). The effects of a walking exercise program on physical function and emotional state of elderly Korean women. *Public Health Nursing*, *16*(2), 146-54.

Starkweather, A. R. (2007). The effects of exercise on perceived stress and IL-6 levels among older adults. *Biological Research for Nursing, 8*(3), 186-94.

Stewart, A. L. et al. (1997). Evaluation of CHAMPS, a physical activity promotion program for older adults. *Annals of Behavioral Medicine*, *19*(4), 353-61.

Stewart, M., Craig, D., & MacPherson, K. (2001). Promoting positive affect and diminishing loneliness of widowed seniors through a support intervention. *Public Health Nursing*, *18*(1), 54-63.

Stiggelbout, M., et al. (2007). Once a week is not enough: Effects of a widely implemented group based exercise programme for older adults; A randomised controlled trial. *Journal of Epidemiology & Community Health*, *58*(2), 83-88.

Tanaka, H., Taira, K., Arakawa, M., Urasaki, C., Yamamoto, Y., Okuma, H., Uezu, E. et al. (2002). Short naps and exercise improve sleep quality and mental health in the elderly. *Psychiatry & Clinical Neurosciences, 56*(3), 233-4.

Topp, R., & Stevenson, J. S. (1994). The effects of attendance and effort on outcomes among older adults in a long-term exercise program. *Research in Nursing & Health*, *17*(1), 15-24.

Trefler, E., et al. (2004). Outcomes of wheelchair systems intervention with residents of long-term care facilities. *Assistive Technology*, *16*(1), 18-27.

Tsutsumi, T., et al., (1998). Comparison of high and moderate intensity of strength training on mood and anxiety in older adults. *Perceptual & Motor Skills*, *87*(3 Pt 1), 1003-11.

Tsutsumi, T., et al. (1997). Physical fitness and psychological benefits of strength training in community dwelling older adults. *Applied Human Science*, *16*(6), 257-66.

Watanabe, E., Takeshima, N., Okada, A., & Inomata, K. (2001). Comparison of water- and land-based exercise in the reduction of state anxiety among older adults. *Perceptual & Motor Skills*, *91*(1), 97-104.

Wheeler, J. A., Gorey, K. M., Greenblatt, B. (1998). The beneficial effects of volunteering for older volunteers and the people they serve: A meta-analysis. *International Journal of Aging & Human Development*, *47*(1), 69-79.

White, H., et al. (2002). A randomized controlled trial of the psychosocial impact of providing internet training and access to older adults. *Aging & Mental Health*, *6*(3), 213-221.

White, H., McConnell, E., Clipp, E., Bynum, L., Teague, C., Navas, L., et al. (1999). Surfing the net in later life: A review of the literature and pilot study of computer use and quality of life. *Journal of Applied Gerontology*, *18*(3), 358-378.

Wikstrom, B. M., Theorell, T., & Sandstrom, S. (1993). Medical health and emotional effects of art stimulation in old age. A controlled intervention study concerning the effects of visual stimulation provided in the form of pictures. *Psychotherapy & Psychosomatics*, *60*(3-4), 195-206.

Wilcock, K. (2006a*). Live Choices: supporting older homeless people to engage in meaningful occupation in the community*. London: Help the Aged

Wilcock, K. (2006b). *Moving On: the meaning of activity for older homeless people.* London: Help the Aged

Williams, K., Gill, D., Butki, B., & Kim, B. (2000). A home-based intervention to improve balance, gait and self-confidence in older women. *Activities, Adaptation and Aging*, *24*(3), 57-70.

Williams, P. & Lord, S. R. (1997). Effects of group exercise on cognitive functioning and mood in older women. *Australian & New Zealand Journal of Public Health*, *21*(1), 45-52.

Wolinsky, F. D., Unverzagt, F. W. *et al.* (2006a). The effects of the ACTIVE cognitive training trial on clinically relevant declines in health-related quality of life. *Journals of Gerontology Series B – Psychological Sciences & Social Sciences, 61B*(5), S281-S287.

Wolinsky, F. D., Unverzagt, F. W., Smith, D. M., Jones, R., Stoddard, A., & Tennstedt, S. L. (2006b). ACTIVE cognitive training trial and health-related quality of life: Protection that lasts for 5 years. *Journals of Gerontology Series A – Biological Sciences and Medical Sciences, 61A*(12), 1324-1328.

Yuen, H. K. (2002). Impact of an altruistic activity on life satisfaction in institutionalized elders: A pilot study. *Physical & Occupational Therapy in Geriatrics*, *20*(3-4), 125-135.

Zauszniewski, J. A. (1997). Teaching resourcefulness skills to older adults. *Journal of Gerontological Nursing*, *23*(2), 14-20.

Appendix A Search strategy

For the effectiveness review the NHS Centre for Reviews & Dissemination (CRD) at the University of York ran the following strategy in Medline and translated it to run in the other databases included in this review.

#	Search stategy	Yield	Display
1	"Aged, 80 and over"/ or Aged/	1575647	DISPLAY
2	Frail Elderly/	3648	DISPLAY
3	((old$ or elderly) adj (people or person$ or woman or man or women or men)).ti,ab.	146823	DISPLAY
4	(elders or geriatric).ti,ab.	20195	DISPLAY
5	pensioner$.ti,ab.	553	DISPLAY
6	(aged people or aged person$).ti,ab.	1827	DISPLAY
7	seniors.ti,ab.	2240	DISPLAY
8	senior citizen$.ti,ab.	799	DISPLAY
9	(later-life or later life).ti,ab.	2907	DISPLAY
10	old age.ti,ab.	12140	DISPLAY
11	third age.ti,ab.	147	DISPLAY
12	(homeless persons/ or ethnic groups/ or poverty/ or prisoners/ or caregivers/) and (aged/ or "aged, 80 and over"/ or frail elderly/)	12957	DISPLAY
13	(homosexuality, female/ or homosexuality, male/ or bisexuality/ or transsexualism/ or transvestism/) and (aged/ or "aged, 80 and over"/ or frail elderly/)	616	DISPLAY
14	(lesbian$ or gay or gays or transgender$).ti. and (aged/ or "aged, 80 and over"/ or frail elderly/)	215	DISPLAY
15	or/1-14	1671021	DISPLAY
16	health promotion/ or preventive health services/	35634	DISPLAY
17	health education/	40836	DISPLAY
18	health behavior/	15154	DISPLAY

