The effectiveness and cost effectiveness of methods of protecting and promoting the health of older workers

Evidence Review for Research Question 1

Annette Cox, Jim Hillage, Luke Fletcher, Rosa Marvell, Sally Wilson, Linda Miller, Sam Swift
Institute for Employment Studies

Tyna Taskila, Zofia Bajorek, Anthony Hind
The Work Foundation

Jenny Brine
Lancaster University

THE WORK FOUNDATION
PART OF LANCASTER UNIVERSITY

ies
institute for employment studies

Institute for Employment Studies

IES is an independent, apolitical, international centre of research and consultancy in HR issues. It works closely with employers in all sectors, government departments, agencies, professional bodies and associations. IES is a focus of knowledge and practical experience in employment and training policy, the operation of labour markets, and HR planning and development. IES is a not-for-profit organisation.

The Work Foundation

Since its inception as The Boy's Welfare Society in 1918 to its present day alliance with Lancaster University, The Work Foundation has conducted work which supports organisations to improve the quality of working life. Since 2002, The Work Foundation has concentrated on producing applied research on workplace health and well-being which enables policy-makers, employers and clinicians to appreciate the links between workforce health, productivity and social inclusion, including the way people are managed at work, the way their jobs are designed, the culture and climate of the organisation and the efforts which employers put into physical and psychological well-being can make a crucial difference to both productivity and well-being.

Institute for Employment Studies
City Gate
185 Dyke Road
Brighton BN3 1TL
UK

Telephone: +44 (0)1273 763400
Email: askies@employment-studies.co.uk
Website: www.employment-studies.co.uk

IES project code: 3469

Contents

Executive Summary

The National Institute for Health and Care Excellence (NICE) has been asked by the Department of Health to develop public health guidance for employers and employees on effective and cost effective ways of promoting and protecting the health of older workers, covering workplace adaptations and adjustments to their changing needs in order to extend working lives and prepare for retirement.

The Institute for Employment Studies (IES) in partnership with The Work Foundation (TWF), Lancaster University, York Health Economics Consortium (YHEC), and University of Loughborough (UoL) have been contracted to undertake the evidence reviews of relevant effectiveness and qualitative studies and the economic analysis.
Three research questions were developed and following the search process, evidence has been found to address Research Questions 1 and 3:

- 'What are the most effective and cost-effective methods of protecting and promoting the health and wellbeing of older workers, and of supporting workers who wish to continue in employment up to and beyond state pensionable age? What supports, or prevents, implementation of these methods?' (RQ 1)
- 'What factors facilitate or constrain workplaces to enhance the wellbeing of older workers, to support them in continuing to work up to and beyond state pensionable age and affect the quality and outcomes of pre-retirement planning?' (RQ3)
- No evidence was found to address Research Question 2
- What are the most effective and cost-effective ways of helping older workers plan and prepare for retirement? What supports, or prevents, implementation of these methods?

This report presents the first of two reviews based on effectiveness studies which examined workplace policies and practices to protect and promote the health and wellbeing of older workers, and supporting workers who wish to continue in employment up to and beyond state pensionable age. A subsequent qualitative review will cover workplace policies and practices on pre-retirement planning and

will examine the factors affecting the health and well-being of older workers, both in work and in subsequent retirement.

It was agreed with NICE project team at the outset that a joint search strategy would be adopted for all three research questions which would cover:

- A search of key literature databases
- A search of the websites of relevant organisations
- Citation searches of material included in the reviews
- A review of material submitted through the NICE Call for Evidence
- Writing to any known researchers and experts in the field not already contacted during the Call for Evidence to ask for relevant material.

All the papers were reviewed against inclusion and exclusion criteria agreed with the NICE project team. Included studies were those that had an experimental or observational design, were published in English since 2005, set in an OECD country including European countries which acceded to membership of the EU on or before 2004, which examined a workplace intervention, policy or practice aimed at protecting and promoting the health and wellbeing of workers aged at least 50. Interventions or support that employees access on their own, statutory provision or interventions to promote physical activity, mental wellbeing and smoking cessation in the workplace, and to manage sickness absence were excluded. Managing long-term sickness absence, promotion of physical activity and smoking cessation are already covered by existing NICE guidance.

The 27,738 titles and abstracts identified through the initial search process were screened through a two-stage process to identify papers that should be considered for full paper screening, using a checklist based on the inclusion/exclusion criteria. Articles were identified at this stage as being relevant for Review Question 1, 2 or 3.

The full papers of all the studies that came through the initial screening process were ordered. Retrieved papers were appraised by two members of the review team using the full inclusion/exclusion checklist to assess the content of the articles and whether they should be included in the review (see Appendix 3).

The 34 papers identified for full paper screening for Review Question 1 have been screened and extracted. During the screening process seven papers were identified for inclusion in this review and an additional three for Review Question 3.

The seven papers identified for inclusion in this review were assessed for quality and the data extracted and presented in an evidence table by two separate members

of the review team. Papers were assessed using a checklist based on the quality assessment in the NICE Public Health Guidance Methods Manual (NICE, 2012). Depending on how they met the criteria behind the checklist papers were graded either: '++', '+' or '-'.

Findings

One study (Harma et al. 2005) found that moving from a backwards rotating shift system to a rapidly forwards rotating shift system had positive significant associations with beneficial results for psychomotor test outcomes, objective sleep measures and self-reported sleepiness and quality of life indicators among workers aged at least 45.

Evidence Statement 1: shift patterns

There is weak evidence from one (-) study[1] 'before and after' non-random controlled and longitudinal study set in Finland on male aircraft maintenance workers aged at least 45 that changing from a backwards to a rapidly forwards rotating shift system can result in significant positive changes in self-reported sleepiness after the morning shifts (Age: df 1,413, F =6.1, p <0.01) , sleep quality, quality of life indicators including sleep and vigilance (group * time * age: df 4, 404, F = 9.5,p <0.0001), general well-being at work (group *time *age: df 4, 413, F =10.0, p <0.001), social life (group *time *age: df 4, 416, F =6.4, p <0.0001), family life (group * time *age: df 4, 408, F =5.0, p <0.0006), hobbies (group *time *age: df 4, 416, F =3.2, p <0.01), and psychomotor performance with a significant decrease of the median reaction times at the end of the night shift among the older workers (mean ±s.e. from 376 ±18 to 353±15 ms).

This evidence appears to be mostly applicable to the UK because of likely standardisation in work content and processes due to international regulation of aircraft maintenance, but concerns about its quality in terms of the small sample sizes and participant selection need to be taken into account.

[1] Harma et al. 2005 (-)

One study (Rutanen et al. 2014) found that regular physical exercise for symptomatic menopausal women aged between 44 and 62 can result in significant positive changes in self-reported mental resources and decreased daily physical work strain. It is possible that improved mental resources could be attributed to contact between intervention participants and the research team in fortnightly meetings during the intervention because these may have an impact through feedback, motivation and perceived emotional support.

Evidence Statement 2: physical activity

There is weak evidence from one (+) study[2] using a randomised controlled trial on working symptomatic menopausal women set in Finland that regular physical exercise for this group can result in significant positive changes in self-reported mental resources (coefficient 0.58, 95% CI = 0.17- 0.00, p < 0.01) and decreased daily physical work strain (coefficient -0.26, 95% CI = -0.45 - -0.07, p < 0.01).

This evidence appears to be fully applicable to the UK because there do not appear to be institutional differences which would mitigate the implementation of the intervention, although lack of blinding in allocation of participants to control and intervention groups may have resulted in changed behaviour among control group members, and early assessment of an outcome measure intended to be used 12 months rather than 6 months after the intervention may mean the study did not accurately assess the full potential impact of the intervention.

[2] Rutanen et al. (2014) (+)

One study (Wagner et al. 2007) found that developing and applying group-based problem-solving techniques can result in significant positive improvement in objective and subjectively assessed memory function and work-related attitudes among inpatients aged 50-59 in a clinic providing psychological therapies. The authors were unable to determine the mechanisms underlying the increased memory performance in the intervention group and a long-term outcome was not assessed so it is not possible to comment on actual transfer of techniques learned to daily work and life performance.

Evidence Statement 3: psychological support

There is weak evidence from one (+) study[3] using a controlled trial among inpatients aged 50-59 in a psychological treatment facility in Germany that developing and applying group-based problem-solving techniques can result in significant positive improvement following the intervention on ability to schedule appointments ($F=15.06$, $p<0.001$), a memory function test ($F=4.95$, $p<0.05$), reduced anxiety about everyday memory function ($t=-2.83$, $p<0.01$) and decreased pursuit of perfection ($t=3.23$, $p<0.1$) and reduced level of exhaustion ($t=-4.17$, $p<.001$) in a questionnaire of work-related attitudes.

The study has unknown applicability to UK workplaces because the nature of work undertaken by the participants is unknown so may not be directly comparable to jobs in the UK labour market, it is not known whether the work attitudes test and clinic-specific memory questionnaire would achieve validity in the English language, and there may be potential differences in the administration of psychological inpatient treatment.

[3] Wagner et al. 2007 (+)

One longitudinal study (Wegner et al. 2008) aimed to investigate whether an inpatient psychotherapeutic treatment with a job-specific element showed long-acting success on burn-out of schoolteachers aged between 27 and 64 (mean age 51)

in a rural area of Germany and whether there was any variation in impact by gender and type of school. The authors suggest that intervention could have greater effects if men especially were more willing to seek treatment and that the long-term effects of the intervention, particularly on delaying retirement age, could be greater if the intervention was begun earlier.

Evidence Statement 4: psychotherapy support

There is weak evidence from one (-) longitudinal study that an inpatient psychotherapeutic treatment with a job-specific element had a long-term impact on burn-out of schoolteachers following inpatient treatment in a rural area of Germany. Improvements in teacher health indicators post-treatment were found in increased incidence of teachers without sickness absence in the previous quarter from 29.5% to 51.8% ($p < .001$), disappearance of a statistically significant difference between burnout scores among high school teachers compared to those in other schools after treatment from 37.7 high school teachers and 26.5 other school teachers ($p < .05$) to 26.5 for high school teachers and 24.8 for other teachers ($p = .599$), disappearance of a statistically significant differences in higher depersonalisation ($p < .0001$) and lower personal accomplishment scores ($p < .05$) in men compared with women after treatment.

There was also improvement in both sexes in scores of high emotional exhaustion, high depersonalisation, and low personal accomplishment and the percentage of participants who had retired or were no longer teaching was positively related to older age in the follow-up survey.

The authors speculated in the conclusions that older workers would benefit from earlier intervention to prolong working lives.

This study has limited applicability to the UK. The authors note that teachers in Germany have special entitlement to this kind of inpatient intervention for burnout through the terms of civil service employment contracts. In contrast teachers in the state education system in the UK have no specific occupational healthcare entitlements.

[4] Wegner et al. 2008 (-)

One pooled cross-sectional study using employer survey and matched employee administrative data at two time points and difference-in-difference comparisons found evidence of a drop in the likelihood of sickness absence from 2001 to 2007 for employees using preventive measures and evidence of positive impact of such measures specifically for employees in the public sector. Detail on the nature of the interventions was not clear in the published paper so the review team contacted the lead author by email for clarification. The lead author responded by stating that the intervention could be any one of 12 possible measures including reduced working hours (with or without reduced pay), temporary or permanent change of occupation and free physical therapy, massage or exercise within working hours. The three most common measures implemented were work adaptation, changed work tasks and technical equipment which were implemented – alone or in combination – in workplaces covering 70% of the employees who had access to at least one

intervention. One in five workers in establishments implementing the interventions were covered by four or more measures.

The odds for sickness absence levels were about 20% higher for employees in establishments with at least one preventive measure compared to establishments without preventive measures and suggest this may reflect the introduction of measures in response to perceived problems with sickness absence. Levels of sick leave were explained in most other sectors by adjusting for industry with no impact from the presence of preventive measures.

Evidence Statement 5: workplace measures

There is weak statistically significant evidence from one cross-sectional pooled study[5] set in workplaces in Norway of a 10% drop in the odds for sickness absence in the period from 2001 to 2007 among employees aged over 50 in establishments using preventive measures (unfortunately not possible to identify in the paper) (measures OR 1.20 CI 95% = 1.12-1.28, change 2001 – 2007 OR = 0.97, CI 95% = 0.9101.03, measure x change OR = 0.89, CI 95% 0.81-0.97). There is positive and statistically significant evidence that adoption of at least one measure has contributed to reducing sickness absence among employees aged 50 years or older in public sector workplaces (measures OR =1.70, CI 95% = 1.37–2.11; change 2001 – 2007 OR = 1.27 CI 95% = 1.06–1.52; measure x change OR = 0.60, CI 95% = 0.45–0.79).

This evidence is weakly applicable to the UK because while Norway has a similar economy there is limited information on the nature of the interventions and how they were designed and applied in an industrial relations context which is different from the UK.

[5] Midtsundstad and Nielsen (2014) (+)

Lastly two studies of health promotion activities showed positive links with changes in health behaviours among older workers.

Evidence Statement 6: health promotion

There is moderate evidence from two (+) RCT studies[1,2] that health promotion programmes aimed at older workers can have positive effects on participants' diet and level of exercise.

One RCT[1] set among employees aged 45 and over in two academic hospitals in the Netherlands that a worksite vitality intervention (comprising exercise and yoga sessions, free fruit and visits from a coach) significantly increased participants' weekly sports activities (ß = 40.4 minutes per week, $p<0.05$) and fruit intake (ß = 2.7 pieces per week, $p<0.05$), when compared to the control group. The intervention also favourably affected the need for recovery after a day of work (ß= - 3.5 points on a 100-point scale derived from the Experience and Evaluation of Work survey, $p<0.05$). No effects were observed for vigorous, aerobic capacity and mental health.

A second RCT[2] set among workers aged 40 and over in a university in Chicago, USA found that computerised health risk assessments combined with individualised, negotiated health improvement action plans and ongoing support and reinforcement from a coach

had a positive effect on participants' diet (z = 3.55, p = <0.001) and physical activity (z = 2.22, p= 0.13) compared with a control group (who received printed health promotion materials). No effects were found for on measures of stress, smoking and weight. No positive effects compared with a control group were found for a second parallel intervention in which participants undertook an automated health risk assessment accompanied by self-directed use of on-line health modules and receipt of generic health tips by email.

[1] Strijk et al. (2012) (+)

[2] Hughes et al. (2011) (+)

The interventions are not particularly intensive in delivery and could be provided to the entire workforce, consistent with a life course perspective of preventing worker ill health at any age. However, the outcome measures used in the studies are intermediate and it would be helpful to understand impact on health outcomes.

While the studies overall generally indicate that interventions can have a positive association with the wellbeing of older workers, they tend to focus on very specific interventions or types of older workers, only three focus on interventions made directly by employers in the workplace and do not contain data on the relative costs and benefits of each type of initiative. Therefore it is difficult to draw any general conclusions of interventions that should be recommended to employers on the basis of this evidence alone.

None of the studies are set in the UK and in some cases their applicability to a UK setting is limited. The relevance of the evidence base is also limited by the quality of the interventions, of which two are rated (-) and five are rated (+).

Despite increasing policy interest in how the health and wellbeing of older workers can be supported, especially in the context of a decline in the proportions of people in some younger segments of the population and planned increases to the age for state pension eligibility, on the face of the evidence so far, there are very few intervention studies on this topic. This review urges research commissioners to prioritise funding of high quality studies into the impact of workplace level interventions on older workers' health and well-being outcomes which will seek to track the health and well-being of individuals during the lifespan of the intervention and onwards to the end of their working lives.

1 Introduction

The National Institute for Health and Care Excellence (NICE) has been asked by the Department of Health to develop public health guidance for employers and employees on effective and cost effective ways of promoting and protecting the health of older workers, covering workplace adaptations and adjustments to their changing needs in order to extend working lives and prepare for retirement. As part of the process of developing the guidance, NICE has commissioned a series of evidence reviews and an economic evaluation.

The Institute for Employment Studies (IES) in partnership with The Work Foundation (TWF), Lancaster University, York Health Economics Consortium (YHEC), and University of Loughborough (UoL) have been contracted to undertake the evidence reviews of relevant effectiveness and qualitative studies and the economic analysis.

This report presents the first of these reviews based on effectiveness studies which examined workplace policies and practices to protect and promote the health and wellbeing of older workers, and of support workers who wish to continue in employment up to and beyond state pensionable age. Subsequent reviews will cover the effectiveness of workplace policies and practices on pre-retirement planning and a qualitative review of studies which examine the factors affecting the health and well-being of older workers, both in work and in subsequent retirement.