19	health knowledge, attitudes, practice/	32621	DISPLAY
20	exp communications media/	147296	DISPLAY
21	(television or video or radio or internet or book$ or booklet$ or leaflet$ or pamphlet$ or newspaper$ or magazine$).ti,ab.	82166	DISPLAY
22	(mental health adj3 (interven$ or promot$)).ti,ab.	897	DISPLAY
23	activities of daily living/ or leisure activities/ or exercise therapy/ or exp exercise movement techniques/	99466	DISPLAY
24	((daily or personal or group$ or arts or art or craft or crafts or exercise or social or physical or community or leisure or mental or living or learning or sport$ or preventive or preventative or health or healthy or everyday or famil$ or neighborhood or neighbourhood) adj3 (activities or activity)).ti,ab.	64725	DISPLAY
25	(interpersonal relation$ adj3 (interven$ or benefit$ or chang$ or develop$ or effect$ or enhanc$ or gain or improv$ or increas$ or maintain$ or maximis$ or maximiz$ or promot$ or rais$ or sustain$)).ti,ab.	232	DISPLAY
26	(social inclusion adj3 (interven$ or benefit$ or chang$ or develop$ or effect$ or enhanc$ or gain or improv$ or increas$ or maintain$ or maximis$ or maximiz$ or promot$ or rais$ or sustain$)).ti,ab.	20	DISPLAY
27	((daily or personal or group$ or art or arts or craft or crafts or exercise or social or physical or community or leisure or mental or living or learning or sport$ or preventive or preventative or health or healthy or everyday or famil$ or neighborhood or neighbourhood) adj3 (participat$ or involv$ or engage$)).ti,ab.	38617	DISPLAY
28	peer support.ti,ab. or peer group/	7985	DISPLAY
29	volunteering.ti,ab. or voluntary workers/	5462	DISPLAY
30	((information or advice or advocacy) adj3 (service$ or provid$ or provision$)).ti,ab.	68586	DISPLAY
31	((spiritual$ or religio$ or faith) adj3 (interven$ or benefit$ or chang$ or develop$ or effect$ or enhanc$ or gain or improv$ or increas$ or maintain$ or maximis$ or maximiz$ or promot$ or rais$ or sustain$)).ti,ab.	1106	DISPLAY
32	(life skills adj3 (interven$ or benefit$ or chang$ or develop$ or effect$ or enhanc$ or gain or improv$ or increas$ or maintain$ or maximis$ or maximiz$ or promot$ or rais$ or sustain$)).ti,ab.	54	DISPLAY
33	((befriending or intimacy or friendship) adj3 (interven$ or benefit$ or chang$ or develop$ or effect$ or enhanc$ or gain or improv$ or increas$ or maintain$ or maximis$ or maximiz$ or promot$ or rais$ or sustain$)).ti,ab.	205	DISPLAY

34	(friends/ or sexuality/ or sexual behavior/) and (interven$ or benefit$ or chang$ or develop$ or effect$ or enhanc$ or gain or improv$ or increas$ or maintain$ or maximis$ or maximiz$ or promot$ or rais$ or sustain$).ti,ab.	12640	DISPLAY
35	social support/ or family relations/ or community networks/ or intergenerational relations/	35359	DISPLAY
36	counseling/ and (interven$ or benefit$ or chang$ or develop$ or effect$ or enhanc$ or gain or improv$ or increas$ or maintain$ or maximis$ or maximiz$ or promot$ or rais$ or sustain$).ti,ab.	7944	DISPLAY
37	((counselling or therapy or bereavement counsel$) adj3 (interven$ or benefit$ or chang$ or develop$ or effect$ or enhanc$ or gain or improv$ or increas$ or maintain$ or maximis$ or maximiz$ or promot$ or rais$ or sustain$)).ti,ab.	78266	DISPLAY
38	(home visit$ or telecare).ti,ab. or house calls/ or day centre$.ti,ab. or day center$.ti,ab. or day care/	8312	DISPLAY
39	hobbies/ or gardening/ or art therapy/ or aromatherapy/ or tai ji/ or tai chi.ti,ab. or reminiscence.ti,ab.	2557	DISPLAY
40	self care/ or self-help groups/ or self help.ti,ab. or independen$ living.ti,ab. or functional ability.ti,ab.	23497	DISPLAY
41	(companion animal$ or (pet adj2 own$) or pet dog$ or pet cat$).ti,ab. or animals,domestic/	11199	DISPLAY
42	self help devices/ or (assistive device$ or assistive tech$).ti,ab.	2891	DISPLAY
43	(transport or travel).ti,ab.	175960	DISPLAY
44	((hous$ or home) adj3 adapt$).ti,ab. or architectural accessibilty/ or housing for the elderly/ or telecare.ti,ab.	1492	DISPLAY
45	carer$.ti,ab. or caregiver/ or domiciliary care.ti,ab. or home care services/	33402	DISPLAY
46	diet/ or alcohol drinking/	105106	DISPLAY
47	or/16-46	989134	DISPLAY
48	"Quality of Life"/	57282	DISPLAY
49	happiness/	1250	DISPLAY
50	(healthy aging or positive aging).ti,ab.	513	DISPLAY
51	((benefit$ or chang$ or develop$ or effect$ or enhanc$ or gain or improv$ or increas$ or maintain$ or maximis$ or maximiz$ or promot$ or rais$ or sustain$) adj3 (wellbeing or well-being or well being)).ti,ab.	3305	DISPLAY

52	((benefit$ or chang$ or develop$ or effect$ or enhanc$ or gain or improv$ or increas$ or maintain$ or maximis$ or maximiz$ or promot$ or rais$ or sustain$) adj3 (self esteem or life satisfaction or purpose in life or mastery or resilience or autonomy)).ti,ab.	3058	DISPLAY
53	((benefit$ or chang$ or develop$ or effect$ or enhanc$ or gain or improv$ or increas$ or maintain$ or maximis$ or maximiz$ or promot$ or rais$ or sustain$) adj3 (personal growth or wellness or morale or positive outlook or positive mental)).ti,ab.	997	DISPLAY
54	(benefit$ or chang$ or develop$ or effect$ or enhanc$ or gain or improv$ or increas$ or maintain$ or maximis$ or maximiz$ or promot$ or rais$ or sustain$).ti,ab. and (self concept/ or affect/ or personal autonomy/ or personal satisfaction/ or life change events/ or socialization/)	32919	DISPLAY
55	((prevent$ or reduc$ or minimis$ or minimiz$ or restrict$ or limit$) adj3 (dependence or discrimination or loneliness or exclusion or anxiety or distress or ageism)).ti,ab.	6529	DISPLAY
56	(prevent$ or reduc$ or minimis$ or minimiz$ or restrict$ or limit$).ti,ab.	2237227	DISPLAY
57	social isolation/ or prejudice/ or social alienation/	23761	DISPLAY
58	56 and 57	2733	DISPLAY
59	48 or 49 or 50 or 51 or 52 or 53 or 54 or 55 or 58	102951	DISPLAY
60	15 and 47 and 59	10983	DISPLAY
61	*Alzheimer Disease/	31975	DISPLAY
62	*Parkinson Disease/	23521	DISPLAY
63	*Dementia/	18036	DISPLAY
64	*Bipolar Disorder/	14629	DISPLAY
65	*Psychotic Disorders/	15939	DISPLAY
66	*Obsessive-Compulsive Disorder/	5100	DISPLAY
67	*Mental Disorders/	65631	DISPLAY
68	*Palliative Care/	12196	DISPLAY
69	or/61-68	180896	DISPLAY

70	60 not 69	10019	DISPLAY
71	exp adolescent/ or exp child/ or exp infant/	2125382	DISPLAY
72	70 not 71	8689	DISPLAY
73	limit 72 to (english language and yr="1990 - 2007")	7057	DISPLAY

For the cost-effectiveness review CRD ran essentially the same search, but replaced the effectiveness filter with the following quality of life filter:

1 (utilit$ approach$ or health gain or hui or hui2 or hui3).ti,ab.
2 (health measurement$ scale$ or health measurement$ questionnaire$).ti,ab.
3 (standard gamble$ or categor$ scal$ or linear scal$ or linear analog$ or visual scal$ or magnitude estimat$).ti,ab.
4 (time trade off$ or rosser$ classif$ or rosser$ matrix or rosser$ distress$ or hrqol).ti,ab.
5 (index of wellbeing or quality of wellbeing or qwb).ti,ab.
6 (multiattribute$ health ind$ or multi attribute$ health ind$).ti,ab.
7 (health utilit$ index or health utilit$ indices).ti,ab.
8 (multiattribute$ theor$ or multi attribute$ theor$ or multiattribute$ analys$ or multi attribute$ analys$).ti,ab.
9 (health utilit$ scale$ or classification of illness state$).ti,ab.
10 health state$ utilit$.ti,ab.
11 well year$.ti,ab.
12 (multiattribute$ utilit$ or multi attribute$ utilit$).ti,ab.
13 health utilit$ scale$.ti,ab.
14 (euro qual or euro qol or eq-5d or eq5d or eq 5d or euroqual or euroqol).ti,ab.
15 (qualy or qaly or qualys or qalys or quality adjusted life year$).ti,ab.
16 willingness to pay.ti,ab.
17 (hye or hyes or health$ year$ equivalent$).ti,ab.
18 (person trade off$ or persn tradeoff$ or time tradeoff$ or time trade off$).ti,ab.
19 theory utilit$.ti,ab.
20 (sf36 or sf 36).ti,ab.
21 (short form 36 or shortform 36 or sf thirtysix or sf thirty six or shortform thirtysix or shortform thirty six or short form thirtysix or short form thirty six).ti,ab.
22 (sf 6d or sf6d or short form 6d or shortform 6d or sf six$ or shortform six$ or short form six$).ti,ab.
23 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22

Appendix B Papers excluded from effectiveness review

Aday, R. H., Kehoe, G. C. et al. (2006). Impact of senior center friendships on aging women who live alone. *Journal of Women & Aging, 18*(1), 57-73.	No intervention. Paper presents panel data.
Age Concern & The Mental Health Foundation. (2006). *Promoting mental health and well-being in later life*. London: Age Concern.	The report does not contain information on specific interventions. It makes suggestions for promoting mental well-being.
Alt-White, A. C. (1993). An interdisciplinary approach to improving the quality of life for nursing home residents. Nursingconnections, 6(4), 51-59.	Arrived too late.
Anderson, R. T., King, A. et al. (2005). Physical activity counseling in primary care and patient well-being: Do patients benefit? *Annals of Behavioral Medicine, 30*(2), 146-54.	Results presented for all ages
Antunes, H. K., Stella, S. G. et al. (2005). Depression, anxiety and quality of life scores in seniors after an endurance exercise program. *Revista Brasileira de Psiquiatria, 27*(4), 266-71.	In Portuguese.
Anonymous (2005). Regular health assessments increase nursing home admissions and improve quality of life in elderly people. Evidence-based Healthcare & Public Health, 9(2), 135-6.	Not traced.
Audette, J. F., Jin, Y. S. et al. (2006). Tai Chi versus brisk walking in elderly women. *Age & Ageing, 35*(4), 388-393.	No results presented for measures of mental well-being.
Baker, M. K., Kennedy, D. J. et al. (2007). Efficacy and feasibility of a novel tri-modal robust exercise prescription in a retirement community: A randomized, controlled trial. *Journal of the American Geriatrics Society, 55*(1), 1-10.	No measure of mental well-being.
Barrett, C.J. & Smerdely, P. (2002). A comparison of community-based resistance exercise and flexibility exercise for seniors. Australian Journal of Physiotherapy, 48(3), 215-9.	Not traced.
Blair, C. E. (1999). Effect of self-care ADLs on self-esteem of intact nursing home residents. *Issues in Mental Health Nursing, 20*(6), 559-70.	Intervention is clinical.
Bowen, D., Kestin, M., McTiernan, A., Carrell, D. & Green, P. (1995). Effects of dietary fat intervention on mental health in women. *Cancer Epidemiology, Biomarkers & Prevention, 4*(5), 555-9.	Age range of 45-69 without distinction between subgroups.
Bradley, N. & Poppen, W. (2003). Assistive technology, computers & internet decrease sense of isolation for homebound elderly and disabled persons. *Technology & Disability, 15*(1), 19-25.	Few participants over 60 without clinical disability.