1.1 Background

The health of the working population is vital to the economy and to society, but due to changing demographics of the workforce, western societies are facing great challenges to maintain economic growth and competiveness. The workforce is ageing in the UK. It has been estimated that approximately one third of the labour force will be aged 50 or over by 2020 (Taylor 2007). Ignoring the skills, knowledge and contribution that older workers are capable of making to organisational performance has been described as a high-risk strategy (Foresight Mental Capital and Wellbeing Project 2008). The number of working age adults across Europe has begun to decline and some sectors of the European economy are beginning to report significant skills shortages. Furthermore both employers and governments face increasing difficulties meeting the financial costs of their pension commitments. In

response, many European governments have increased state pension ages or reduced the generosity of state pensions to address this issue (Sinclair et al. 2013). Partly as a result, the workforce is older and more likely to face health problems with more people living with a long standing health problem or disability. According to The Labour Force Survey (2011), of 7.2 million aged 50-64 who are employed, 42% are living with a health condition or disability in the UK (Sinclair et al. 2013). It is likely that chronic disease rates will continue to rise; much of this is due to an increase in poor life style factors, such as poor diet, smoking and lack of exercise. Older people in disadvantaged groups more commonly face health problems at an earlier age, and are more likely to face difficulties in finding and keeping jobs, partly due to lower educational attainment and lower skill levels (Bloomer , 2014).

Ill-health represents a major economic burden for society due to increased healthcare costs, loss in productivity and sickness absence. Both males and females over the age of 55 take more days off work due to self-reported ill health caused or made worse by work. The most common sources of new cases of work-related illness reported were musculoskeletal complaints and stress, depression or anxiety, with those over 45 having the highest estimated prevalence rate (Crawford et al. 2009). Mental ill-health is associated with both physical and mental decline which is more common among older groups (Foresight Mental Capital and Wellbeing Project 2008). Besides poor health, the reasons for ceasing economic activity at age 50+ include limited skills and increased caring responsibilities (Marmot 2010). An evidence based review on the health, safety and health promotion needs of older workers (Crawford et al. 2009) identified that although there is an increased risk with age of developing a disease, this is not necessarily a reason to exclude an individual from work. Certain diseases, such as heart disease or diabetes, can be controlled and reasonable adjustments can be made to keep the individual at work.

The health of employees is a major factor in an organisation's competitiveness. Although absence rates have been falling in recent years, it has been estimated that annual costs of sickness absence for UK businesses is nearly £14 billion a year (Vaughan-Jones & Barham 2009). Employees in good health can be up to three times as productive as those in poor health; they can experience fewer motivational problems; they are more resilient to change; and they are more likely to be engaged with the business's priorities (Vaughan-Jones & Barham 2010). In Dame Carol Black's review of the health of Britain's working age population it was calculated that improved workplace health could generate cost savings to the government of over £60 billion – the equivalent of nearly two thirds of the NHS budget for England (Black 2008).

It has been recognised that improved workplace health has the potential to make a significant contribution to the economy, to public finances and to reducing levels of

disease and illness in society (Waddell and Burton 2006). Employers play a key role in helping to protect health and prevent future ill health of the working population and NICE Public Health Guidelines (2009) recommend a strategic and coordinated approach to promoting employees' mental health wellbeing. One of the biggest challenges facing the working longer agenda is poor health of older workers. However, until recently, relatively few initiatives by governments or employers have been established to explicitly improve the health of older workers (Sinclair et al. 2013). In fact, according to research from the Chartered Institute of Personnel and Development (CIPD) and the Chartered Management Institute (CMI) into age management, UK employers are still 'woefully unprepared for the impact workforce demographics will have on their businesses' (Macleod et al. 2010).

Despite these barriers, the number of employed people aged 65 or over has more than doubled over the past two decades, from 425,000 in March to May 1994 to 1.1 million March to May 2014 (ONS 2014).

Survey research of 1,500 older workers by the Equality and Human Rights Commission (Smeaton et al. 2009) found that 60% of older workers wanted to carry on working after retirement age either in the same or different jobs. This is often because they cannot afford to retire. Whilst economic considerations are a key factor, personal fulfilment is also important to older workers, with re-entering the workforce for enjoyment or company at work (Parry & Harris 2011). The decision of whether or not to continue working is complex and influenced not only by a combination of individual factors but also by organisational culture and policies.

Although there has been increasing research interest in the well-being of older workers (eg Crawford et al. 2009) and 'pre-retirement' training (Foresight Mental Capital and Wellbeing Project 2008), systematic evaluation of the best approach to the management of age diversity at the workplace is lacking. As more employers recognise the need to promote the health and wellbeing of ageing employees, it is important that they have access to guidelines which help them to provide healthy and good quality working environments in a cost effective way and using evidence-based interventions. Therefore NICE have commissioned systematic evaluation of the evidence on the effective policies and approaches for promoting and protecting the health of older workers to underpin the development of guidance for employers and others.

1.2 Aims and objectives of the review

The overall aim of this review is to identify, appraise and summarise research evidence to support the development of guidance for employers and employees on effective management practices to improve the health of older workers (aged 50 or over). The guidance will be aimed at human resources professionals, trade unions

and professional bodies. It will also be aimed at health professionals (particularly those working in occupational health), and commissioners and managers with public health as part of their remit. It will be of interest to people who are self-employed and other members of the public. The guidance will cover organisational policies and initiatives for older employees, changes to the way work is organised and the work environment, activities to challenge or counteract ageism, retirement planning and training for mentors and older workers and any initiatives by organisations representing employers or the wider business community to promote the above.

The specific aim of this first review is to examine the following research question (RQ1):

> *What are the most effective and cost-effective methods of protecting and promoting the health and wellbeing of older workers, and of supporting workers who wish to continue in employment up to and beyond state pensionable age?*

In addition the following secondary question will also be considered

> *What supports, or prevents, the implementation of these methods?*

1.3 Structure of the report

This report covers:

- The methodology we adopted to conduct this review
- The findings from the review
- A discussion of the evidence.

In addition a series of Appendices provide further information on our approach and a bibliography of the studies included and excluded from this review.

2 Methodology

2.1 The review team

The review was conducted by the Institute for Employment Studies (IES) in partnership with The Work Foundation (TWF), the York Health Economics Consortium, and the University of Loughborough. The review team was led by Dr Annette Cox, Associate Director at IES, and included Jim Hillage from IES, Dr Tyna Taskila from The Work Foundation, Dr Matthew Taylor from York Health Economics Research Consortium and Professor Cheryl Haslam from the University of Loughborough.

2.2 Overall search strategy

It was agreed with NICE project team at the outset that a joint search strategy would be adopted for all three research questions which would cover:

- Effectiveness studies (for Review Questions 1 and 2)
- Qualitative studies (for Review Question 3)
- Economic studies (for the Economics review)

The search for relevant evidence covered a number of elements:

- A search of key literature databases
- A search of the websites of relevant organisations
- Citation searches of material included in the reviews
- A review of material submitted through the NICE Call for Evidence
- Writing to any known researchers and experts in the field not already contacted during the Call for Evidence to ask for relevant material.

2.3 Inclusion and exclusion criteria

All the papers were reviewed against inclusion and exclusion criteria agreed with the NICE project team in relation to the research questions for the study which were:

A primary question of:

> *What are the most effective and cost-effective methods of protecting and promoting the health and wellbeing of older workers, and of supporting workers who wish to continue in employment up to and beyond state pensionable age?*

A secondary question of:

> *What supports, or prevents, the implementation of these methods?*

2.3.1 Inclusion criteria

Populations to be included

- All adults aged at least 50 in full or part-time employment, both paid and unpaid, self-employed people working in micro, small, medium and large organisations with an appointed line manager, and volunteers
- All employers in the public, private and 'not for profit' sectors who employ at least one employee

Interventions and policies to be included

- Interventions intended to address the research question primarily involving or aimed at employees aged over 50
- Interventions addressing entire workforces where at least 51% of employees are aged over 50
- Interventions targeted at 'older' workers aged below 50 where the intervention has an impact on them at age 50 or above
- Interventions delivered by third party organisations commissioned by organisations to deliver these within the workplace

Locations to be included

- Developed/OECD countries, major European countries outside the EU, and European countries which acceded to the EU in or before 2004 – please see list in Appendix 2
- Workplace settings or community level interventions aimed at workers rather than general population

Time period

- Studies published since 2005

Study types

- Experimental quantitative studies including:
  - before and after studies
  - non-randomised controlled studies (NRCS)
  - randomised controlled trials (RCT)
  - systematic reviews or meta-analyses
- Observational quantitative studies:
  - before-and-after studies
  - cohort studies
  - interrupted time studies
- Economic studies
  - cost–benefit analyses
  - cost-effectiveness analyses

2.3.2 Exclusion criteria

Excluded population groups

- Self-employed individuals working in organisations without appointed line managers
- Sole traders
- Unemployed individuals

- Interventions aimed at the general public rather than people working in specific organisations
- Studies covering interventions aimed at all employees where the majority (at least 51%) are aged under 50, unless a specific differential impact (either positive or negative) is found for workers aged at least 50

Interventions and policies that are excluded

- Intervention or support that employees accesses on their own initiative, without prompting from the employer, organisation or line manager or other third party (eg trade union).
- Statutory provision to employees
- The effectiveness of specific interventions to promote physical activity, smoking cessation in the workplace, to manage sickness absence and the return to work of those who have been on long-term sick leave, and mental wellbeing of which the first three topics are already covered by NICE guidance
- Interventions delivered without targeting specific worker populations

Locations to be excluded

- Developing and non-OECD countries
- Countries which acceded to membership of the EU later than 2004.

Study types to be excluded

- Non English language studies
- Qualitative studies

2.4 Outcomes

The outcomes of interest to this review include the following:

- **Organisation:** employee health and wellbeing and engagement; levels of employee recruitment and retention for the relevant age group; days lost to sickness absence (and reasons for absence); presenteeism; changes to work content, working time volume/patterns, flexible working practices; organisational measures of productivity; uptake of support services; return to work rates, job retention, measures of work ability, length of service, equality and diversity monitoring data (eg composition of workforce with health

conditions/disabilities); organisational HR data with relevance to staff wellbeing (eg survey results pertaining to HSE's Management Standards, staff surveys more generally); RIDDOR data indicating health and safety outcomes; incidence of age-related discrimination grievances/disciplinaries/employment tribunal claims; all available economic data; business outcomes such as labour turnover, productivity; customer service; profitability; health related behaviours/diseases.

- **Employee**: individual levels of health and wellbeing, motivation, individual performance, stress and job satisfaction;; perceptions of fair treatment; awareness, availability and uptake of training and support services; changes in work patterns and tasks (including changes in work/life balance); knowledge and awareness among managers and rest of workforce; impact on knowledge, skills and behaviour, including outcomes post-retirement such as financial status, social inclusion/isolation, civic participation, loneliness/mental health, physical health, self-reported quality of life.

2.5 The search for evidence

A single search to cover RQs 1, 2 and, 3 and the economic evaluation was conducted of key databases in health and medicine, social studies and business and management. A separate search for theses and dissertations was undertaken but due to the volume of material, theses and dissertations were not taken forward for inclusion in the sifting as it was judged that significant findings of publishable quality would picked up through the search of peer reviewed journal articles and grey literature.

2.5.1 Databases searched

General

- Academic Search Complete (via Ebsco)
- Scopus (Elsevier)
- Web of Science (includes SSCI) (Thomson Reuters)

Business and social science

- ABI/Inform (via Proquest)
- AgeInfo and NDAR (Ce`ntre for Policy on Ageing)
- Assia (via Proquest)

- Business Source Premier (via Ebsco)
- Campbell Collaboration (Native interface)
- International Bibliography of the Social Sciences (via Proquest)
- EconLit (via Ebsco)
- EPPICentre databases – DoPHER and TRoPHI (Native interface)
- SCIE (Native interface)
- Social Policy and Practice (via NHS Evidence)
- Sociological Abstracts (via Proquest)
- XPertHR (Native interface)

Health and Medicine

- AMED (Ebsco)
- Cochrane (Wiley)
- EMBASE (OVID)
- HMIC (HDAS)
- Health Business Elite(HDAS)
- Medline (OVID)
- PsycINFO (Ebsco)

2.5.2 Additional cost effectiveness search

In addition to the general searches for RQs 1-3, a specific cost effectiveness search for the economic evaluation is being conducted using the following sources:

- Cost-effectiveness Analysis (CEA) Registry (https://research.tufts-nemc.org/cear4);
- EconLit
- Embase (via OvidSP)
- Health Economic Evaluations Database (HEED)

- MEDLINE (via OvidSP)
- NHS Economic Evaluation Database (NHS EED)
- RePEc (Research Papers in Economics) (http://repec.org/)

2.5.3 Grey literature search

In addition to searching traditional academic databases the search process also covered 'grey literature', ie material that was not published in academic media or was in the process of publication. The following approach was adopted to the search through grey literature:

- A thorough search using the deep web search engine MEDNAR was conducted
- A thorough search of Google Scholar was conducted to identify grey literature, unpublished although peer reviewed conference papers, policy reports and theses E-mail alerts were set up to automatically notify the team of any new publications or grey items within the search parameters
- BASE (http://www.base-search.net/) was searched, specifically for material in institutional repositories
- Resources and directories available through Greynet International (www.greynet.org) were examined to locate any other compendia and direct links to grey literature not covered by other sources

Websites

A range of relevant policy and other agencies were searched, including the following UK sites:

- Acas: http://www.acas.org.uk/
- Age UK: http://www.ageuk.org.uk/
- Association of Chartered Physiotherapists in Occupational medicine http://www.csp.org.uk/tagged/association-chartered-physiotherapists-occupational-health-ergonomics-acpohe
- British Chambers of Commerce (BCC): http://www.britishchambers.org.uk/
- British Psychological Society: http://www.bps.org.uk/
- Centre for Employment Studies Research: http://www1.uwe.ac.uk/bl/bbs/research/cesr.aspx

- Centre for Mental Health: http://www.centreformentalhealth.org.uk/
- Chartered Institute of Environmental Health: http://www.cieh.org/
- Chartered Management Institute: http://www.managers.org.uk/
- CIPD: http://www.cipd.co.uk/
- College of occupational therapy –work section http://www.cot.co.uk/cotss-work/cot-ss-work
- Department for Work and Pensions: https://www.gov.uk/government/organisations/department-for-work-pensions
- Department of Health: https://www.gov.uk/government/organisations/department-of-health
- DWP Fuller Working Lives: A Framework for Action https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/319872/fuller-working-lives.pdf.
- EEF: http://www.eef.org.uk/
- Employers' Forum on Age (part of the Employer Network for Equality and Inclusion): http://www.efa.org.uk/
- HSE: http://www.hse.gov.uk/
- Investors in People: http://www.investorsinpeople.co.uk/about-us/our-organisation-achieving-success-through-people
- IOSH: http://www.iosh.co.uk/
- London Health Commission: http://www.londonhealthcommission.org.uk/
- National Audit Office: http://www.nao.org.uk/
- NHS Working Longer Review http://www.nhsemployers.org/PayAndContracts/NHSPensionSchemeReview/ImpactofWorkingLongerReview/Pages/NHSWorkingLongerReview.aspx.
- NICE (including former Health Development Agency document search) and NHS Evidence: http://www.nice.org.uk/
- Oxford Health Alliance: http://www.oxha.org/

- Public Health Observatories: http://www.apho.org.uk/
- Scottish Government: http://www.scotland.gov.uk/
- Sloan Centre for Ageing at Work http://capricorn.bc.edu/agingandwork/database/browse/facts/fact_record/5670/all
- UK Commission for Employment and Skills: https://www.gov.uk/government/organisations/uk-commission-for-employment-and-skills
- Welsh Government: http://wales.gov.uk/
- 'Working Late' research programme on the New Dynamics of Ageing www.workinglate.org/
- Xpert HR: http://www.xperthr.co.uk/

In addition we searched the sites of the following international bodies:

- Cedefop: http://www.cedefop.europa.eu/
- Eurofound: http://www.eurofound.europa.eu/
- European Commission: http://ec.europa.eu/index_en.htm
- EU-OSHA:https://osha.europa.eu/
- EuroHealthNet: http://eurohealthnet.eu/
- Finnish Institute of Occupational Health: http://www.ttl.fi/en/Pages/default.aspx
- Institute for Work and Health: http://www.iwh.on.ca/
- International Commission of Occupational Health: http://www.icohweb.org/
- International Labour Organisation: http://www.ilo.org/global/lang--en/index.htm
- Liberty Mutual Research Institute for Safety: http://www.libertymutualgroup.com/omapps/ContentServer?pagename=LMGroup/Views/LMG&ft=2&fid=1138356633468&ln=en
- Organisation for Economic Co-operation and Development: http://www.oecd.org/unitedkingdom/

- The National Institute for Occupational Safety and Health: http://www.cdc.gov/niosh/
- World Health Organisation: http://www.who.int/en/

2.5.4 Call for Evidence

The NICE project team issued a Call for Evidence on 10 June 2014 which closed on 10 July 2014 and asked for interested parties to send in evidence of relevance to the reviews. NICE issued a second Call for Evidence on 9 March 2015 which closed on 27 March 2015 with a specific focus on evidence gaps identified through the search and review process.