Brown, V.M., Allen, A. C., Dwozan, M., Mercer, I., & Warren, K. (2004). Indoor gardening and older adults: effects on socialization, activities of daily living, and loneliness. *Journal of Gerontological Nursing*, 30(10), 34-42.	No measure of mental well-being.
Brown, C. (2006). Systematic review concluded that there is an evidence base for occupational therapy interventions with people over 60 living independently in the community. *Occupational Therapy Now*, 8(1), 13-4.	Not traced.
Bruyere, O., Wuidart, M. et al. (2005). Controlled whole body vibration to decrease fall risk and improve health-related quality of life of nursing home residents. *Archives of Physical Medicine and Rehabilitation, 86*(2), 303-7.	Targeted on falls reduction (already covered by NICE guidelines)
Butts, J. B. (2001). Outcomes of comfort touch in institutionalized elderly female residents. *Geriatric Nursing, 22(*4), 180-184.	Intervention is clinical.
Campbell, J., & Aday, R. H. (2001). Benefits of a nurse-managed wellness program: a senior center model. *Journal of Gerontological Nursing, 27*(3), 34-43.	Measures of association only.
Cant, B. & Taket, A. (2005). Promoting social support and social networks among Irish pensioners in South London, UK. *Diversity in Health and Social Care, 4,*	No measure of mental well-being.
Carlson, M., Fanchiang, S. P. et al. (1996). A meta-analysis of the effectiveness of occupational therapy for older persons. *American Journal of Occupational Therapy 50*(2): 89-98.	As included studies focus on range of outcomes, it is not possible to ascertain effects on mental well-being.
Carlson, M., Clark, F. et al. (1998). Practical contributions of occupational science to the art of successful ageing: how to sculpt a meaningful life in older adulthood. *Journal of Occupational Science: Australia, 5*(3), 107-18.	General overview without data or information about measures.
Caserta, M. S. & Lund, D. A. (1993). Intrapersonal resources and the effectiveness of self-help groups for bereaved older adults. *Gerontologist, 33*(5), 619-629.	No measure of mental well-being.
Cassidy, J. & Jackson, J. (1999). Occupational therapy works for wellness. *Assisted Living Today*, 6 (1), 30-31	Not obtained in time.
Cattan, M., White, M., et al. (2005). Preventing social isolation and loneliness among older people: a systematic review of health promotion interventions. *Ageing & Society, 25*(1), 41-67.	As included studies focus on range of outcomes, it is not possible to ascertain effects on mental well-being.
Chambers, M. & McCammon, C. (1998). Using technology to enhance psychological well-being. *J Psychiatric & Mental Health Nursing, 5*(5), 423-4.	Work in progress; no data presented.

Chen, K., Snyder, M. & Krichbaum, K. (2002). Tai Chi & well-being of Taiwanese community-dwelling elders. *Clinical Gerontologist*, *24*(3-4), 137-56.	No intervention.
Chi, I. & Leung, E. M. (1998). Health promotion for e elderly persons in Hong Kong. *Journal of Health & Social Policy, 10*(3), 37-51.	Uncontrolled after study cannot estimate effectiveness of intervention.
Chin, A.P.M.J., et al. (2007). Physical exercise or micronutrient supplementation for the wellbeing of the frail elderly? A randomised controlled trial. *British Journal of Sports Medicine, 36*(2), 126-31.	Incorrect reference for Paw et al. (2007), included in this review.
Clark, F. A., et al. (2003). Lifestyle redesign improves health and is cost-effective. *OT Practice, 8*(2), 9-13.	Practitioner summary of Clark et al. (1997 & 2001), both included in this review.
Clark, H., Dyer, S.et al. (1998). *That bit of help: the high value of low-level preventative services for older people.* Bristol: The Policy Press.	Out of print.
Cohon, D., Hines, L., Cooper, B. A., Packman, W., & Siggins, E. (2003). Preliminary study of an intervention with kin caregivers. *Journal of Intergenerational Relationships, 1(3).*	Wide age range without distinction between subgroups.
Conn, V. S. et al. (2003). Integrative review of physical activity intervention research with aging adults. *Journal of the American Geriatrics Society, 51*(8), 1159-68.	No measure of mental well-being.
Cress, M. E., Buchner, D. M. et al. (2005). Best practices for physical activity programs and behavior counseling in older populations. *Journal of Aging and Physical Activity, 13*(1), 61-74.	General overview without intervention or measure of mental well-being.
Cyarto, E., Moorhead, G. & Brown, W. (2004). Updating the evidence relating to physical activity intervention studies in older people. *Journal of Science and Medicine in Sport, 7*(1Supp), 30-8.	No measure of mental well-being.
Demiris, G. et al. (2004). Older adults' attitudes towards and perceptions of "smart home" technologies: A pilot study. *Medical Informatics & the Internet in Medicine, 29*(2), 87-94.	No intervention or measure of mental well-being.
Dixon-Woods, M. et al. (2004). Organising services for influenza vaccination for older people. *Journal of Health Services & Research Policy*, 9(2), 85-90.	No measure of mental well-being.
Eichelbaum, E. (2000). Use of art therapy to deal with low self-esteem among the aged: selected case studies. *Medical Sciences*, 56(4), M200-5.	Not traced.
Elavsky, S. et al. (2005). Physical activity-related well-being in older adults: social cognitive influences. *Psychology & Aging, 20*(2), 295-302.	No measure of mental well-being.
Elkan, R., Kendrick, D., Dewey, M., Hewitt, M., Robinson, J., Blair, M. et al. (2001). Effectiveness of home-based support for older people: systematic review & meta-analysis. *BMJ, 323*(7315), 719-25.	No meta-analysis of mental well-being.