2.5.5 Contacting experts

To supplement the Call for Evidence a range of key academics, researchers and commentators in the field, known to the research team, PHAC members or recommended by the NICE project team were contacted and asked for any appropriate references.

2.5.6 Reference searching

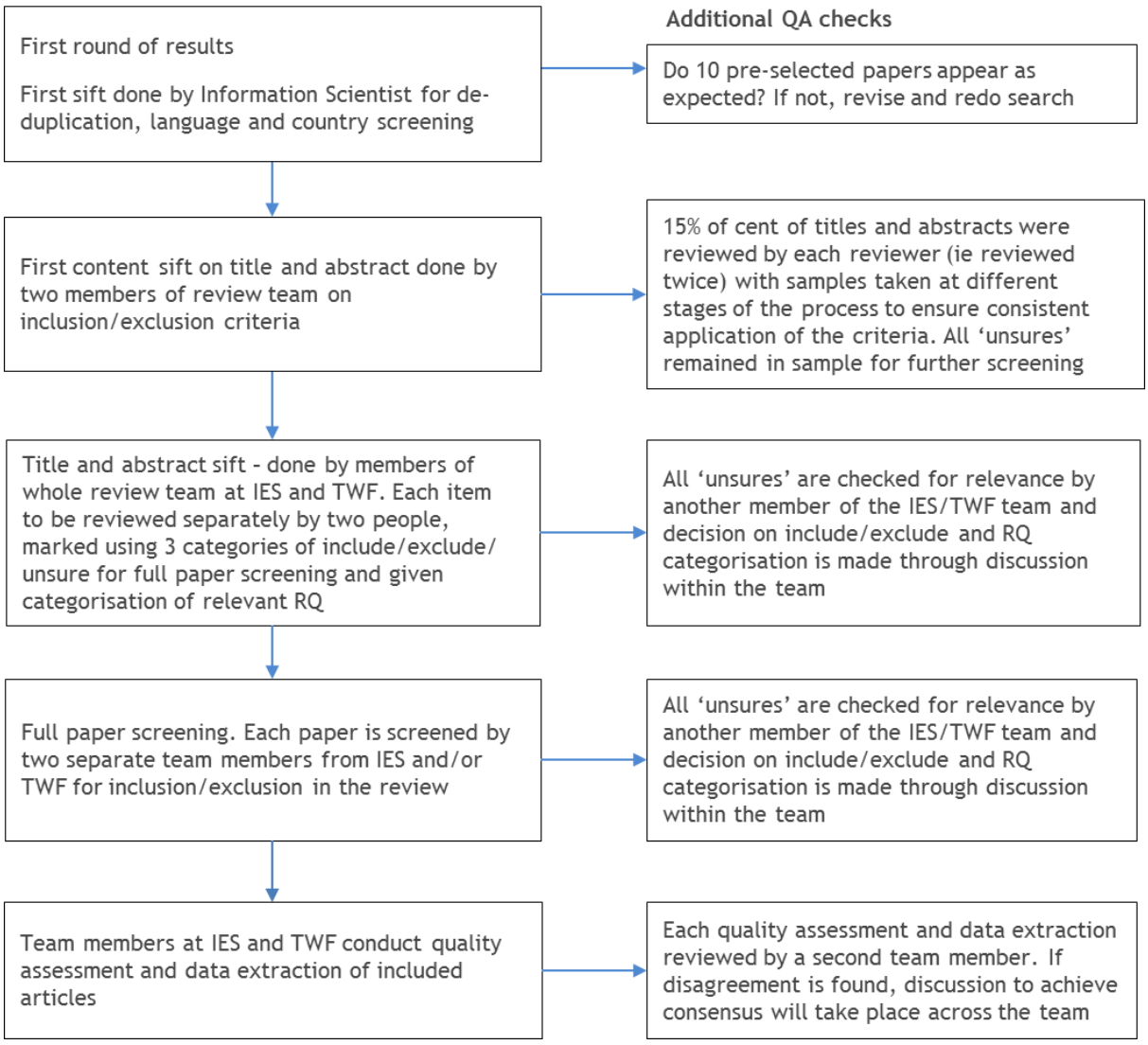
Once papers for initial inclusion were identified, the reference lists of these articles will be checked for any additional references. These articles were checked in Web of Science and GoogleScholar to identify citing articles.

2.6 Screening and data extraction

The process for sifting and screening material identified through the search and extracting the relevant evidence is summarised in Figure 2.1. The titles and abstracts of the papers identified through the initial search were downloaded into EndNote and screened for relevance using the inclusion and exclusion criteria, using a three-stage process involving:

- An initial sift based on title and abstract
- A second screening stage based on title and abstract and allocation to RQ1, 2 or 3
- A full paper screening.

Figure 2.1: Outline of sift and screening process



Initial sift

The titles of all material identified through the search were de-duplicated, checked that they conform to the inclusion criteria on language, date and country and quickly reviewed against the inclusion and exclusion criteria by two members of the review team. Fifteen per cent of the titles and abstracts were reviewed by each reviewer (ie reviewed twice) with samples taken at different stages of the process to ensure consistent application of the criteria[1].

[1] The first 1,000 titles and abstracts were reviewed by both researchers and the kappa statistic was 74 per cent. The papers where the two reviewers disagreed were discussed and an understanding

Second title and abstract screening

The titles and abstracts of all papers which came through the initial sift were separately reviewed against a checklist based on the full inclusion and exclusion criteria by two members of the review team (ie reviewed twice) and identified for full paper screening and the results recorded in the review database.

At this point, the included papers were tagged according to whether the paper was relevant for RQ 1, 2 or 3 and/or the economics review. Where there was disagreement between the reviewers a third member of the team reviewed the paper and reached a consensus with the other two reviewers.

Full paper screening

Each full paper was separately screened against a checklist based on the full inclusion and exclusion criteria by two members of the review team and identified for inclusion (or exclusion) for one of the reviews. Where there was disagreement a third member of the team also reviewed the paper and a consensus was reached with the other two reviewers.

2.7 Outcomes of the search process

A series of databases were searched by an Information Scientist at the Lancaster University library between 21 July and 16 August 2014, see Table 2.1.

reached on what met and did not meet the inclusion and exclusion criteria. Subsequent two further batches of 600 and 570 papers were double sifted and the results compared with kappa statistics of 87 per cent and 89 per cent respectively.

Table 2.1: Summary of literature databases searched

Database Name	Platform	Number of titles and abstracts downloaded to EndNote database
Academic Search Complete	EBSCO	5,956
Scopus	Elsevier	1,227
Web of Science (includes SSCI)	Thomson Reuters	2,692
Business and social science		
ABI/Inform	ProQuest	624
AgeInfo (Centre for Policy on Ageing)	Native	56
Assia	ProQuest	3,598
Business Source Premier	EBSCO	1,568
Campbell Collaboration	Native	0
EconLit	EBSCO	217
EPPICentre databases	Native	0
International Bibliography of the Social Sciences	ProQuest	206
Social Care Online (from SCIE)	Native	0
Social Policy and Practice	OVID	1,386
Sociological Abstracts *searched with ASSIA*	ProQuest	
XPertHR	Native	3
Health and Medicine		
Cochrane (Wiley)	Native	101
EMBASE	OVID	817
HMIC	HDAS	103
Health Business Elite	HDAS	861
Medline	OVID	5,781
Medline-in-process	OVID	50
PsycINFO	EBSCO	1,948
Theses and Dissertations		
Index to Theses	Native	19
Digital Dissertations	ProQuest	525
Total		27,738

Source: IES/Work Foundation/Lancaster University, 2014

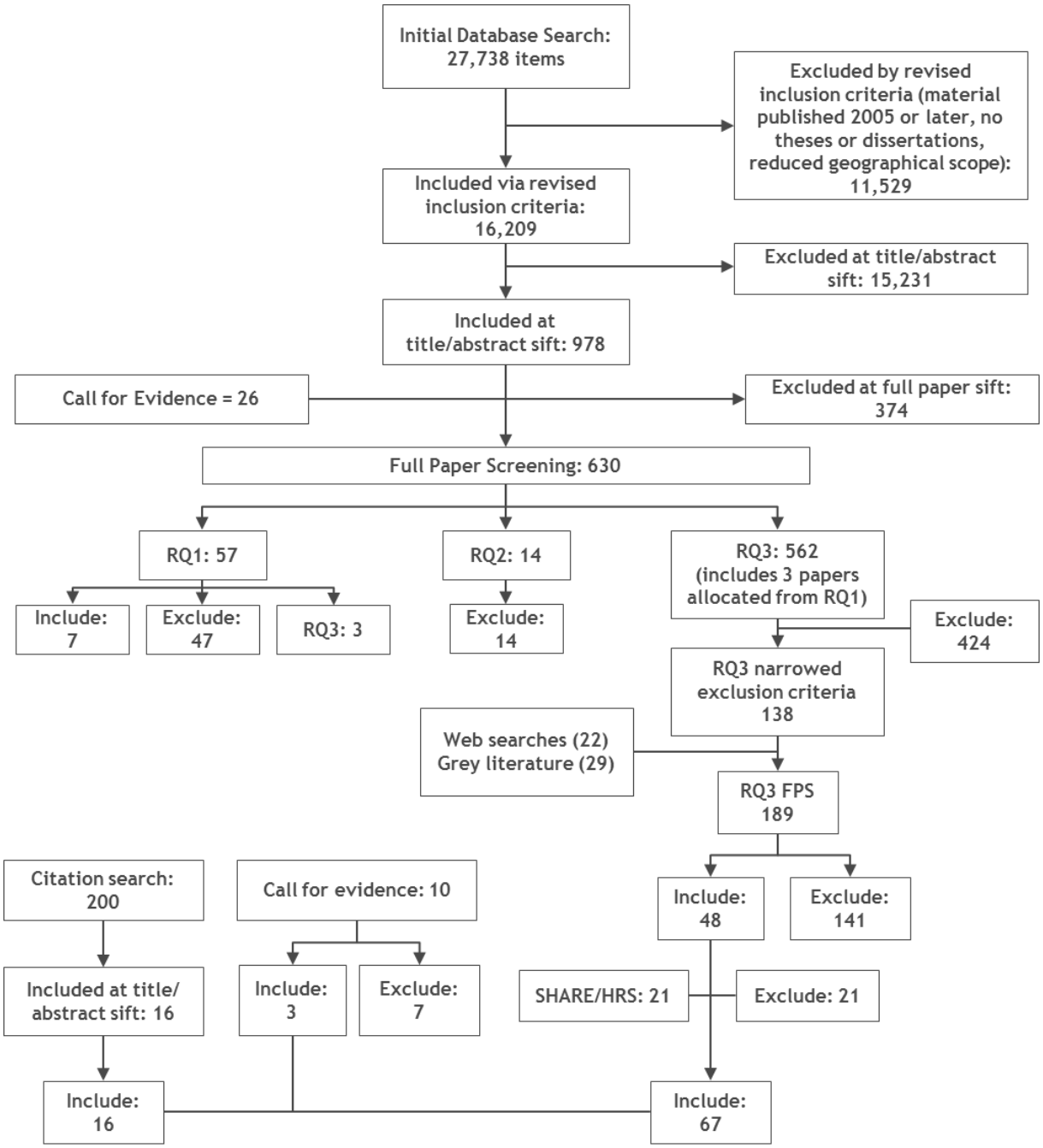
The search strategies were designed to cover: workplace interventions to support the health, well-being and continued employment beyond normal retirement age of older workers, pre-retirement training, advice, guidance and mentoring; (cost-) effectiveness and health and well-being outcomes. Examples of the strategies used are set out in Appendix 4 and the results set out in Table 2.1. The titles and abstracts identified through the searches were recorded in an EndNote database.

Following the searching and screening process a total number of 630 papers were identified for full paper screening. This represents a considerable reduction from the original volume of papers identified through the search strategy. To manage the volume of literature gathered, additional criteria were introduced to focus the scope of the research to papers published since 2005, exclude dissertations and theses since

data from them would have made its way into peer reviewed journals and to focus on OECD countries and European countries joining the EU in or before 2007. In practice, large volumes of the papers returned by the searches proved not to be relevant to the review. A large volume of literature consisted of technical papers on retirement or pensions legislation, another large segment dealt with the domiciliary or residential care of older people, a further segment dealt with national policy on retirement ages or pensions policies and a further segment consisted of news items reporting the imminent or actual retirement of prominent business figures.

All of the papers put forward for full paper screening have been obtained and screened and the results are summarised in Figure 2.2. Seven papers were included in this first review.

Figure 2.2: Outcome of search process for Review Questions 1, 2 and 3



Source: IES, TWF, Lancaster University

2.8 Data extraction

The seven papers identified for inclusion in this review were assessed for quality and the data extracted and presented in an evidence table. The evidence from each paper was extracted and the quality of the paper appraised by a member of the IES/TWF review team and then checked and re-appraised by another. A narrative summary of the evidence table was also produced.

2.8.1 Quality appraisal

Papers were assessed using a checklist based on the quality assessment in the NICE Public Health Guidance Methods Manual (NICE, 2012). As a result papers were graded either:

++ All or most of the checklist criteria have been fulfilled; where they have not been fulfilled the conclusions are very unlikely to alter

+ Some of the checklist criteria have been fulfilled, where they have not been fulfilled, or not adequately described, the conclusions are unlikely to alter and

- Few or no checklist criteria have been fulfilled and the conclusions are likely or very likely to alter.

The checklist is included in Appendix 2.

2.8.2 Data extraction

For each paper the evidence table, which follows the format set out in Methods for development of NICE public health guidance (third edition, 2012) summarises:

- The key research aims
- The study quality rating
- The research design and methodology
- The findings that contribute to the research questions
- Limitations and gaps
- Summary information about authors, publication etc.

2.9 Evidence synthesis

The findings from studies have been synthesised and where appropriate grouped thematically and an evidence statement(s) generated for each theme (Chapter 4).

During development of the evidence statements and synthesis the relevance of the findings to the UK context was also assessed, based on the following criteria:

- The population involved
- The setting, including the country or countries and type of workplaces in which the study took place

- The intervention and whether it would be appropriate for the UK
- The reported outcomes.

2.10 Excluded studies

Appendix 7 provides the reference details of 26 excluded studies at the full paper screening stage for Review Question 1. Studies were excluded because they failed to meet at least one of the inclusion criteria. As soon as they failed to meet one of the criteria they were excluded. In the appendix the references are ordered by the criterion by which they were excluded. They may have failed against other criteria too.

Six were excluded because the methodology was either a qualitative study, did not explicitly measure health and wellbeing, had no control group or longitudinal element and seven were rejected on grounds of relevance, eg they did not study the influence of interventions to protect or promote older workers' wellbeing and working capacity beyond normal retirement age. Ten were excluded because they did not focus on an intervention or they did not focus on an intervention being applied to workers aged at least 50. Three were excluded because of their focus on chronic illnesses.

3 Findings

A total of seven studies met the criteria for inclusion in this first review and focussed on workplace policies, practices or interventions implemented in employing organisations that contained evidence about the effectiveness or cost effectiveness of interventions to protect and promote the health and wellbeing of older workers, and to support workers who wish to continue in employment up to and beyond state pensionable age.

The studies are summarised below and the implications of the findings discussed in Chapter 4.

3.1 Summaries of the included studies

Härmä et al. (2006)

This (-) study involves a non-randomised controlled trial designed to assess the impact of changing from a backward rotating shift system of evening, morning, nights to a rapid forward rotating shift system of morning, evening, nights on sleep-wakefulness and well-being among shift workers aged 44 or younger and 45 or older in maintenance, inspection and supervisory roles in a Finnish aircraft maintenance setting.

Participation in the study was voluntary and employee representatives together with occupational health experts and managers selected the new shift system. Among an original baseline sample of 273 workers (60% of worker population within the employer), 40 participants in the new shift system responded to a baseline questionnaire, completed diary studies of sleep patterns and provided actigraph readings of waking and sleeping patterns. From these workers 24 took part in the follow-up questionnaire and nine took part in the field study compared to a control group of 116 workers who answered the baseline and follow-up questionnaire, of which nine also provided field measurements. The subjects were divided into younger (24–44 years) and older (45–61 years) but their distribution within the control and intervention groups is not stated in the study.