Elley, C. R., et al. (2003). Effectiveness of counselling patients on physical activity in general practice: Cluster randomised controlled trial. *British Medical Journal, 326*(7393), 793.	Age range 40-79 years. No results by age group. Difficult to determine any effects for our target group.
Elliott, H. (2005). The worth of mirth: A pilot study on the physiological and psychological effect of humour on older adults. *Perspectives, 29*(1), 6-9.	This is not a public health intervention; rather it is a pilot experiment comparing the effects of three types of humorous films.
Evans C. E, Goodman, C. & Redern, S. (2003). Maintaining independence in the cognitively intact elderly care home population - a systematic review of intervention trials. *Reviews in Clinical Gerontology, 13*(2), 163-174.	No measures of mental well-being. It looks at independence but this is measured by ADL.
Friedrich, D. (2003). Personal and societal intervention strategies for successful ageing. *Public Policy, Ageing International*, 28(1), 3-36.	No intervention. This is a commentary on ageing research.
Gallagher, B. (2003). Tai Chi Chuan and Qigong: physical and mental practice for functional mobility. *Topics in Geriatric Rehabilitation, 19 (*3), 172-182.	Review paper with no search strategy reported. Only provides an overview. No intervention.
Gallagher, E. M. & Brunt, H. (1996). Head over heels: Impact of a health promotion program to reduce falls in the elderly. *Canadian Journal on Aging-Revue Canadienne Du Vieillissement, 15*(1), 84-96.	Focuses on falls reduction - already a NICE guideline on this.
Galper, D. I., Trivedi, M. H. et al. (2006). Inverse association between physical inactivity and mental health in men and women. *Medicine and Science in Sports and Exercise, 38*(1), 173-178.	All ages (range 20-88) but results not broken down by age group
Garcia-Martin, M. A., Gomez-Jacinto, L. & Martimportugues-Goyenechea, C. (2004). A structural model of the effects of organized leisure activities on the well-being of elder adults in Spain. *Activities, Adaptation & Aging, 28*(3), 19-34.	Initially included, but on closer inspection the paper does not report the results of an intervention. The results are associations with well-being.
Gattuso, S. (2001). Healthy and Wise: Promoting mental health in Australian rural elders. *Journal of Mental Health & Aging, 7*(4), 425-433.	No outcome measure of mental well being
Gitlin, L. N., Winter, L., et al. (2006). A randomized trial of a multicomponent home intervention to reduce functional difficulties in older adults. *Journal of the American Geriatrics Society, 54*(5), 809-816.	Intervention is clinical.
Goff, K. (1993). Creativity and life satisfaction of older adults. *Educational Gerontology, 19*(3), 241-250.	The paper focuses on the relationship between creativity and life satisfaction pre and post and not the effects of the intervention per se on life satisfaction.

Gray-Donald, K., Payette, H., Boutier, V., & Page, S. (1994). Evaluation of the dietary intake of homebound elderly and the feasibility of dietary supplementation.[see comment]. Journal of the American College of Nutrition, 13(3), 277-284	Not arrived in time
Greenfield, E. A. & Marks, N. F. (2004). Formal volunteering as a protective factor for older adults' psychological well-being. *Journals of Gerontology Series B-Psychological Sciences & Social Sciences, 59*(5), S258-64.	Panel data. No intervention.
Hardcastle, S. & Taylor, A. H. (2005). Finding an exercise identity in an older body: "It's redefining yourself and working out who you are. *Psychology of Sport and Exercise, 6*(2), 173-188.	Only 3 of 8 participants >65. No individual analysis.
Hay, J., et al., Cost-effectiveness of preventive occupational therapy for independent-living older adults. Journal of the American Geriatrics Society, 2002. 50(8): p. 1381-8.	Referred to health economics team.
Hebert, Rejean; Robichaud, Line; Roy, Pierre-Michel; Bravo, Gina; Voyer, LouisEfficacy of a nurse-led multidimensional preventive programme for older people at risk of functional decline. A randomized controlled trial, Age and Ageing Vol. 30 No. 2 (Mar 2001)47-153	Clinical population.
Hedelin, B., K. Frojdh, et al. (1997). "The Herrhagen Social Network Intervention Programme: impact on depression, anxiety and life satisfaction." Health Care in Later Life 2(4): 251-9.	Unable to obtain in time
Help The Aged and Leeds Metropolitan University. (2006). Supporting older people to overcome social isolation and loneliness.	No measure of mental well-being
Help The Aged. (2006). My Home Life: Quality of life in care homes.	Ordered but not arrived.
Hill, W., Weinert, C., et al. (2006). Influence of a computer intervention on the psychological status of chronically ill rural women: Preliminary results. *Nursing Research, 55*(1), 34-42.	Age range 35-65. No average age reported and no results by age group, although the paper reports that there are 15 women aged 60-65.
Hinkka, K., S. L. Karppi, et al. (2006). "A network-based geriatric rehabilitation programme: Study design and baseline characteristics of the patients." International Journal of Rehabilitation Research 29(2): 97-103.	Not arrived
Hirst, J. (1997). "A time to dance." Community Care 20: 8-9.	Not arrived

Hogan, M. (2005). Physical and cognitive activity and exercise for older adults: a review. *International Journal of Aging & Human Development, 60*(2), 95-126.	Review paper. No methods of reviewing, unable to determine robustness of review.
Jackson, J. L. (2001). The effects of Thera-Band resistance training on health-related quality of life and muscular strength in adults ages 60--80. Unpublished Ed.D. 108 p., (Oklahoma State University).	Unable to obtain
Jackson, J., Carlson, M., et al. (1998). Occupation in lifestyle redesign: the Well Elderly Study Occupational Therapy Program. *American Journal of Occupational Therapy, 52*(5), 326-36.	No outcome measures are presented. The paper describes the study, rather than presents it with data.
Jackson, J., Mandel, D. R., & Zemke, R. (2001). Promoting quality of life in elders: An occupation-based occupational therapy program. *WFOT Bulletin, 43*, 5-12.	The paper is an overview of a specific therapy programme. The primary data it refers to is already included in this review.
Jensen, G. L., et al. (2004). Weight loss intervention for obese older women: improvements in performance and function. *Obesity Research, 12*(11), 1814-20.	Obese participants = clinical population
Johnson, W. L. (1997). The effects of pet encounter therapy on mood states and social facilitation in nursing home residents, (California School of Professional Psychology - San Diego).	Unable to obtain
Jones, A. (1997). Volunteers combat social isolation in older people. *Nursing Times, 93*(33), 52-3.	The paper is a topical summary and does not present any data from which conclusions can be drawn. There is little detail of the methods used.
Kara, B., Pinar, L., et al. (2005). Correlations between aerobic capacity, pulmonary and cognitive functioning in the older women. *International Journal of Sports Medicine, 26*(3), 220-224.	No measure of mental well-being
Katula, J. A., Blissmer, B. J., & McAuley, E. (1999). Exercise intensity and self-efficacy effects on anxiety reduction in healthy, older adults. *Journal of Behavioral Medicine, 22*(3), 233-247.	Measures of self-efficacy and state anxiety, not PWB or SWB
Kelley, S. J., Yorker, B. C., Whitley, D. M., & Sipe, T. A. (2001). Multimodal intervention for grandparents raising grandchildren: Results of an exploratory study. *Child Welfare Journal, 80(*1), 27-50.	Wrong population <65 (Mean age suggests this)
Kemp, C. A. (2004). Qigong as a therapeutic intervention with older adults. *Journal of Holistic Nursing, 22*(4), 351-73.	Review paper. Reports review methods, but the studies reviewed are clinical populations.

Keysor, J. J., & Jette, A. M. (2001). Have we oversold the benefit of late-life exercise? *Journals of Gerontology Series A-Biological Sciences and Medical Sciences, 56*(7), M412-M423.	The paper focuses on emotional disability not mental well-being
Kline-Mangione, K. (1995). Frailty research: A review of the FICSIT trials. *Topics in Geriatric Rehabilitation*, *11*(2), 61-70.	The studies reviewed primarily focus on falls. Most have no measure of mental well-being.
Krondl, M., Coleman, P. H., Bradley, C. L., Lau, D., & Ryan, N. (1999). Subjectively healthy elderly consuming a liquid nutrition supplement maintained body mass index and improved some nutritional parameters and perceived well-being. *Journal of the American Dietetic Association, 99*(12), 1542-1548.	Clinical intervention
Li, F., Harmer, P., McAuley, E., Fisher, K. J., Duncan, T. E., & Duncan, S. C. (2001). Tai Chi, self-efficacy, and physical function in the elderly. *Prevention Science*, *2*(4), 229-239.	No outcome measure of mental well-being.
Li, F., McAuley, E., Harmer, P., Duncan, T. E., & Chaumeton, N. R. (2001). Tai Chi enhances self-efficacy and exercise behavior in older adults. *Journal of Aging & Physical Activity, 9*(2), 161-171.	The measure of self efficacy is related to physical performance.
Li, I-chuan (2004). Effectiveness of a health promotion program for the low-income elderly in Taipei, Taiwan. *Journal of Community Health, 29*(6), 511-525.	No outcome measure of mental well-being.
Liddle, J., March, L. et al. (1996). Can occupational therapy intervention play a part in maintaining independence and quality of life in older people? A randomised controlled trial. *Australian & New Zealand Journal of Public Health, 20*(6), 574-8.	Intervention is clinical.
Likourezos, A., Burack, O. R. & Lantz, M. S. (2002). The therapeutic use of companion animals. *Clinical Geriatrics, 10*(4), 31-3.	Paper is too brief - the results are not described in terms of validated measures (there is no measure of mental well-being) the findings are described from a secondary perspective.
Marinelli, R., & Plummer, O. (1999). Healthy aging: beyond exercise. *Activities, Adaptation and Aging, 23*(4), 1-11.	Not enough information relevant to exclusion criteria. No measure of PWB.
McAuley, E., et al. (2003). Exercise self-efficacy in older adults: social, affective, and behavioral influences. *Annals of Behavioral Medicine, 25*(1), 1-7.	No outcome measure of mental well-being. The efficacy that is measured in the study is specific to exercise.

McAuley, E., et al. (2002). Physical activity and physique anxiety in older adults: Fitness, and efficacy influences. *Aging & Mental Health, 6*(3), 222-230.	No outcome measure of mental well-being. The measure of self efficacy is related to physical performance.
McAuley, E., Blissmer, B., Katula, J., Duncan, T. E., & Mihalko, S. L. (2000). Physical activity, self-esteem, and self-efficacy relationships in older adults: a randomized controlled trial. Annals of Behaviourl Medicine, 22(2), 131-139	The paper does not examine the global measure of self esteem.
McAuley, E., Katula, J., Mihalko, S. L., Blissmer, B., Duncan, T. E., Pena, M., et al. (1999). Mode of physical activity and self-efficacy in older adults: A latent growth curve analysis. Journals of Gerontology Series B-Psychological Sciences & Social Sciences, 54(5), P283-P292	No intervention
McConatha, D., McConatha, J. T., & Dermigny, R. (1994). The use of interactive computer services to enhance the quality of life for long-term care residents. *Gerontologist, 34*(4), 553-556.	Sample consists of people with dementia. Not clear how to separate these from the rest.
McMurdo, M. & Burnett, L. (1992) Randomised controlled trial of exercise in the elderly. *Gerontology, 38*(5), 292-298.	Pre 1993
Morrow-Howell, N., et al. (2003). Effects of volunteering on the well-being of older adults. *Journals of Gerontology Series B-Psychological Sciences & Social Sciences, 58*(3), S137-45.	No intervention. Paper presents panel data.
Moroz, A., et al., Convalescence care for seniors of lower Manhattan: an interdisciplinary outreach, rehabilitation, and education model. International Journal of Rehabilitation Research, 2004. 27(1): p. 75-8.	Not arrived in time
Motl, R. W., Konopack, J. F. et al. (2005). Depressive symptoms among older adults: long-term reduction after a physical activity intervention. *Journal of Behavioral Medicine, 28*(4), 385-94.	No measure of mental well-being
Mummery, K., Schofield, G. & Caperchione, C. (2004). Physical activity dose-response effects on mental health status in older adults. *Australian & New Zealand Journal of Public Health, 28*(2), 188-92.	No intervention
Nasman, O., Ilmarinen, J. (1999). Metal-Age: A process for improving well-being and total productivity. *Experimental Aging Research, 25*(4),	No results presented. Overview only of a proposed idea for an intervention.