Survey data were collected from the participants in both intervention and control groups approximately 1.5 years before intervention and 6 months after covering:

- Sleep diary including estimated time taken to falling asleep (sleep latency), the number of wakings, estimated awaking time between waking and falling asleep, feeling of waking too early and feeling of insufficient sleep using a five point scale ranging from fully/enough to clearly/not enough
- Subjective measures of sleepiness using the Karolinska Sleepiness Scale (KSS) (Akerstedt and Gillberg, 1990)
- Self-reported general wellbeing measures covering questions on 'How does the current shift system affect your, a) sleep and vigilance, b) well-being at work, c) general health, d) social life, e) family life and f) hobbies' using a 5 point Likert type scale from 'improves considerably' to 'disturbs considerably' (modified from Barton et al. 1995)
- Additional direct questions to the whole intervention group to combat absence of baseline data for 16 subjects eg 'How did the new shift system affect your alertness, sleep and performance in the different shifts?' using a 3 point scale of decline, no change or improvement.

The field study consisted of:

- Portable actigraph readings of waking and rest patterns made using a wrist monitor during one complete shift cycle of 2-5 days
- Assessments of objective cognitive-motor performance via a computer-based Psychomotor Vigilance Task (PVT) using speed and accuracy of hand reaction responses to an LED visual stimulus.

The data were analysed using a linear mixed model to assess the presence of associations between a range of variables including the group (intervention vs. control), time (before or after the intervention), age (44 years or younger, 45 years and older) and different shifts, with subjective perceptions of sleepiness and well-being from questionnaire responses, actigraph readings and results of PVT test.

Outcomes

The change of the shift system was associated with an improved perception of the effects of the shift system on sleep quality including insomnia, health, well-being at work, and free-time activities among workers of all ages with:

a. greater perceived improvement among workers aged at least 45 of sleep and vigilance (group * time * age: df 4, 404, $F = 9.5, p < 0.0001$), general well-being at work (group *time *age: df 4, 413, $F = 10.0$, $p < 0.001$), social life (group *time *age:

df 4, 416, F =6.4, p <0.0001), family life (group * time *age: df 4, 408, F =5.0, p <0.0006) and hobbies (group *time *age: df 4, 416, F =3.2, p <0.01)

b. positive and significant association between shift changes, sleep efficiency and sleep fragmentation (se: df 32/791, F =1.58, p <0.0223; FI: df 32/ 797, F =1.50, p <0.0389) with small improvements in among workers aged at least 45

c. significantly decreased severe sleepiness during free time after night shifts for younger and older workers (group*time: df 3/385, F =4.9, p <0.03)

d. significantly lower self-reported sleepiness during free-time after the morning shifts among workers aged at least 45 (Age: df 1,413, F =6.1, p <0.01)

e. a positive and significant five-way interaction between group, time, age, shift and time of the shift on self-reported sleepiness (df 1/80, F =2.64, p <0.001) with greatest decrease in sleepiness among workers aged at least 45 during the night shift

f. a positive and significant four-way interaction of group, time, shift and age on the PVT test (df 41/590, F =4.04, p <0.0001) showing especially a decrease of the median reaction times at the end of the night shift among workers aged at least 45 (mean ±s.e. from 376 ±18 to 353 ±15 ms)

Limitations of the study

The study was rated '-', because the sample sizes involved are small, participation in the intervention was determined by self-selection, some participants in the intervention joined the study after the baseline round of research was completed and attrition in research participants during the study is not explained. PVT readings were taken only during the first and last two hours of each night shift and the change in the shift systems means that PVT the reaction times at the start of the night shift are not directly comparable between the two shift systems.

Applicability to the UK

This study appears to be partially applicable to the UK. It is set in an international airline company in Finland where work tasks and practices in different countries are similar due to international regulation and there is nothing specific to the nature of the intervention itself to exacerbate difficult of transferring the intervention. However, different industrial relations cultures in Finland and the UK may mean the process of change and implementation would not be precisely replicable.

Evidence Statement 1: shift patterns

There is weak evidence from one (-) study[1] 'before and after' non-random controlled and longitudinal study set in Finland on male aircraft maintenance workers aged at least 45 that changing from a backwards to a rapidly forwards rotating shift system can result in significant positive changes in self-reported sleepiness after the morning shifts (Age: df 1,413, F =6.1, p <0.01) , sleep quality, quality of life indicators including sleep and vigilance (group * time * age: df 4, 404, F = 9.5,p <0.0001), general well-being at work (group *time *age: df 4, 413, F =10.0, p <0.001), social life (group *time *age: df 4, 416, F =6.4, p <0.0001), family life (group * time *age: df 4, 408, F =5.0, p <0.0006), hobbies (group *time *age: df 4, 416, F =3.2, p <0.01), and psychomotor performance with a significant decrease of the median reaction times at the end of the night shift among the older workers (mean ±s.e. from 376 ±18 to 353±15 ms).

This evidence appears to be mostly applicable to the UK because of likely standardisation in work content and processes due to international regulation of aircraft maintenance, but concerns about its quality in terms of the small sample sizes and participant selection need to be taken into account.

[1] Harma et al. 2005 (-)

Rutanen et al. (2014)

This (+) rated study was a randomised controlled trial to assess the effects of an aerobic physical exercise intervention on perceived work ability, daily strain and mental resources among working women aged between 42 and 60 reporting menopausal symptoms in Finland.

The intervention included 6 months of aerobic exercise training in sessions lasting 50 minutes four times a week, with a progressive increase in intensity, of which at least two weekly sessions were supposed to be walking or Nordic walking and the other two could be jogging, cycling, swimming, skiing, aerobics/step aerobics or other gymnastic exercise.

Women were recruited to the study via an advertisement in the local newspaper with 56 allocated to the intervention group and 53 to the control group from a total of 176 original applicants, as women who did not work at least seven hours per week and women who exercised at least twice a week were excluded from the study. There were no significant differences between intervention and control groups at the baseline point except higher physical work demands for the intervention group.

Survey data obtained through daily mobile phone questionnaires and measures of physical fitness were collected from subjects in the intervention and control groups immediately prior to the intervention and after 6 months participation as follows:

a. self-reported questionnaire of seven items making up the Work Ability Index (WAI) including 1) Work ability in relation to lifetime best rated on a Likert-type 10 point scale, 2) Work ability in relation to physical and mental work demands rated on a Likert-type 5 point scale, 3) number of diagnosed diseases on a scale from 1 to 5, 4) work impairment from diseases rated on a Likert-type 6 point scale, 5) self-reported sick leave in the past 12 months on a scale from 1 to 5, 6) individual prognosis of work ability after 2 years (1, 4, 7), and 7) mental resources

b. questionnaires on physical and mental work strain were filled out in the evenings at baseline and end using a 5 point Likert-type scale (0 = very little, 5 = very much).

c. cardio-respiratory fitness was assessed by the UKK walking test measuring heart rate, walking time and BMI and estimates of maximal oxygen consumption.

Differences in questionnaire and fitness results at baseline between control and intervention groups were assessed by t-tests for normally distributed continuous variables, Mann-Whitney tests when non-normally distributed and Chi-square tests for categorical variables. Linear, ordinal and multinomial regression models were applied to test for post-intervention effects.

Outcomes

The results showed that physical exercise in the intervention group was associated with self-reported improvement in mental resources and a decrease daily physical work strain with statistically significant and positive improvements for the intervention group in reported mental resources (coefficient 0.58, 95% CI = 0.17- 1.00, $p < 0.01$) and physical work strain (coefficient -0.26, 95% CI = -0.45 - -0.07, $p < 0.01$). There were no significant findings for any other outcome.

Limitations of the study

The study was rated (+) because while attrition was lower than 20% in both intervention and control groups and the study adopted a randomised allocation process, this was not blinded and control group members were likely to know the intervention group members in a small community so could have increased their physical activity even though they did not report this. In addition, respondents reported relatively high levels of work ability at baseline which limited opportunities for further increase and the work ability index was developed to monitor change over a year so the six month time period used may not be sufficient to detect change.

Applicability to the UK

This study appears to be applicable to the UK as the intervention can be delivered in a similar way to a similar group of target participants and there appears to be no specific feature of the health or employment system which would potentially prevent this. Examples of exercise groups promoted in a workplace setting by employers have been implemented supplemented by campaigns run by employee representative organisations to support female employees experiencing symptoms of menopause.

Evidence statement 2: physical activity

There is weak evidence from one (+) study[2] using a randomised controlled trial on working symptomatic menopausal women set in Finland that regular physical exercise for this group can result in significant positive changes in self-reported mental resources (coefficient 0.58, 95% CI = 0.17-1.00, $p < 0.01$) and decreased daily physical work strain (coefficient -0.26, 95% CI = -0.45 - -0.07, $p < 0.01$).

This evidence appears to be fully applicable to the UK because there do not appear to be institutional differences which would mitigate the implementation of the intervention, although lack of blinding in allocation of participants to control and intervention groups may have resulted in changed behaviour among control group members, and early assessment of an outcome measure intended to be used 12 months rather than 6 months after the intervention may mean the study did not accurately assess the full potential impact of the intervention.

[2] Rutanen et al. (2014) (+)

Wagner et al. , 2007

This (+) study using a controlled trial was set in an inpatient clinic for mental therapies in Germany and sought to assess the impact of a cognitive training programme on objective and subjective memory performance and work attitudes in a sample of 92 patients aged 50-59 with mild cognitive impairment arising from memory, anxiety and depressive disorders.

Patients were referred to the clinic by their GP or health insurance company and screened for the type of disorder on admission, at which point those with dementia were excluded from the study and participants were then allocated to intervention and control groups through unspecified methods. Fewer than 8% of participants dropped out during the study and there were no significant differences between those who declined to participate and participants, or between intervention and control groups at baseline. Of 92 participants, 27 had a memory disorder, 31 suffered from additional impairments in other cognitive areas of functioning and 34 had cognitive impairments not involving memory.

The cognitive training programme tool place in interactive, closed group of four to eight participants over seven sessions lasting up to 90 minutes each based on behaviour analysis relating to prospective memory and structured processing of new information. Problem-solving techniques were collectively identified by the group to improve future performance. Participants were asked to develop ways to transfer and apply techniques in their everyday lives, and expected to practise and consolidate skills acquired during their daily routine at the clinic and through homework assignments.

Cognitive performance and attitudes to work during the first week after clinic admission and on discharge were assessed using:

- The Logical Memory I and II subsets of the Wechsler Memory Scale
- An Appointment Test which has correlations with other memory and attention tests and is an everyday simulation test to assess prospective memory, in which participants have to remember future appointments
- A Memory Assessment questionnaire developed by the clinic consisting of 49 items on a Likert-type 5 point scale about memory in daily life situations eg 'How would you rate your memory compared to when it was at its best?' (1= much more and 5 = much better), 'How worried are you about your memory right now?' (1= very worried and 5 = not worried at all),
- A work-related attitudes questionnaire covering issues such as occupational ambition, pursuit of perfection, ability to distance oneself from work, subjective significance of work, perception of inner balance, satisfaction with life, experience of success at work and resignation when faced with failure,

Data were analysed using SPSS with parametric and non-parametric procedures (t-test and ANCOVA) and results were adjusted for age and level of education.

Outcomes

The results showed that the intervention group demonstrated significant positive improvement compared to the control group following the intervention on:

- The Appointment Test ($F=15.06$, $p<0.001$)
- Logical Memory II test ($F=4.95$, $p<0.05$)
- Reduced anxiety about everyday memory function ($t=-2.83$, $p<0.01$) from the memory assessment questionnaire

- Decrease of pursuit of perfection ($t=3.23$, $p<0.1$) and reduced level of exhaustion ($t=-4.17$, $p<.001$) in the work-related attitudes questionnaire.

Responses to a subjective questionnaire at the end of the training showed that 82% of the intervention group felt able to apply the problem-solving strategies acquired during training at their workplace, 74% felt able to analyse difficulties with cognitive demands, 71% of the intervention group had learned strategies to help them to remember appointments and 70% had acquired methods to structure new information in a useful manner. 68% felt better able to accept variations in job performance, as a consequence of training, and 70% of the intervention group rated their overall cognitive ability as improved. It could be possible that reduction in broader psychological disorders such as depression contributed to cognitive improvements but the control group which had been exposed to the same treatment regime for these conditions with the exception of the cognitive performance intervention had also experienced a reduction in depressive symptoms but showed no increase in cognitive performance.

Limitations of the study

The study was rated (+) because while a controlled intervention was used, methods of allocation to the control group and intervention group are unknown, the authors were unable to determine the mechanisms underlying the increased memory performance in the intervention group and a long-term outcome was not assessed so it is not possible to comment on actual transfer to daily work and life performance.

Applicability to the UK

This study appears to be of weak applicability to the UK because the nature of work undertaken by the participants is unknown so may not be directly comparable to jobs in the UK labour market, it is not known whether the work attitudes test and clinic-specific memory questionnaire would achieve validity in the English language, and there may be potential differences in the traditions and systems of administering and referring people to psychological support, though employers may play a similar role in referring workers to psychological therapies through Employee Assistance Programmes and private health insurance schemes.

Evidence Statement 3: psychological support

There is weak evidence from one (+) study[3] using a controlled trial among inpatients aged 50-59 in a psychological treatment facility in Germany that developing and applying group-based problem-solving techniques can result in significant positive improvement following the intervention on ability to schedule appointments ($F=15.06$, $p<0.001$), a memory function test ($F=4.95$, $p<0.05$), reduced anxiety about everyday memory function ($t=-2.83$, $p<0.01$) and decreased pursuit of perfection ($t=3.23$, $p<0.1$) and reduced level of exhaustion ($t=-4.17$, $p<.001$) in a questionnaire of work-related attitudes.

The study has unknown applicability to UK workplaces because the nature of work undertaken by the participants is unknown so may not be directly comparable to jobs in the UK labour market, it is not known whether the work attitudes test and clinic-specific memory questionnaire would achieve validity in the English language, and there may be potential differences in the administration of psychological inpatient treatment.

[3] Wagner et al. 2007 (+)

Wegner et al. 2008

This (-) longitudinal study aimed to investigate whether an inpatient psychotherapeutic treatment with a job-specific element showed long-acting success on burn-out of schoolteachers in a rural area of Germany and whether there was any variation in impact by gender and type of school.

Participants aged between 27 and 64 years (mean 51.1 and SD ± 6.7 years) were referred by a medical practitioner and all of the 200 referred consented to participate in the treatment. The treatment consisted of a programme lasting around seven weeks delivered through a holistic mixture of physical and psychological interventions, including a 100 minute weekly group therapy meeting focussing on work related problems, three sessions of gestalt psychotherapy and physiotherapy, and two depth psychology discussions of 50 minutes each.

A survey of 60 items was administered at an unspecified time point at the beginning of the intervention and approximately two years afterwards, to which 150 of the original 200 participants responded giving a 25% attrition rate. Non-respondents to the follow-up questionnaire showed no significant differences from respondents, except for men with the diagnosis of personality disorders of which 20% were non-respondents versus 9% respondents of the total sample.

Data was analysed using t-tests with paired random samples to compare the results of survey periods, and t-tests with unpaired random samples for group comparisons as well as corresponding Chi square tests to check frequency differences.

The survey included: details on demographic data, job context, job performance, working time, medical history and the Maslach Burnout Inventory (MBI) in its

German translation covering 22 statements about feelings and attitudes that assess the three aspects of burnout: emotional exhaustion, depersonalisation and personal accomplishment measured on a 7 point Likert scale. Items included: 'I feel emotionally drained from my work', 'I have accomplished many worthwhile things in this job'.

Outcomes

Improvements in teacher health indicators post-treatment were found in the follow-up survey as follows:

- A statistically significant increase in the percentage of teachers who were not ill in the last quarter from 29.5% to 51.8% ($p < .001$)
- Disappearance of a statistically significant difference between burnout scores among high school teachers compared to those in other schools after treatment from 37.7 high school teachers and 26.5 other school teachers ($p < .05$) to 26.5 for high school teachers and 24.8 for other teachers ($p = .599$)
- Disappearance of a statistically significant difference of higher depersonalisation ($p < .0001$) and lower personal accomplishment scores ($p < .05$) in men compared with women after treatment
- Improvement in both sexes in scores of high emotional exhaustion, high depersonalisation, and low personal accomplishment.

The percentage of participants who had retired or were no longer teaching was positively related to older age in the follow-up survey.

The authors suggest that intervention could have greater effects if men especially were more willing to seek treatment and that the long-term effects of the intervention, particularly on delaying retirement age, could be greater if the intervention was begun earlier.

Limitations of the study

The study is rated as (-) for a variety of reasons. The authors note that the number of participating high school teachers compared to those from other school types is relatively low and the lack of control group reduces the validity of the results. The suggestion made by the authors that older workers would benefit from earlier support is not empirically tested within the scope of the study and requires validation.

Applicability to the UK

This study has limited applicability to the UK. The authors note that teachers in Germany have special entitlement to this kind of inpatient intervention for burnout due to their status as civil servants, whereas teachers in the state education system in the UK have no specific occupational healthcare entitlements.

Evidence Statement 4: psychotherapy support

There is weak evidence from one (-) longitudinal study[1] that an inpatient psychotherapeutic treatment with a job-specific element had a long-term impact on burn-out of schoolteachers following inpatient treatment in a rural area of Germany. Improvements in teacher health indicators post-treatment were found in increased incidence of teachers without sickness absence in the previous quarter from 29.5% to 51.8% ($p < .001$), disappearance of a statistically significant difference between burnout scores among high school teachers compared to those in other schools after treatment from 37.7 high school teachers and 26.5 other school teachers ($p < .05$) to 26.5 for high school teachers and 24.8 for other teachers ($p = .599$), disappearance of a statistically significant differences in higher depersonalisation ($p < .0001$) and lower personal accomplishment scores ($p < .05$) in men compared with women after treatment.

There was also improvement in both sexes in scores of high emotional exhaustion, high depersonalisation, and low personal accomplishment and the percentage of participants who had retired or were no longer teaching was positively related to older age in the follow-up survey.

The authors speculated in the conclusions that older workers would benefit from earlier intervention to prolong working lives.

This study has limited applicability to the UK. The authors note that teachers in Germany have special entitlement to this kind of inpatient intervention for burnout through the terms of civil service employment contracts. In contrast teachers in the state education system in the UK have no specific occupational healthcare entitlements.

[1] Wegner et al. 2008 (-)

Midtsundstad and Nielsen (2014)

This pooled cross-sectional study using employer survey and matched employee administrative data at two time points used difference-in-difference comparisons to examine the effect of employer-initiated measures to reduce worker illness on the probability of sickness absence among workers aged over 50 in workplaces across a variety of industrial sectors in Norway.

Data was taken from a random sample survey of Norwegian establishments with a 73% response rate from sectors including: manufacturing, construction, retail, hotels and restaurants, public administration, education, health and social services, and 'other' industries and a cross section of employees aged 50 or older from national administrative data.

Forty-one per cent of establishments had some form of preventive workplace measure in place in 2007 with unknown start dates. The workplaces all employed at least 10 staff and had at least one employee aged 60 or older in 2005. Establishments with and without preventive measures were similar regarding distribution of gender, mean age of workers, their educational level and percentage of staff with a partial disability.

Detail on the nature of the interventions was not clear in the published paper so the review team contacted the lead author by email for clarification. The lead author responded by stating that the intervention could be any one of 12 possible measures including reduced working hours (with or without reduced pay), temporary or permanent change of occupation and free physical therapy, massage or exercise within working hours. The three most common measures implemented were work adaptation, changed work tasks and technical equipment which were implemented – alone or in combination – in workplaces covering 70% of the employees who had access to at least one intervention. Among employees in a workplace with one or more interventions in 2007, about 40% were in establishments using a single measure, 28% in establishments with two measures, and 12% were in establishments with three measures. Thus, one in five workers in establishments implementing the interventions were covered by four or more measures.

The sample size of workers in establishments with at least one intervention was 5885 at baseline in 2001 and 7957 in 2007, while the sample of workers in establishments without any of the interventions was 8376 in 2001 and 11,003 in 2007. The total sample of employees was 14,261 in 2001 and 18,960 in 2007.

The outcome measure chosen was sickness absence lasting at least 16 days certified by a medical practitioner identifiable through the administrative data on each employee.

Results were analysed using a difference-in-differences approach to assess changes in likelihood of sickness absence over time between with and those without access to preventive work place measures. The method used was Logistic regression and as a control, linear probability models used to substantiate reported estimates. Models were run, to adjust for employee characteristics such as age, income, disability and gender, and establishment characteristics, and separate models were run for each sector.

Outcomes

- The odds for sickness absence levels were about 20% higher for employees in establishments with at least one preventive measure compared to establishments without preventive measures. The authors suggest this may reflect the

introduction of measures in response to perceived problems with sickness absence.

- There is positive and statistically significant evidence that employees in establishments with at least one preventive measure experienced a 10% drop in the odds for sickness absence in the period from 2001 to 2007 (measures OR 1.20 CI 95% = 1.12-1.28, change 2001 – 2007 OR = 0.97, CI 95% = 0.9101.03, measure x change OR = 0.89, CI 95% 0.81-0.97). There was no change to the results after adjustment for individual characteristics.

- In public sector establishments with at least one preventive measure, there is positive and statistically significant evidence that the measures themselves have contributed to reducing sickness absence among employees aged 50 years or older (measures OR =1.70, CI 95% = 1.37–2.11; change 2001 – 2007 OR = 1.27 CI 95% = 1.06–1.52; measure x change OR = 0.60, CI 95% = 0.45–0.79)

- Levels of sick leave were explained in most other sectors by adjusting for industry with no impact from the presence of preventive measures. Sick-leave levels were the highest among employees in manufacturing, construction and in health- and social services. Levels were high in large establishments and low in establishments exposed to competition. In contrast, the presence of an HR-professional, experience of down-sizing within the last 5 years and signing up to a working life agreement negotiated between government and social partners which commits employers to reduce sickness absence rates by 20% from the 2001 rates do not have a significant impact on individual sick leave probability.

Limitations of the study

The study limitations include a lack of information on the timing of the introduction of the workplace measures, and the study only assessed the presence rather than the type of measures implemented. Importantly, it is not possible to tell whether the individuals surveyed were making use of the measures available. The authors state that establishments using and not using the measures are unlikely to be completely similar, neither are employees in each type of establishment and the authors may not have been able to control for all unobserved differences such as health status and precise working conditions.

Applicability to the UK

This study has weak applicability to the UK. Norway is similar to the UK as a developed Western economy with some similar industries but has a different industrial relations system and potentially different management approaches and attitudes to managing older workers. Because of the limited detail on the nature of

some of the more substantive interventions, it is not clear how easily they could be applied for the benefit of older workers in the UK.

Evidence Statement 5: workplace measures

There is weak statistically significant evidence from one cross-sectional pooled study[5] set in workplaces in Norway of a 10% drop in the odds for sickness absence in the period from 2001 to 2007 among employees aged over 50 in establishments using preventive measures (unfortunately not possible to identify in the paper) (measures OR 1.20 CI 95% = 1.12-1.28, change 2001 – 2007 OR = 0.97, CI 95% = 0.9101.03, measure x change OR = 0.89, CI 95% 0.81-0.97). There is positive and statistically significant evidence that adoption of at least one measure has contributed to reducing sickness absence among employees aged 50 years or older in public sector workplaces (measures OR =1.70, CI 95% = 1.37–2.11; change 2001 – 2007 OR = 1.27 CI 95% = 1.06–1.52; measure x change OR = 0.60, CI 95% = 0.45–0.79).

This evidence is weakly applicable to the UK because while Norway has a similar economy there is limited information on the nature of the interventions and how they were designed and applied in an industrial relations context which is different from the UK.

[5] Midtsundstad and Nielsen (2014) (+)

Two studies of health promotion activities showed positive links with changes in health behaviours among older workers.

Strijk et al. 2012

This (+) rated study set in two academic hospitals in the Netherlands sought to evaluate the effectiveness of a worksite vitality intervention on vigorous physical activity (VPA), fruit intake, aerobic capacity, mental health and need for recovery after work among hospital workers aged at least 45 years.

The sample consisted of people invited to participate who were screened to select those working at least 16 hours per week with no risk for developing adverse health effects. Workers were required to give written informed consent and those participating were allocated using random allocation software. In the intervention group 75% were female with a mean age of 52.5 years (SD=4.8). In the control group 76% were female with a mean age of 52.3 years (SD=4.9).

The 6-month intervention consisted of (1) a Vitality Exercise Program (VEP) with (2) provision of free fruit and combined with (3) three visits to a Personal Vitality Coach (PVC). The VEP consisted of a weekly 45 min: (1) yoga session, (2) workout session and (3) unsupervised aerobic exercise session.

The measures used were as follows:

- Physical activity (PA): measured subjectively through a questionnaire and objectively using accelerometers. Outcome measures were total minutes per week of: (1) sports activities, (2) VPA, and (3) total moderate-to vigorous physical activities (MVPA).
- Weekly fruit intake: self reported via questionnaire
- Aerobic capacity (VO2max): estimated using the UKK 2 km walk test. Workers walked briskly for 2 km, with heart rate and performance time monitored, from which VO2 max was estimated.
- Mental health: Questionnaire that refer to the past 4 weeks: 'Did you feel… (1) nervous, (2) down in dumps, (3) peaceful, (4) sad and (5) happy'.
- Need for recovery (NFR) was assessed with a questionnaire consisting of 11 statements (yes/no) concerning the recovery period after a day's work.

At baseline, data on potential confounders and effect modifiers were assessed by questionnaire including age, gender, education, chronic disease status, smoking, intervention location, type of work and marital status.

Analysis tested for differences were using independent t tests for continuous variables and Pearson's tests for categorical and dichotomous variables. Differences in change over time between the intervention and control group were analysed using linear regression. For the sensitivity analyses, missing data were imputed using multiple imputations based on Multivariate Imputation by Chained Equations.

Outcomes

The results showed that the intervention significantly increased participants' weekly sports activities (ß = 40.4 minutes per week, $p<0.05$) and fruit intake (ß = 2.7 pieces per week, $p<0.05$), when compared to the control group. The intervention also favourably affected the need for recovery after a day of work (ß= - 3.5 points on a 100-point scale derived from the Experience and Evaluation of Work survey, $p<0.05$). No effects were observed for vigorous, aerobic capacity and mental health measures. A significant relationship was found between sports activity levels and high compliance to the guided yoga (b=49.6 min, 95% CI 13.9 to 85.2) and workout sessions (b=72.9 min/week, 95% CI 36.1 to 109.8) when compared to the control group. Also for fruit intake, effects were stronger in the high compliance group of both the yoga (b=3.8 pieces, 95% CI 1.1 to 6.4) and the workout sessions (b=4.0 pieces/week, 95% CI 1.1 to 6.4).

Limitations

The authors noted that the sample consisted of a relatively healthy population of older workers, mainly consisting of female workers, making it more difficult to generalise the study results. The study failed to ensure vigorous intensity physical activity compliance during the guided workout session which would be required to improve aerobic capacity. The review team noted that inclusion/exclusion criteria were not listed, making replication of the study potentially difficult.

Applicability to the UK

The findings are partially applicable to the UK. They cannot be generalised to all sectors as the research was focussed on two healthcare organisations.

Hughes et al. 2012

This (+)rated study on workers aged over 40 in a university in the USA set out to examine the effects of two worksite health-promotion interventions (compared with a health-education control) on older workers' health behaviours and health outcomes.

Participants were recruited through a broad array of recruitment strategies, including conducting recruitment events at locations frequently visited by the target population, mass e-mails, and campus listservs that sent biweekly messages to staff. Participants were allocated to intervention and control group through randomization sequences determined with custom software designed to achieve a balanced allocation of cases to conditions stratified by education and race/ethnicity.

The sample consisted of 423 participants of which 150 were in the COACH intervention group, 135 were in the RealAge intervention group, and 138 were in the control group. Their mean age was 51 (range 40-68), 82% were female, and 62% were from an ethnic minority group.

The interventions were:

a) COACH: Contact with a single coach trained in principles of behaviour change and motivational interviewing. Named the COACH programme because it was thought to have greater appeal to a working-age population. The coach was involved in health-risk assessment and discussion about potential behavioural change. The coach reviewed health-related goals and negotiated an action plan to meet them; this could be altered over time. For a week after the assessment, the coach asked participants about accessing resources needed to implement the plan, and revised the plans for those who were having difficulty doing so. The coach also contacted participants

biweekly during months 1-6, and monthly during months 7-12 to keep the goals up to date.

b) RealAge: Participants took the RealAge test online, which reviewed numerous health factors. After participants completed the test, the website generated individual risk profiles and indicated areas that could be improved. Participants used the website to select behaviours and create plans to meet behavioural goals. The site tracked each time a participant used it, and forwarded this information to the study team at regular intervals.

The outcomes were assessed using the following measures:

- Dietary behaviours: changes in per cent energy from fat and in fruit and vegetable intake from baseline to 6 months and 12 months.
- Physical activity (PA): Changes in vigorous activity, moderate activity, and level of exercise participation from baseline to 6 months and 12 months.
- Stress: Change from baseline to 6 months and 12 months in 4 measures of stress using scales developed to assess health-related stress, and coping behaviours.
- Smoking. Smoking cessation was defined as a minimum of 6 months of total abstinence from tobacco use at 6 months and 12 months after baseline, among participants who were current smokers at baseline or 6 months.
- Body mass index, waist circumference, and weight: For all 3 measures change from baseline at 6 months and 12 months assessed.

The analysis was conducted using mixed-effects regression models involving 2 between subjects' variables (group assignment and state of change (SOC)) and 1 within-subject variable (time). Group assignment was represented by 2 indicator variables for COACH and RealAge, with the control as the reference group. Time was represented with indicator variables for 6 months and 12 months relative to baseline. SOC indicated pre-contemplation, contemplation, or preparation (coded 1) versus action or maintenance (coded 0).

Participants varied with regard to the health behaviour; a respondent might be coded differently on SOC for different outcomes. The two-way interactions between group and time allowed determination of how patterns of change over time varied by group. 3-way interactions allowed assessment of whether differences over time between groups varied by SOC level. One-tailed tests of significance were used because seven prior studies of similar interventions had shown positive effects. Accordingly, a directional hypothesis at the conventional .05 significance level was chosen to detect a null or positive effect.

Outcomes

The study found that computerised health risk assessments combined with individualised, negotiated health improvement action plans and ongoing support and reinforcement from a coach had a positive effect on participants' diet ($z = 3.55$, $p = <0.001$) and physical activity ($z = 2.22$, $p= 0.13$) compared with a control group who received printed health promotion materials. However no effects were found for stress, smoking and weight. In addition, no positive effects compared with the control group were found for a second parallel intervention in which participants undertook an automated health risk assessment accompanied by self-directed use of on-line health modules and receipt of generic health e-mail tips.

Limitations

The authors noted that the interventions were tested with staff at a university who may have had higher levels of education than workers in other industries and so the generalizability of the findings to workers in other settings requires further testing. Cost/benefit analysis was not undertaken. Both interventions were light interventions and further study could include more intense interventions. It is also necessary to consider whether a dose-response relationship exists between comprehensiveness of services offered, and whether certain programme components, such as incentives, are more effective and more critical than others. The review team identified that effect size was not reported, confounding factors were not identified, and the generalizability also suffers from the imbalanced gender and ethnic minority breakdowns of the study.

Applicability to the UK

The findings are of limited applicability to the UK because the focus on workers in a single organisation limits the generalizability of the results to other organisations and sectors.

Evidence Statement 6: health promotion

There is moderate evidence from two (+) RCT studies[1,2] that health promotion programmes aimed at older workers can have positive effects on participants' diet and level of exercise.

One RCT[1] set among employees aged 45 and over in two academic hospitals in the Netherlands that a worksite vitality intervention (comprising exercise and yoga sessions, free fruit and visits from a coach) significantly increased participants' weekly sports activities (ß = 40.4 minutes per week, $p<0.05$) and fruit intake (ß = 2.7 pieces per week, $p<0.05$), when compared to the control group. The intervention also favourably affected the need for recovery after a day of work (ß= - 3.5 points on a 100-point scale derived from the Experience and Evaluation of Work survey, $p<0.05$). No effects were observed for vigorous, aerobic capacity and mental health.

A second RCT[2] set among workers aged 40 and over in a university in Chicago, USA found that computerised health risk assessments combined with individualised, negotiated health improvement action plans and ongoing support and reinforcement from a coach had a positive effect on participants' diet ($z = 3.55$, $p = <0.001$) and physical activity ($z = 2.22$, $p= 0.13$) compared with a control group (who received printed health promotion materials). No effects were found for on measures of stress, smoking and weight. No positive effects compared with a control group were found for a second parallel intervention in which participants undertook an automated health risk assessment accompanied by self-directed use of on-line health modules and receipt of generic health tips by email.

[1] Strijk et al. (2012) (+)

[2] Hughes et al. (2011) (+)

The interventions are not particularly intensive in delivery and could be provided to the entire workforce, consistent with a life course perspective of preventing worker ill health at any age. However, the outcome measures used in the studies are intermediate and it would be helpful to understand impact on health outcomes.

4 Discussion

This review includes evidence from seven studies about the way in which workplace interventions can affect the health and wellbeing of older workers. While the studies generally indicate that interventions can have a positive association with the wellbeing of older workers, they tend to focus on very specific interventions or types of older workers, only four focus on interventions made directly by employers in the workplace and do not contain data on the relative costs and benefits of each type of initiative. Therefore it is difficult to draw any general conclusions of interventions that should be recommended to employers on the basis of this evidence alone.