379-384.	
Oken, B. S., D. Zajdel, et al. (2006). "Randomized, controlled, six-month trial of yoga in healthy seniors: effects on cognition and quality of life." Alternative Therapies in Health & Medicine 12(1): 40-7.	Arrived too late
Okun, M. A., Olding, R. W., & Cohn, C. M. (1990). A meta-analysis of subjective well-being interventions among elders. *Psychological Bulletin, 108*(2), 257-266.	Pre 1993
Park, J.S., [The effects of an elderly health promotion program on health promotion lifestyles, health status and quality of life in the elderly]. Taehan Kanho Hakhoe chi, 2004. 34(7): p. 1194-204.	Unable to obtain
Parkinson, S. (2006). The impact of physical exercise on psychological well-being in older adults, Parkinson, Susan: Walden U., US.	Unable to obtain
Pollin, R., et al., Effects of participation in a community exercise program on the self efficacy of geriatric adults. Journal of Geriatric Physical Therapy, 2004. 27(3): p. 120	Arrived too late
Reuben, D. B., Frank, J. C., Hirsch, S. H., McGuigan, K. A., & Maly, R. C. (1999). Randomized clinical trial of outpatient comprehensive geriatric assessment coupled with an intervention to increase adherence to recommendations. *Journal of the American Geriatrics Society, 47(*3). Pages?	Intervention is an assessment for long term care.
Rowland, L., Promoting the well-being of elderly women through exercise 1997, Cranfield Univ. (GB). p. [np].	Insufficient citation given
Ross, M. C., Bohannon, A. S., Davis, D. C., & Gurchiek, L. (1999). The effects of a short-term exercise program on movement, pain, and mood in the elderly. Results of a pilot study. *Journal of Holistic Nursing, 17*(2), 139-147.	No measure of mental well-being
Schultz, C. L. (1993). Caring for family caregivers in Australia: A model of psychoeducational support. *Ageing and Society*, *15*(1), 1-25.	
Shapiro, A. & Taylor, M. (2002). Effects of a community-based early intervention program on the subjective well-being, institutionalization, and mortality of low-income elders. *Gerontologist, 42*(3), 334-341.	Assessment for long term care

Shaw, D. C. Systematic review of the literature on respite care for the frail elderly.	Not due for publication until May 2008. clinical population
Staats, S. (1991). Quality of life and affect in older persons: Hope, time frames, and training effects. *Current Psychology: Research and Reviews,10*(1-2), 1-30.	Wrong date (too early)
Steptoe, A., et al. (2004). Quality of life and self-rated health in relation to changes in fruit and vegetable intake and in plasma vitamins C and E in a randomised trial of behavioural and nutritional education counselling. *British Journal of Nutrition*, *92*(1), 177-84.	All ages. Mean age = 43. No sub groups so not possible to determine any effects relevant to our age group.
Stevens, W., M. Hillsdon, et al. (1998). "Cost-effectiveness of a primary care based physical activity intervention in 45-74 year old men and women - a randomised controlled trialJournal title British Journal of Sports Medicine 32#March 1998: 236-241.	Not arrived
Steultjens, E. M. J., Dekker, J., Bouter, L. M., Jellema, S., Bakker, E. B., van den Ende, C. H. M. (2004). Occupational therapy for community dwelling elderly people: A systematic review. *Age and Ageing, 33*(5), 453-460.	The review did not focus on mental well-being outcomes. Two of the 26 included studies examined either the SF-36 or the life satisfaction index. No reference is made to these outcomes. The main outcome is functional ability.
Stevens, N. (1997). Friendship and well being: A course for older women. *Tijdschrift voor Gerontologie en Geriatrie, 28*(1), 18-26.	Paper is in German.
Stevens, N. (2001). Combating loneliness: A friendship enrichment programme for older women. *Ageing and Society*, *21*(2), 183-202.	Uncontrolled after study cannot estimate effectiveness of intervention; no measure of mental well-being.
Strauss-Blasche, G., Ekmekcioglu, C., Klammer, N., & Marktl, W. (2000). The change of well-being associated with spa therapy. Forschende Komplementarmedizin und Klassische Naturheilkunde, 7(6), 269-274.	In German
Tamaki, M., Shirota, A., Tanaka, H., Hayashi, M., & Hori, T. (1999). Effects of daytime nap in the aged. Psychiatry & Clinical Neurosciences, 53(2), 273-5.	No measure of mental well-being
Tsutsumi, T., The effects of strength training on mood, self-efficacy, cardiovascular reactivity and quality of life in older adults. 1997, Tsutsumi, Toshihiko: Boston U, US.	Unable to obtain
Tuck, I., Alleyne, R. et al. (2006). Spirituality and stress management in healthy adults. *Journal of Holistic Nursing, 24*(4), 245-53.	All ages (20-77) with a mean age of 52 (SD = 11.3)

van Haastregt. J., Diederiks, J., van Rossum, E., de Witte. L., & Crebolder, H. (2000). Effects of preventive home visits to elderly people living in the community: systematic review. *British Medical Journal, 320*(7237), 754-758.	Interventions included in the review are assessments for long term care.
Vela, R. H. (2006). The effects of relaxation and meditation on assisted living residents, their neighbors, and the overall quality of life of the residential facility. University of Kansas, US.	Unable to obtain
Victor, C. & Howse, K. (2001) Promoting the health of older people: setting a research agenda.	Review article. No details of methodology provided.
Vuori, I. (1998). Does physical activity enhance health? *Patient Education and Counseling, 33,* S95-S103.	Review paper with no methods described; no measure of mental wellbeing.
Washburn, A. M., Luxenberg, J., Brod, M., Steinhauer, M., & Katsap, M. (2001). A club of friends: Enrolling nursing home residents in an adult day program. *Journal of the American Medical Directors Association, 2*(5), 225-229.	Clinical population (nearly 50%)
Watanabe, E., Takeshima, N., Okada, A., & Inomata, K. (2001). Effects of increasing expenditure of energy during exercise on psychological well-being in older adults. *Perceptual & Motor Skills, 92*(1), 288-298.	No measure of mental well-being
Wilcox, S., Dowda, M. et al. (2006). Results of 1st year of Active for Life: translation of 2 evidence-based physical activity programs for older adults into community settings. *American Journal of Public Health, 96*(7), 1201-9.	No measure of mental well-being.
Wolf, S. L., et al., (1996). Reducing frailty and falls in older persons: an investigation of Tai Chi and computerized balance training. *Journal of the American Geriatrics Society*, *44*(5), 489-497.	No measure of mental well-being reported.
Wood, C. (1993). Mood change and perceptions of vitality: a comparison of the effects of relaxation, visualisation and yoga. *Journal of the Royal Society of Medicine, 86*(5), 254-9.	Age range 21-76. No results by age sub-groups. Difficult to determine effects to our target population.
Zeltzer, B.B. et al. (2003). Arts therapies promote wellness in elders. *Behavioral Healthcare Tomorrow*, 12(2), 7-12.	Alzheimers sufferers