None of the studies are set in the UK and in some cases their applicability to a UK setting is limited due to differences in eligibility for and referral to psychological interventions. The relevance of the evidence base is also limited by the methodological quality of the studies, of which two are rated (-) and three are rated (+).

The lack of papers included in this review reflects a challenge that despite increasing policy interest in how the health and wellbeing of older workers can be supported, especially in the context of a decline in the proportions of people in some younger segments of the population and planned increases to the age for state pension eligibility, very few intervention studies were located. The wider specialist management trade press provides evidence of interventions being implemented to support older workers at an organisational level but such management-level interventions are rarely evaluated sufficiently rigorously to pass the inclusion criteria that we have applied to this evidence review. We expect more comprehensive and illuminating evidence to be generated by the next review which is likely to include a wider range of papers as Research Question 3 will include qualitative studies which are ineligible for this review. This review urges research commissioners to prioritise funding of high quality studies into the impact of workplace level interventions on health and well-being outcomes which will seek to track the health and well-being of individuals during the lifespan of the intervention and onwards to the end of their working lives.

Appendix 1: List of countries eligible for inclusion in the study

AUSTRALIA (OECD)

AUSTRIA (OECD, Europe)

BELGIUM (OECD, Europe)

CANADA (OECD)

CYPRUS (OECD)

CZECH REPUBLIC (OECD, Europe)

DENMARK (OECD, Europe)

ESTONIA (EUROPE)

FINLAND (OECD, Europe)

FRANCE (OECD, Europe)

GERMANY (OECD, Europe)

GREECE (OECD, Europe)

HUNGARY (OECD, Europe)

IRELAND (OECD, Europe)

ISRAEL (OECD)

ITALY (OECD, Europe)

JAPAN (OECD)

KOREA (OECD)

LATVIA (EUROPE)

LITHUANIA (EUROPE)

LUXEMBOURG (OECD, Europe)

MALTA (EUROPE)

NETHERLANDS (OECD, Europe)

NEW ZEALAND (OECD)

NORWAY (OECD, Europe)

POLAND (OECD, Europe)

PORTUGAL (OECD, Europe)

SLOVAKIA (Europe)

SLOVENIA (Europe)

SPAIN (OECD, Europe)

SWEDEN (OECD, Europe)

SWITZERLAND (OECD, Europe)

UNITED KINGDOM (OECD, Europe)

UNITED STATES (OECD)

Appendix 2: Quality Assessment Form and Checklist

Checklist items are worded so that 1 of 5 responses is possible:

++	Indicates that for that particular aspect of study design, the study has been designed or conducted in such a way as to minimise the risk of bias.
+	Indicates that either the answer to the checklist question is not clear from the way the study is reported, or that the study may not have addressed all potential sources of bias for that particular aspect of study design.
–	Should be reserved for those aspects of the study design in which significant sources of bias may persist.
Not reported (NR)	Should be reserved for those aspects in which the study under review fails to report how they have (or might have) been considered.
Not applicable (NA)	Should be reserved for those study design aspects that are not applicable given the study design under review (for example, allocation concealment would not be applicable for case control studies).

In addition, the reviewer is requested to complete in detail the comments section of the quality appraisal form so that the grade awarded for each study aspect is as transparent as possible. Each study is then awarded an overall study quality grading for internal validity (IV) and a separate one for external validity (EV):

++	All or most of the checklist criteria have been fulfilled, where they have not been fulfilled the conclusions are very unlikely to alter.
+	Some of the checklist criteria have been fulfilled, where they have not been fulfilled, or not adequately described, the conclusions are unlikely to alter.
-	Few or no checklist criteria have been fulfilled and the conclusions are likely or very likely to alter.

Study identification: (Include full citation details)		
Study design: Refer to the glossary of study designs (Appendix 5) and the algorithm for classifying experimental and observational study designs (Appendix 6) to best describe the paper's underpinning study design		
Guidance topic:		
Assessed by:		
Section 1: Population		
1.1 Is the source population or source area well described? Was the country (eg developed or non-developed, type of healthcare system), setting (primary schools, community centres etc.), location (urban, rural), population demographics etc. adequately described?	Quality Rating	Comments:
1.2 Is the eligible population or area representative of the source population or area? Was the recruitment of individuals, clusters or areas well defined (eg advertisement, birth register)? Was the eligible population representative of the source? Were important groups under-represented?		Comments:
1.3 Do the selected participants or areas represent the eligible population or area? Was the method of selection of participants from the eligible population well described? What % of selected individuals or clusters agreed to participate? Were there any sources of bias? Were the inclusion or exclusion criteria explicit and appropriate?		Comments:

Section 2: Method of allocation to intervention (or comparison)		
2.1 Allocation to intervention (or comparison). How was selection bias minimised? Was allocation to exposure and comparison randomised? Was it truly random ++ or pseudo-randomised + (eg consecutive admissions)? If not randomised, was significant confounding likely (−) or not (+)? If a cross-over, was order of intervention randomised?		Comments:
2.2 Were interventions (and comparisons) well described and appropriate? Were interventions and comparisons described in sufficient detail (ie enough for study to be replicated)? Was comparisons appropriate (eg usual practice rather than no intervention)?		Comments:
2.3 Was the allocation concealed? Could the person(s) determining allocation of participants or clusters to intervention or comparison groups have influenced the allocation? Adequate allocation concealment (++) would include centralised allocation or computerised allocation systems.		Comments:
2.4 Were participants or investigators blind to exposure and comparison? Were participants **and** investigators – those delivering or assessing the intervention kept blind to intervention allocation? (Triple or double blinding score ++) If lack of blinding is likely to cause important bias, score −.		Comments:
2.5 Was the exposure to the intervention and comparison adequate? Is reduced exposure to intervention or control related to the intervention (eg adverse effects leading to reduced compliance) or fidelity of implementation (eg reduced adherence to protocol)? Was lack of exposure sufficient to cause important bias?		Comments:

2.6 Was contamination acceptably low? Did any in the comparison group receive the intervention or vice versa? If so, was it sufficient to cause important bias? If a cross-over trial, was there a sufficient wash-out period between interventions?		Comments:.
2.7 Were other interventions similar in both groups? Did either group receive additional interventions or have services provided in a different manner? Were the groups treated equally by researchers or other professionals? Was this sufficient to cause important bias?		Comments:
2.8 Were all participants accounted for at study conclusion? Were those lost-to-follow-up (ie dropped or lost pre-,during or post-intervention) acceptably low (ie typically <20%)? Did the proportion dropped differ by group? For example, were drop-outs related to the adverse effects of the intervention?		Comments:
2.9 Did the setting reflect usual UK practice? Did the setting in which the intervention or comparison was delivered differ significantly from usual practice in the UK? For example, did participants receive intervention (or comparison) condition in a hospital rather than a community-based setting?		Comments:
2.10 Did the intervention or control comparison reflect usual UK practice? Did the intervention or comparison differ significantly from usual practice in the UK? For example, did participants receive intervention (or comparison) delivered by specialists rather than GPs? Were participants monitored more closely?		Comments:

Section 3: Outcomes		
3.1 Were outcome measures reliable? Were outcome measures subjective or objective (eg biochemically validated nicotine levels ++ vs self-reported smoking −)? How reliable were outcome measures (eg inter- or intra-rater reliability scores)? Was there any indication that measures had been validated (eg validated against a gold standard measure or assessed for content validity)?		Comments:
3.2 Were all outcome measurements complete? Were all or most study participants who met the defined study outcome definitions likely to have been identified?		Comments:
3.3 Were all important outcomes assessed? Were all important benefits and harms assessed? Was it possible to determine the overall balance of benefits and harms of the intervention versus comparison?		Comments:
3.4 Were outcomes relevant? Where surrogate outcome measures were used, did they measure what they set out to measure? (eg a study to assess impact on physical activity assesses gym membership – a potentially objective outcome measure – but is it a reliable predictor of physical activity?)		Comments: .
3.5 Were there similar follow-up times in exposure and comparison groups? If groups are followed for different lengths of time, then more events are likely to occur in the group followed-up for longer distorting the comparison. Analyses can be adjusted to allow for differences in length of follow-up (eg using person-years).		Comments:

3.6 Was follow-up time meaningful? Was follow-up long enough to assess long-term benefits or harms? Was it too long, eg participants lost to follow-up?		Comments:
Section 4: Analyses		
4.1 Were exposure and comparison groups similar at baseline? If not, were these adjusted? Were there any differences between groups in important confounders at baseline? If so, were these adjusted for in the analyses (eg multivariate analyses or stratification). Were there likely to be any residual differences of relevance?		Comments:
4.2 Was intention to treat (ITT) analysis conducted? Were all participants (including those that dropped out or did not fully complete the intervention course) analysed in the groups (ie intervention or comparison) to which they were originally allocated?		Comments:
4.3 Was the study sufficiently powered to detect an intervention effect (if one exists)? A power of 0.8 (that is, it is likely to see an effect of a given size if one exists, 80% of the time) is the conventionally accepted standard. Is a power calculation presented? If not, what is the expected effect size? Is the sample size adequate?		Comments:
4.4 Were the estimates of effect size given or calculable? Were effect estimates (eg relative risks, absolute risks) given or possible to calculate?		Comments:

4.5 Were the analytical methods appropriate? Were important differences in follow-up time and likely confounders adjusted for? If a cluster design, were analyses of sample size (and power), and effect size performed on clusters (and not individuals)? Were subgroup analyses pre-specified?		Comments:
4.6 Was the precision of intervention effects given or calculable? Were they meaningful? Were confidence intervals or p values for effect estimates given or possible to calculate? Were CI's wide or were they sufficiently precise to aid decision-making? If precision is lacking, is this because the study is under-powered?		Comments:
Section 5: Summary		
5.1 Are the study results internally valid (ie unbiased)? How well did the study minimise sources of bias (ie adjusting for potential confounders)? Were there significant flaws in the study design?		Comments:
5.2 Are the findings generalisable to the source population (ie externally valid)? Are there sufficient details given about the study to determine if the findings are generalisable to the source population? Consider: participants, interventions and comparisons, outcomes, resource and policy implications.		Comments:

The following sections outline the checklist questions, the prompts provided as pop-up boxes in the electronic version (highlighted in boxes) and additional guidance notes to aid the reviewer in assessing the study's internal and external validity.

Section 1

This section seeks to assess the key population criteria for determining the study's **external validity**.

Although there are checklists for assessing external validity of RCTs (with a particular focus on clinical interventions) (see for example [Rothwell 2005]), there don't appear to be any checklists specific for public health interventions.

The questions asked in this section ask the reviewer to identify and describe the **source** population of the study (that is, those the study aims to represent), the **eligible** population (those that meet the study eligibility criteria), and the **study participants** (those that agreed to participate in the study). Where a study assesses an intervention delivered to a particular geographical setting or area (rather than delivered to individuals), the questions in this section relate to describing the source area or setting, and how the study areas or settings were chosen. For example, a study might assess the effect on health outcomes of neighbourhood renewal schemes and this section seeks to identify and describe how those neighbourhoods were chosen and whether they are representative of the neighbourhoods the study seeks to represent.

External validity is defined as the extent to which the findings of a study are generalisable beyond the confines of the study itself to the source population. So, for example, findings from a study conducted in a school setting in the USA might be generalisable to other schools in the USA (the source population of the study). An assessment of external validity will consider how representative of the source population the study population is and whether or not there are any specific population, demographic or geographic features of the selected population that might limit or support generalisability. Also important are considerations of the setting, intervention and outcomes assessed. These factors will be considered in sections 2 and 3 of the checklist.

1.1 Is the source population or source area well described?

Was the source population or area described in sufficient detail? For example, country (developed or non-developed, type of healthcare system), setting (for example, primary school, community centre), location (urban, rural) and population demographics.

This question seeks to determine the study's source population or area (that is, to whom or what the study aims to represent). The source population is usually best identified by referring to the study's original research question.

It is important to consider those population demographic characteristics such as age, sex, sexual orientation, disability, ethnicity, religion, place of residence, occupation, education, socioeconomic position and social capital[2] that can help to assess the impact of interventions on health inequalities and may help guide recommendations for specific population subgroups.

1.2 Is the eligible population or area representative of the source population or area?

Was the recruitment of individuals, clusters or areas well defined (for example, advertisement, birth register, class list, area)?

Was the eligible population or area representative of the source or were important groups under-represented?

To determine if the eligible population or area (for example, smokers responding to a media advertisement, areas of high density housing in a particular catchment area) are representative of the source population (for example, smokers or areas of high density housing), consider the means by which the eligible population was defined or identified and the implicit or explicit inclusion and exclusion criteria used. Were important groups likely to have been missed or under-represented? For example, were recruitment strategies geared toward more affluent or motivated groups? (For example, recruitment from more affluent areas or local fitness centres.) Were significant numbers of potentially eligible participants likely to have been inadvertently excluded? (For example, through referral to practitioners not involved in the research study.)

[2] Demographic criteria as outlined by the PROGRESS-Plus categorisation (Kavanagh et al. 2008).

1.3 Do the selected participants or areas represent the eligible population or area?

Was the method of selection of participants from the eligible population well described?
What percentage of selected individuals or clusters agreed to participate? Were there any sources of bias?
Were the inclusion or exclusion criteria explicit and appropriate?

Consider whether the method of selection of participants or areas from the eligible population or area was well described (for example, consecutive cases or random sampling). Were any significant sources of biases likely to have been introduced? Consider what proportion of selected individuals or clusters agreed to participate. Was there a bias toward more healthier or motivated individuals or wealthier areas?

Also consider whether the inclusion and exclusion criteria were well described and whether they were appropriate given the study objectives and the source population. Strict eligibility criteria can limit the external validity of intervention studies if the selected participants are not representative of the eligible population. This has been well-documented for RCTs where recruited participants have been found to differ from those who are eligible but not recruited, in terms of age, sex, race, severity of disease, educational status, social class and place of residence (Rothwell 2005).

Finally, consider whether sufficient detail of the demographic (for example, age, education, socioeconomic status, employment) or personal health-related (for example, smoking, physical activity levels) characteristics of the selected participants were presented. Are selected participants representative of the eligible population?

Section 2: method of allocation to intervention (or comparison)

This section aims to assess the likelihood of selection bias and confounding being introduced into a study.

Selection bias exists when there are systematic differences between the participants in the different intervention groups. As a result, the differences in the outcome observed may be explained by pre-existing differences between the groups, rather than because of the intervention itself. For example, if the people in 1 group are generally in poorer health compared with the second group, then they are more likely to have a worse outcome, regardless of the effect of the intervention. The intervention groups should be similar at the start of the study so that the only difference between the groups should be the intervention received.

2.1 Allocation to intervention or comparison. How was confounding minimised?

Was allocation to exposure and comparison randomised? Was it truly random ++ or pseudo-randomised + (for example, consecutive admissions)?

If not randomised, was significant confounding likely (−) or not (+)?

If a crossover, was order of intervention randomised?

Consider the method by which individuals were allocated to either intervention or control conditions. Random allocation of individuals (as in RCTs) to receive 1 or other of the interventions under investigation, is considered the most reliable means of minimising the risk of selection bias and confounding.

If an appropriate method of randomisation has been used, each participant should have an equal chance of ending up in each of the intervention groups. Examples of random allocation sequences include random numbers generated by computer, tables of random numbers and drawing of lots or envelopes. However, if the description of randomisation is poor, or the process used is not truly random (for example, if the allocation sequence is predictable, such as date of birth or alternating between 1 group and another) or can otherwise be seen as flawed, this component should be given a lower quality rating.

2.2 Were the interventions (and comparisons) well-described and appropriate?

Were interventions and comparisons described in sufficient detail (that is, enough for study to be replicated)?

Were comparisons appropriate (for example, usual practice rather than no treatment)?

2.3 Was the allocation concealed?

Could the person(s) determining the allocation of participants or clusters to intervention or comparison groups have influenced the allocation?

Adequate allocation concealment (++) would include centralised allocation or computerised allocation systems.

If investigators are aware of the allocation group for the next individual to be enrolled in the study, there is potential for people to be enrolled in an order that results in imbalances in important characteristics. For example, a practitioner might feel that people with mild rather than severe mental health problems would be more likely to do better on a new, behavioural intervention and be tempted to only enrol such individuals when they know they will be allocated to that group. This would result in

the intervention group being, on average, less severe at baseline than control group. Concealment of treatment group may not always be feasible but concealment of allocation up until the point of enrolment in the study should always be possible.

Information should be presented in the paper that provides some assurance that allocations were not known until at least the point of allocation. Centralised allocation, computerised allocation systems and the use of coded identical containers would all be regarded as adequate methods of concealment. Sealed envelopes can be considered as adequate concealment if the envelopes are serially numbered, sealed and opaque, and allocation is performed by a third party. Poor methods of allocation concealment include alternation, or the use of case record numbers, date of birth or day of the week.