Appendix C Papers excluded from cost-effectiveness review

STUDY	COMMENTS
Cost effectiveness and cost papers (9)	
Baumgarten, M., Lebel, P. et al. (2002). Adult day care for the frail elderly: outcomes, satisfaction, and cost. Journal of Aging & Health, 14(2), 237-59.	Clinical population; no measure of mental well-being
Burney, J. & Haughton, B. (2002). EFNEP: a nutrition education program that demonstrates cost-benefit. Journal of the American Dietetic Association, 102(1), 39-45.	Wrong population (< 60 years)
Dalziel, K.., Segal, L. et al. (2006). Cost utility analysis of physical activity counselling in general practice. Australian & New Zealand Journal of Public Health, 30(1), 57-63.	Wrong population (40-79 years) without age breakdown
Elley, R. Kerse, N., Arroll, B., Swinburn, B., Ashton, T., Robinson, E. (2004). Cost-effectiveness of physical activity counselling in general practice. New Zealand Medical Journal, 117(1207), 1216-30.	Wrong population (40-79 years) without age breakdown
Huang, N., Pietsch, J. et al. (2004). Victorian Active Script Programme: promising signs for general practitioners, population health, & promotion of physical activity. British Journal of Sports Medicine, 38(1), 19-25.	No measure of mental well-being
Isaacs, A.J. et al. (2007). Exercise Evaluation Randomised Trial (EXERT): randomised trial comparing GP referral for leisure centre-based exercise, community-based walking and advice only. Health Technology Assessment, 11(10).	Wrong population (40-74 years) without age breakdown
Lansley, P., McCreadie, C. et al. (2004). Can adapting the homes of older people and providing assistive technology pay its way? Age and Ageing, 33(6), 571-6.	No measure of mental well-being
Mann, W. C., Ottenbacher, K. J. et al. (1999). Effectiveness of assistive technology and environmental interventions in maintaining independence and reducing home care costs for the frail elderly: a randomized controlled trial. Archives of Family Medicine, 8(3), 210-217.	No measure of mental well-being
Vestergaard, S., Andersen, C. K. et al. (2006). Exercise intervention of 65+-year-old men and women: functional ability and health care costs. Aging – Clinical and Experimental Research, 18(3), 227-34.	No measure of mental well-being
Utility papers (19)	
Antunes, H. K., Stella, S. G. et al. (2005). Depression, anxiety and quality of life scores in seniors after an endurance exercise program. Revista Brasileira de Psiquiatria, 27(4), 266-71.	See effectiveness review; also likely to be in Portuguese
Barnett, A., Smith, B. et al. (2003). Community-based group exercise improves balance and reduces falls in at-risk older people: a randomised controlled trial. Age & Ageing, 32(4), 407-14.	No measure of mental well-being
Barrett, C. J. & Smerdely, P. (2002). A comparison of community-based resistance exercise and flexibility exercise for seniors. Australian Journal of Physiotherapy, 48(3), 215-9.	See effectiveness review
Cress, M. E., Buchner, D. M. et al. (1999). Exercise: effects on physical functional performance in independent older adults. Journals of Gerontology A – Biological Sciences & Medical Sciences, 54(5), M242-8.	No measure of mental well-being
Damush, T. M. & Damush, J. G. (1999). The effects of strength training on strength and health-related quality of life in older adult women. Gerontologist, 39(6), 705-10.	See effectiveness review
Elley, C. R., Kerse, N. et al. (2003). Effectiveness of counselling patients on physical activity in general practice: cluster randomised controlled trial. BMJ, 326(7393), 793-6.	Wrong population (40-79 years) without age breakdown
Helbostad, J. L., Sletvold, O. et al. (2004). Home training with and without additional group training in physically frail old people living at home: effect on health-related quality of life and ambulation. Clinical Rehabilitation, 18(5): 498-508.	See effectiveness review
Jensen, G. L., Roy, M. A. et al. (2004). Weight loss intervention for obese older women: improvements in performance and function. Obesity Research, 12(11), 1814-20.	See effectiveness review

Kerse, N., Elley, C. R. et al. (2005). Is physical activity counseling effective for older people? A cluster randomized, controlled trial in primary care. Journal of the American Geriatrics Society, 53(11), 1951-6.	See effectiveness review
Kingston, P., Bernard, M. et al. (2001). Assessing the health impact of age-specific housing. Health & Social Care in the Community, 9(4), 228-34.	No intervention
Kolt, G. S., Oliver, M. et al. (2006). An overview and process evaluation of TeleWalk: a telephone-based counseling intervention to encourage walking in older adults. Health Promotion International, 21(3), 201-8.	No measure of mental health
Latham, N., Anderson, C. et al. (2003). Progressive resistance strength training for physical disability in older people. Cochrane Database of Systematic Reviews, 2, CD002759.	No measure of mental health well-being
Lee, C. & Russell, A. (2003). Effects of physical activity on emotional well-being among older Australian women: cross-sectional and longitudinal analyses. Journal of Psychosomatic Research, 54(2), 155-60.	No intervention
Matuska, K., Giles-Heinz, A.. et al. (2003). Outcomes of a pilot occupational therapy wellness program for older adults. American Journal of Occupational Therapy, 57(2), 220-4.	See effectiveness review
Myers, A. M., Malott, O. W. et al. (1999). Measuring accumulated health-related benefits of exercise participation for older adults: the Vitality Plus Scale. Journals of Gerontology Series A – Biological Sciences & Medical Sciences, 54(9), M456-66.	Wrong population no measure of mental well-being
Oken, B. S., Zajdel, D. et al. (2006). Randomized, controlled, six-month trial of yoga in healthy seniors: effects on cognition and quality of life. Alternative Therapies in Health & Medicine, 12(1), 40-7.	See effectiveness review
Patrick, D. L., Grembowski, D. et al. (1999). Cost and outcomes of Medicare reimbursement for HMO preventive services. Health Care Financing Review, 20(4), 25-43.	No measure of mental well-being
Stiggelbout, M., Popkema, D. Y. et al. (2004). Once a week is not enough: effects of a widely implemented group based exercise programme for older adults; a randomised controlled trial. Journal of Epidemiology & Community Health, 58(2), 83-8.	See effectiveness review
Whitehurst, M. A., Johnson, B. L. et al. (2005). The benefits of a functional exercise circuit for older adults. Journal of Strength & Conditioning Research, 19(3), 647-51.	No measure of mental well-being