If the method of allocation concealment used is regarded as poor, or relatively easy to subvert, the study should be given a lower quality rating. If a study does not report any concealment approach, this should be scored as 'not reported'.

2.4 Were participants and investigators blind to exposure and comparison?

Were participants AND investigators – those delivering or assessing the intervention kept blind to intervention allocation? (Triple or double-blinding score ++).

If lack of blinding is likely to cause important bias, score −.

Blinding refers to the process of withholding information about treatment allocation or exposure status from those involved in the study who could potentially be influenced by this information. This can include participants, investigators, those administering care and those involved in data collection and analysis.

Unblinded individuals can bias the results of studies, either intentionally or unintentionally, through the use of other effective co-interventions, decisions about withdrawal, differential reporting of symptoms, or influencing concordance with treatment.

The terms 'single blind', 'double blind' and even 'triple blind' are sometimes used in studies. Unfortunately, they are not always used consistently. Commonly, when a study is described as 'single blind', only the participants are blind to their group allocation. When both participants and investigators are blind to group allocation the study is often described as 'double blind'. It is preferable to record exactly who was blinded, if reported, to avoid misunderstanding.

It is important to note that blinding of participants and researchers is not always possible, and it is important to think about the likely size and direction of bias caused by failure to blind in making an assessment of this component.

2.5 Is the exposure to the intervention and comparison adequate?

Is reduced exposure to the intervention or control related to the intervention (for example, adverse effects leading to reduced compliance) or fidelity of implementation (for example, reduced adherence to protocol)?

Was lack of exposure sufficient to cause important bias?

2.6 Is contamination acceptably low?

Did any in the comparison group receive the intervention or vice versa?

If so, was it sufficient to cause important bias?

If a crossover trial, was there a sufficient wash-out period between interventions?

2.7 Were other interventions similar in both groups?

Did either group receive additional interventions or have services provided in a different manner?

Were the groups treated equally by researchers or other professionals?

Was this sufficient to cause important bias?

This question seeks to establish if there were any important differences between the intervention groups aside from the intervention received. If some patients received additional intervention (known as 'co-intervention'), this additional intervention is a potential confounding factor in the presence of which can make it difficult to attribute any observed effect to the intervention rather than to the other factors.

2.8 Were there other confounding factors?

Were there likely to be other confounding factors not considered or appropriately adjusted for?

Was this sufficient to cause important bias?

2.9 Were all participants accounted for at study conclusion?

Were those lost to follow-up (that is, dropped or lost pre-, during or post- intervention) acceptably low (that is, typically less than 20%)?

Did the proportion dropped differ by group? For example, were drop-outs related to the adverse effects of intervention?

Section 2 also aims to assess the likelihood of attrition bias being introduced into a study.

Attrition bias occurs when there are systematic differences between the comparison groups with respect to participants lost, or differences between participants lost to the study and those who remain. Attrition can occur at any point after participants have been allocated to their intervention groups. As such, it includes participants who are excluded post-allocation (and may indicate a violation of eligibility criteria), those who fail to complete the intervention and those who fail to complete outcome measurement (regardless of whether or not the intervention was completed).

It is a concern if the number of participants who were lost to follow-up (that is, dropped out) is high – typically >20%, although it is not unreasonable to expect a higher drop-out rate in studies conducted over a longer period of time.

Consideration should also be given to the reasons why participants dropped out. Participants who dropped out of a study may differ in some significant way from those who remained in the study. Drop-out rates and reasons for dropping out should be similar across all treatment groups. In good quality studies, the proportion of participants lost after allocation is reported and the possibility of attrition bias considered in the analysis.

2.10 Did the setting reflect usual UK practice?

Did the setting in which the intervention or comparison was delivered differ significantly from usual practice in the UK? For example, did participants receive intervention (or comparison) condition in a hospital rather than a community-based setting?

2.11 Did the intervention or control comparison reflect usual UK practice?

Did the intervention or comparison differ significantly from usual practice in the UK? For example, did participants receive intervention (or comparison) delivered by specialists rather than GPs? Were participants monitored more closely?

Section 3: outcomes

Some of the items on this checklist may need to be filled in separately for each of the different outcomes reported by the study. For example, a study may report only 1 outcome of interest, measured by 1 tool, at 1 point in time, in which case each of the components (for example, reliability of outcome measure, relevance, withdrawals and drop-outs) can be assessed based on that 1 tool. However, if a study reports multiple outcomes of interest, scored by multiple tools (for example, self-report AND biochemically validated measures), at multiple points in time (for example, 6-month follow-up AND 1-year follow-up) individual components will need to be assessed for each outcome of interest.

It is important, therefore, that the reviewer has a clear idea of what the important outcomes are and over what timeframe, before appraising a study. The important outcomes for a piece of guidance will be identified through consultation with the NICE project team, the public health advisory committee and stakeholders.

3.1 Were the outcome measures reliable?

Were outcome measures subjective or objective (eg biochemically validated nicotine levels ++ versus self-reported smoking)?

How reliable were outcome measures (eg inter- or intra-rater reliability scores)?

Was there any indication that measures had been validated (eg validated against a gold standard measure or assessed for content validity)?

This question seeks to determine how reliable (that is, how consistently the method measures a particular outcome) and valid (that is, the method measures what it claims to measure) the outcome measures were. For example, a study assessing effectiveness of a smoking cessation intervention may report on a number of outcomes using a number of different tools, including self-reported smoking rates (a subjective outcome measure that is often unreliable) and biochemically validated smoking rates (an objective outcome measure that is likely to be more reliable).

If the outcome measures were subjective, it is also important to consider if the participant or researcher was blinded to the intervention or exposure (see question 2.4) as blinding may rescue the reliability of some subjective outcome measures.

3.2 Were the outcome measurements complete?

Were all or most study participants who met the defined study outcome definitions likely to have been identified?

3.3 Were all important outcomes assessed?

Were all important benefits and harms assessed?

Was it possible to determine the overall balance of benefits and harms of the intervention versus comparison?

3.4 Were outcomes relevant?

Where surrogate outcome measures were used, did they measure what they set out to measure? For example, a study to assess impact on physical activity assesses gym membership – a potentially objective outcome measure – but a reliable predictor of physical activity?

3.5 Were there similar follow-up times in exposure and comparison groups?

If groups are followed for different lengths of time, then more events are likely to occur in the group followed up for longer distorting the comparison.

Analyses can be adjusted to allow for differences in length of follow-up (for example, using person-years).

It is possible to overcome differences in the length of follow-up between groups in the analyses, for example, by adjusting the denominator to take the time into account (by using person-years).

3.6 Was follow-up time meaningful?

Was follow-up long enough to assess long-term benefits or harms?

Was it too long, for example, participants lost to follow-up?

The duration of post-intervention follow-up of participants should be of an adequate length to identify the outcome of interest.

Section 4: analyses

4.1 Were the exposure and comparison groups similar at baseline? If not, were these adjusted?

Were there any differences between groups in important confounders at baseline?

If so, were these adjusted for in the analyses (for example, multivariate analyses or stratification)?

Were there likely to be any residual differences of relevance?

Studies may report the distributions or important differences in potential confounding factors between intervention groups. However, formal tests comparing the groups are problematic – failure to detect a difference does not mean a difference does not exist, and multiple comparisons of factors may falsely detect some differences that are not real.

It is important to assess whether all likely confounders have been considered. Confounding factors may differ by outcome, so potential confounding factors for all of the outcomes that are of interest will need to be considered.

4.2 Intention to treat analysis?

Were all participants (including those that dropped out or did not fully complete the intervention course) analysed in the groups (that is, intervention or comparison) to which they were originally allocated?

4.3 Was the study sufficiently powered to detect an intervention effect (if one exists)?

A power of 0.8 (that is, it is likely to see an effect of a given size if one exists, 80% of the time) is the conventionally accepted standard.

Is a power calculation presented? If not, what is the expected effect size? Is the sample size adequate?

For cluster RCTs in particular, it is important to consider whether the cluster design has been appropriately taken into account in calculating required sample size for adequate power.

4.4 Were estimates of effect size given or calculable?

Were effect estimates (for example, relative risks, absolute risks) given or possible to calculate?

4.5 Were the analytical methods appropriate?

Were important differences in follow-up time, and likely confounders, adjusted for?

If a cluster design, were analyses of sample size (and power), and effect size performed on clusters (and not individuals)?

Were subgroup analyses pre-specified?

There are a large number of considerations in deciding whether analytical methods were appropriate. For example, it is important to review the appropriateness of any subgroup analyses (and whether pre-specified or exploratory) that are presented. Although subgroup analyses can often provide valuable information on which to base further research (that is, are often exploratory), it is important that findings of subgroup analyses are not over (or under) emphasised. Meaningful results from subgroup analyses are beset by the problems of multiplicity of testing (in which the risk of a false positive result increases with the number of tests performed) and low statistical power (that is, studies generally only enrol sufficient participants to ensure that testing the primary study hypothesis is adequately powered) (Assmann et al. 2000). In a good quality paper, subgroup analyses are restricted to pre-specified subgroups and are often confined to primary outcome measures. Data are analysed using formal statistical tests of interaction (that assess whether intervention effect differs between subgroups) rather than comparison of subgroup p values. A correction for multiple testing is performed where appropriate (for example, 'Bonferroni correction' where a stricter significance level is used to define statistical significance). The results are delineated carefully, and full details of how analyses were performed are provided (Assmann et al. 2000; Guillemin 2007).

The appropriateness of some analytical methods will also depend on the study design under investigation. For example, with cluster RCTs, because participants are randomised at the group level and are not independent 'units' (as is the case with RCTs based on individuals without clustering), and outcomes are often assessed at the individual level, statistical adjustments are necessary before pooled intervention and control group outcomes can be compared.

Likewise, it is also important to consider whether the degree of similarity or difference in clusters has been considered in analyses of cluster RCTs. Good quality cluster-RCTs will determine the intra-class correlation coefficient of their study (a statistical measure

of the interdependence in each cluster that is calculated by taking the ratio of the variance between groups compared with variance in groups).

Appendix 3: Inclusion and quality checklist

Inclusion/exclusion checklist

Population

Does the study population include:

	Yes	No		
Self-employed persons with no appointed line manager			Yes >	exclude
Sole traders			Yes >	exclude
Unemployed individuals			Yes >	exclude
No adults aged 50 or over			Yes >	Exclude

Publication details

Was the study:

	Yes	No		
Published before 2005			Yes >	exclude
Published in a language other than English			Yes >	exclude
A dissertation or thesis			Yes >	exclude

Setting

Is the study exclusively set in:

	Yes	No		
A workplace or amongst workers			No >	exclude
A country on the checklist (see below)			No >	exclude

Country Checklist
Australia, Austria, Belgium, Canada, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Israel, Italy, Japan, Korea, Latvia, Lithuania, Luxembourg, Malta, Netherlands, New Zealand, Norway, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden, Switzerland, United Kingdom, United States

Relevance

Does the study examine:

	Yes	No		
Age or specific needs of/impact on older workers (must have at least 51% as over 50)			No >	exclude
Organisational/community policies, initiatives and interventions that focus on health and wellbeing, supporting older workers, retirement planning and training, and/or counteracting/challenging ageism			No >	exclude
Issues relevant to the economic evaluation			No >	exclude

Does the study focus on:

	Yes	No		
Changes to employment/health and safety legislation			Yes >	exclude
Changes to organisational structure			Yes >	exclude
Activities for line managers that are NOT about training/mentoring to help managers manage older workers/counteract ageism/assist pre-retirement planning			Yes >	exclude
Whole workforce interventions that focus on physical activity, mental wellbeing, smoking cessation and long-term sickness absence/returning to work			Yes >	exclude
Interventions/support that employees can access on their own			Yes >	exclude
Statutory provision to employees			Yes >	exclude
Does the study focus on chronic illnesses (without considering prevention and specific effects on over 50s)			Yes >	exclude

Intervention

Does the study examine:

	Yes	No		
Employees over 50			No >	exclude
Entire workforces where at least 51% of employees are over 50			No >	exclude
How interventions targeted at 'older' workers under 50 may impact on them at over 50,?			No >	exclude
Interventions commissioned by organisations, but delivered by third party organisations			No >	exclude

Study information

For RQ1 and RQ2, does the study:

	Yes	No		
Employ qualitative methodology			No >	exclude
Examine the effect/impact on health and wellbeing			No >	exclude
Include an explicit measure of health and wellbeing			No >	exclude
Clearly explain its methodology			No >	exclude
Include control group and/or have more than one measure point			No >	exclude

For RQ3 does the study:

	Yes	No		
Include one of the following: document analysis, focus groups, interviews, observations, cross-sectional survey logy			No >	exclude
Clearly explain its methodology			No >	exclude
Make its evidence explicit			No >	exclude

Other information

Is the study:

	Yes	No
A review		
Experimental/observational		
Economic		
Qualitative		
A book		

NB can have more than one study type

Is the study set in:

	Yes	No
USA?		
UK?		
Europe?		
Other OECD?		
Multiple eligible locations?		

Which RQ is the paper relevant for?

	Yes	No
RQ1		
RQ2		
RQ3		

Is the study:

	Yes	No
An economic evaluation		
A systematic review/meta-analysis		
A book/book chapter		

For RQ1 and RQ 2, does the study have:

	Yes	No
A control group		
Two or more time measure points		

Does the sample:

	Yes	No
Include/focus on volunteers		

Appendix 4: Sample search strategies

MEDLINE 1996 to July 2014 (via OVID)

Search strategy 5 August 2014

Set	Searches	Results
1	(over adj2 "50").ti,ab.	9908
2	(over adj2 "55").ti,ab.	1277
3	(over adj2 "60").ti,ab.	7445
4	(over adj2 "65").ti,ab.	4672
5	((age* or old* or elder* or grey or silver or pensioner or senior) adj (worker* or employee* or people* or person* or woman or women or man or men or colleague* or earner* or operative* or volunteer* or population* or workforce or staff* or labourer* or laborer* or executive* or manager* or administrator* or personnel)).ti,ab.	190687
6	"third age*".ti,ab.	229
7	"baby boomer*".ti,ab.	662
8	(later adj2 life adj4 (worker* or employee* or people* or person* or woman or women or man or men or colleague* or earner* or operative* or volunteer* or population* or workforce or staff* or labourer* or laborer* or executive* or manager* or administrator* or personnel)).ti,ab.	308
9	(aged/ or middle aged/) and (worker* or employee* or people* or person* or woman or women or man or men or colleague* or earner* or operative* or volunteer* or population* or workforce or staff* or labourer* or laborer* or executive* or manager* or administrator* or personnel).ti,ab.	787984
10	(third adj2 (career* or job*)).ti,ab.	25
11	((age* or old* or elder* or grey or silver or pensioner or senior) adj2 (nurse* or physician* or doctor* or therapist* or paramedic* or surgeon* or dentist* or midwife or midwives or pharmacist* or lawyer* or teacher* or professor* or academic* or firefighter* or ambulance* or police* or miner* or driver* or trucker*)).ti,ab.	7416
12	(middle adj age* adj (worker* or employee* or people* or person* or woman or women or man or men or colleague* or earner* or operative* or volunteer* or population* or workforce or staff* or labourer* or laborer* or executive* or	5905

	manager* or administrator* or personnel)).ti,ab.	
13	(exp occupational groups/ or exp administrative personnel/ or exp clergy/ or exp doulas/ or exp ethicists/ or exp faculty/ or exp emergency responders/ or exp foreign professional personnel/ or exp health personnel/ or exp allied health personnel/ or exp anatomists/ or exp caregivers/ or exp "coroners and medical examiners"/ or exp dental staff/ or exp dentists/ or exp faculty, dental/ or exp faculty, medical/ or exp faculty, nursing/ or exp health educators/ or exp health facility administrators/ or exp infection control practitioners/ or exp medical chaperones/ or exp medical laboratory personnel/ or exp medical staff/ or exp nurses/ or exp nurse administrators/ or exp nurse anesthetists/ or exp nurse clinicians/ or exp nurse midwives/ or exp nurse practitioners/ or exp nurses, community health/ or exp nurses, international/ or exp nurses, male/ or exp nurses, public health/ or exp nursing staff/ or exp personnel, hospital/ or exp pharmacists/ or exp physician executives/ or exp physicians/ or exp veterinarians/ or exp inventors/ or exp laboratory personnel/ or exp lawyers/ or exp librarians/ or exp military personnel/ or exp "missions and missionaries"/ or exp police/ or exp research personnel/) and (age* or old* or elder* or grey or silver or pensioner or senior).ti,ab.	34746
14	exp Workplace/ or exp Employment/ or exp Work/ or exp Industry/	193247
15	((job* or employ* or work*) adj (place* or site* or setting* or location* or organisation* or organization*)).ti,ab.	4719
16	(workplace* or business* or shop* or factory or factories or company or companies or office* or industry or industries).ti,ab.	149591
17	exp Retirement/	3648
18	(retirement or retired or unretirement or redeployment).ti,ab.	7176
19	((retire* or pre-retire* or unretire*) adj2 (revers* or plan* or decision* or delay* or adjust* or late* or post*)).ti,ab.	588
20	((work or employment or flex* or retire*) adj2 transition).ti,ab.	244
21	((flex* or part-time or "part time") adj4 (career* or employ* or work* or time* or job* or hour* or intervention*)).ti,ab.	5303
22	((third or 3rd or encore or bridge) adj (work or career* or job* or employ*)).ti,ab.	76
23	"fourth pillar".ti,ab.	6
24	((regulat* or adapt* or adjust* or change* or modif* or redesign* or re-design*) adj2 (premise* or building* or work* or equipment or office* or shop* or industry or industries or factory or factories or company or companies or practice* or hour* or responsib* or environment* or job*)).ti,ab.	28260
25	(reasonable adj1 adjustment*).ti,ab.	33
26	(job* adj2 design).ti,ab.	119
27	((employ* or work* or job*) adj3 (training or mentor*)).ti,ab.	4947
28	((employ* or work* or job*) adj2 (pattern* or shift* or rota* or roster*)).ti,ab.	4828
29	((welfare or pension* or benefit* or tax* or work or employment) adj4 (barrier* or facilitat* or incentive* or disincentive* or penalt*)).ti,ab.	3641
30	Ageism/ or (ageism or (age adj2 discriminat*)).ti,ab.	682
31	((job* or work* or employ*) adj2 (shar* or return*)).ti,ab.	5950

32	(engage* and (civi* or job* or work* or employ* or staff* or worker*or workforce*)).ti,ab.	13168
33	(performance adj2 manage*).ti,ab.	645
34	(recruit* adj4 (civi* or job* or work* or employ* or staff* or worker*or workforce*)).ti,ab.	2503
35	exp "Personnel Staffing and Scheduling"/ and (age* or old* or elder* or grey or silver or pensioner or senior).ti,ab.	970
36	exp Accidents, Occupational/ and (age* or old* or elder* or grey or silver or pensioner or senior).ti,ab.	1531
37	exp Occupational Diseases/ and (age* or old* or elder* or grey or silver or pensioner or senior).ti,ab.	8842
38	((retention or retain) adj4 (worker* or employee* or people* or person* or woman or women or man or men or colleague* or earner* or operative* or volunteer* or population* or workforce or staff*)).ti,ab.	2069
39	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	890363
40	14 or 15 or 16	311640
41	17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38	86767
42	39 and 40 and 41	7574
43	limit 42 to (english language and humans and yr="2000 -Current")	5781

Notes:

Set 11 is a free-text search for a number of key professions, including health service personnel, which might not be picked up by using the generic words such as worker or staff

Set 39 represents older workers.

Set 40 represents the workplace

Set 41 covers workplace interventions

Set 42 combines all these three sets

and set 43 limits the results to English language, Humans and 2000 to current.

So set 43 is the results to be downloaded to EndNote and sifted there.

Index to theses

4 August 2014

Searches in All fields and limited to 2000-2014.

Term	Results	Saved to EndNote
Older workers	38	16
Retirement work	126	1
Retirement planning	3	1
Older employees	5	1

Items not retained because on completely different topics or duplicates. [Also covered by the Proquest Dissertations and Theses service.]

Medline via OVID 1996- July Week 4 2014

5 August 2014

Agreed strategy. 5781 downloaded to EndNote

Medline in Process via OVID

5 August 2014

Agreed strategy, but removing final set limits to English, humans and date and adapting MESH headings to keywords.

51 items retrieved, of which one was too old and so 50 downloaded to EndNote.

Embase 1996 to 2014 Week 31 via OVID

5 August 2014

Used agreed strategy, altering MESH terms to EMTREE equivalents. Then restricted to journals not indexed in Medline. 817 downloaded to EndNote

Cochrane – native interface

5 August 2014

There are 101 results from 8586 records for your search on '(age* or old* or elder* or grey or silver or pensioner or senior) near (nurse* or physician* or doctor* or therapist* or paramedic* or surgeon* or dentist* or midwife or midwives or pharmacist* or lawyer* or teacher* or professor* or academic* or firefighter* or ambulance* or police* or miner* or driver* or trucker*worker* or employee* or people* or person* or woman or women or man or men or colleague* or earner* or operative* or volunteer* or population* or workforce or staff* or labourer* or laborer* or executive* or manager* or administrator* or personnel) in Title, Abstract, Keywords and workplace* or business* or shop* or factory or factories or company or companies or office* or industry or

industries in Title, Abstract, Keywords , Publication Year from 2000 to 2014 in Cochrane Reviews'

101 items downloaded to EndNote

PsycINFO (via EBSCO)

5 August 2014

Agreed strategy, adapting MESH headings to APA where appropriate. 1948 hits downloaded to EndNote.

Health Business Elite (via HDAS)

5 August 2014

Adapted the strategy, using keywords only. Also limited to books and periodicals, removing newspaper articles. 861 hits downloaded to EndNote.Seemed to be a lot of irrelevant items.

ENDNOTE LIBRARY: After deduplication, 9414 items. 5th August

HMIC (via HDAS)

6 August 2014

Adapted the strategy, using keywords only. 103 items downloaded.

ASSIA and Sociological Abstracts (via HDAS)

Adapted the strategy, using keywords only.

3598 hits. Downloaded 6th and 7th August.

EndNote: deduplicated – 518 removed, 11978 remain. 6th August

Academic Search Complete (via EBSCO)

7 August 2014

As PsycINFO, but removed newspaper articles and book reviews. 5,956 hits.

Business Search Premier (EBSCO)

7 August 2014

As PsycINFO, but limited to country reports and academic journals (ie excluded trade, magazines, newspapers.) 1,568

ECONLIT (EBSCO)

As PsycINFO. All forms of publication included. 217

IBSS (Proquest)

8 August 2014

Adapted version of Medline search strategy. 262 hits downloaded.

Dissertations and Theses (includes UK) (Proquest)

8 August 2014

Adapted version of Medline strategy, 1711 hits, but weeded first so only 525 downloaded.

ABI Inform (Proquest)

8 August 2014

Adapted version of Medline strategy. Limited to material in academic journals. 624 downloaded.

11 August 2014

Social Policy & Practice (via OVID)

Adapted version of Medline strategy. 1386 downloaded. (Difficulties with getting date field to display correctly and all material downloaded as format journal article – in the end Used the Social Work Abstracts import filter)

Web of Science (via Thomson Reuters)

11 August 2014

Adapted Medline strategy. 5194 items. A lot of clinical material so limited to WoS subject categories (PUBLIC ENVIRONMENTAL OCCUPATIONAL HEALTH OR MANAGEMENT OR BUSINESS OR ECONOMICS OR PSYCHOLOGY APPLIED OR INDUSTRIAL RELATIONS LABOR OR ENVIRONMENTAL SCIENCES OR ENGINEERING INDUSTRIAL OR ERGONOMICS OR OPERATIONS RESEARCH MANAGEMENT SCIENCE OR SOCIOLOGY OR SOCIAL SCIENCES INTERDISCIPLINARY OR PUBLIC ADMINISTRATION) Downloaded 2692 items.

Campbell Collaboration

Native interface

11 August 2014

Search Older = 99, retirement = 8. Non actually relevant so nothing downloaded.

EPPI-Centre

11 August 2014

Native interface

Scanned list for 2000-2014, nothing found.

SCOPUS

12 August 2014

Native interface, at Manchester University Library. Adapted search strategy to focus on key topics, as the combination of adjacency indicators which worked well in other DBs was not accepted. 1227 records downloaded.

XPertHR

13 August 2014

Older

240 hits. Mainly summaries of articles published elsewhere, case reports, news snippets.Saved 3 reports

AgeInfo

Centre for Ageing Research

14 April 2014

Used words for workers, employees, doctors etc.

NOT dementia.

819 hits, went through those for 2000-2014, 56 items downloaded.

Grey search

Mednar

19th August 2014

Advanced search. "Older Workers" limited to 2000-2014. Excluded PubMed and commercial databases, concentrated on most likely US government agencies. 333 hits, 2 retained.

BASE

19th August 2014

Older workers

2000-2014

Limited type of material to: books, reports, papers and theses -

1007 hits, 529 downloaded. After deduplication, 506.

NDAR

19th August 2014

Workplace, workers

14 hits, 5 downloaded

Grey Literature report (NYAM)

http://www.greylit.org/

3 September 2014 Older workers and 2005-2014 – 19 hits – 7 downloaded to EndNote

GoogleScholar

3 September 2014

"older workers" and Report in Title.

Limited to 2005-2014 - 13 records downloaded to EndNote.

Appendix 5 Evidence Tables

Härmä et al. 2006

Study details	Population and setting	Research aims/objectives; research questions/hypotheses Allocation of individuals to intervention/control	Outcomes and methods of analysis	Results	Notes by review team
Authors: Härmä, M., Tarja, H., Irja, K., Mikael, S., Jussi, V., Anne, B., Pertti, M. Year: 2006 Citation: Härmä, M., Tarja, H., Irja, K., Mikael, S., Jussi, V., Anne, B., Pertti, M. (2006). A controlled intervention study on the	Source population/s: Report the following Study was undertaken in Finland Setting : line maintenance unit of a large airline company Location: Urban Sample characteristics: Male technicians of the aircraft technical maintenance unit. 2001 273 subjects responded to a questionnaire (60% response rate) and 49 joined	Research aims/objectives To evaluate the effects of a very rapidly forward rotating shift system, including only a single morning, evening and night shift, on the sleep-wakefulness, well-being and social life of young and older shift workers. Research questions/hypotheses Not reported Method of allocation: Voluntary self-selection – not possible to randomise workers to the intervention and control group (noted as a study limitation)	Outcomes: Self-completed questionnaire looking at the frequency of insomnia complaints, and an insomnia index was calculated from the questions. Questions re: effects of shift system on general health and wellbeing with direct questions being asked in 2001 and 2003. In 2001, the field measurements were conducted during one full shift cycle (1_15 days). In 2003, the field measurements were taken during two full shift cycles (2_5 days). An actigraph and a Pocket PC (handheld computer for field data)	Report results for all relevant outcomes: The change of the shift system was associated with an improved perception of the effects of the shift system on sleep, health, well-being at work, and free-time activities The improvement of general health was greater for the younger group (group *time*age: df 4, 406, F =4.1, p <0.003), while the perceived improvements of sleep and vigilance (group * time * age: df 4, 404, F = 9.5,p <0.0001), general well-being at work (group *time *age: df 4, 413, F =10.0, p <0.001), social life (group *time *age: df 4, 416, F =6.4, p <0.0001), family life (group * time *age: df 4, 408, F =5.0, p <0.0006) and hobbies (group *time *age: df 4,	Limitations identified by author: Intervention studies in field conditions are difficult to conduct, and the behaviour of the subjects may change during long follow-up times. Another problem with the protocol was the different starting and ending times of the shifts. This was not a problem with the sleepiness ratings, but since PVT was measured only during the first and last two hours of the shift, the reaction times in the beginning of the night shift are not directly

Study details	Population and setting	Research aims/objectives; research questions/hypotheses Allocation of individuals to intervention/control	Outcomes and methods of analysis	Results	Notes by review team
effects if a very rapidly forward rotating shift system on sleep-wakefulness and well-being among young and elderly shift workers. *International Journal of Psychophysiology, 59*. 70-79. Country of study: Finland Aim of study: To evaluate the effects of a very rapidly forward rotating shift system, including only a single	(voluntarily) the field studies with a pocket PC and actigraphy recordings. November 2002, 40 voluntary subjects started with new shift system, and in May 2003 38 of them responded to the same questionnaire and 31 joined the field studies. 24 of the 38 subjects in the new shift system had responded the same questionnaire in 2001, and 9 of them had also joined the field studies. 116 subjects of the control group answered the same questionnaire in 2003. 9 subjects also joined the field	Intervention/s description: In the beginning of the study, all the workers had a continuous backward rotating three-shift system with the shift order of EEE- -MMM- -NNN- - (E = evening shift, M = morning shift, N = night shift, - = free day). The shift changing times were mostly at 07:00, 15:00 and 23:00. The new shift system was planned and selected together with the representatives of the employer, employees and the occupational health experts. The order of the shifts in the new schedule was MEN- -, the morning shift being from 06:00 to 16:00 (10 h), the evening	were given to each subject. The Pocket PC was used for the custom PVT-test, sleep and social life diaries, and for the subjective ratings of sleepiness. Subjects wore their actigraph night and day during the whole measurement period. Subjective rating of sleepiness (KSS, Karolinska Sleepiness Scale) and objective performances (with the Psychomotor Vigilance Task, PVT) were recorded. Daily rest and activity cycles were monitored with an actigraph recording unit (Actiwatch AW4, Cambridge Neurotechnology, version 3.24) worn on the wrist. Sleep diary: After each sleep period, the subjects	416, F =3.2, p <0.01) were more prominent among the older group. Age did not have any significant main effects on any of the parameters listed above. The change in shifts was significantly related with the decrease of insomnia symptoms (index) in all shifts. The change in shifts was related to the increase in sleep length after the night shift but not after the other shifts. Before the night shift, 10% of the younger workers in the intervention group took a nap in before and 16% after the intervention. Napping increased similarly in the control group of the younger workers (from 18% to 22%). For the older workers, the corresponding values were from 26% to 47% in the intervention group and from 25% to 35% in the control group. The effect of the new shift schedule on the actual sleep length (SL) depended on age and shift (group *time * shift*age: df	comparable between the two shift systems. It was not possible, however, to randomise workers to the intervention and control groups which means that the volunteers for the new shift system could be self-selected. Several subjects measured at baseline could not be registered at the end of the study while some other subjects measured at the end did not join the study in the beginning. The statistical approach used means that partly different subjects were compared in 2001 and 2003. Some peculiar changes, especially in the rather small subgroups of the intervention and

Study details	Population and setting	Research aims/objectives; research questions/hypotheses Allocation of individuals to intervention/control	Outcomes and methods of analysis	Results	Notes by review team
morning, evening and night shift, on the sleep-wakefulness, well-being and social life of young and older shift workers. Study design: Controlled, on-site intervention study. Quality score[4] -	measurements (the same subjects as in 2001). The subjects were divided into younger (24–44 years) and older (45–61 years). Occupations included maintenance workers, inspectors and supervisors, and work experience varied between 14-32 years. Eligible population: Participants volunteered to take part in the study and questionnaire data on subjective stress, snoring and the amount of sleep apneas showed no differences at baseline between the age groups.	shift from 15:00 to 01:00 (10 h) and the night shift from 21:00 to 06:00 (9 h). The new schedule had thus fewer consecutive morning, evening and night shifts. In the old shift schedule, there were 56 h between the evening and morning shifts and also between the morning and night shifts. Similarly, free-time after the three night shifts in the old shift system was 80 h. In the new shift system, there were 72 h between the single nightshift and the following morning shift.	rated their estimated minutes to falling asleep (sleep latency), the number of awakenings, the estimated awaking time between waking and falling asleep, the feeling of awakening too early from sleep and the feeling of insufficient sleep (five point scale, fully enough–clearly not enough). Subjective sleepiness was rated using the Karolinska Sleepiness Scale (KSS). Objective performances were measured using the Psychomotor Vigilance Task (PVT) Follow-up periods: 2001-2003 Method of analysis: The associations between the group (intervention vs. control), time (before or after the intervention), age (44 years or younger, 45 years and older) and different shifts were	16/784, F =1.84, p <0.0231). The change in shifts affected also sleep efficiency (SE) and sleep fragmentation (FI, fragmentation index), measured by the actigraph recordings. SE depended on shift (df 4/779, F =2.51, p <0.0405), the quality of sleep being lower during the day than during the night. The interaction of group, time, shift and age was significant for both variables (se: df 32/791, F =1.58, p <0.0223; FI: df 32/ 797, F =1.50, p <0.0389). The restorative effect of sleep improved especially before the night shift among the younger age group due to the intervention (group * time * shift: df 13/840, F =2.03, p < 0.0160, group * time * shift * age: df 19/787, F =1.94, p <0.0159). The change in shifts decreased severe sleepiness at free time in both groups after the night shift, (group*time: df	control groups of the actigraph and PVT data, could thus be due to differences in the groups compared. Limitations identified by review team: Generalisability of findings for other sectors Only men were used, and age groups were quite wide – may have been inter-group differences not reported. Source of funding: The study was supported by an EU-grant QLRT2000-00038 of the ‘‘Respect’’-program.

Study details	Population and setting	Research aims/objectives; research questions/hypotheses Allocation of individuals to intervention/control	Outcomes and methods of analysis	Results	Notes by review team
			tested with a linear mixed model for repeated measurements. All the statistical analyses were carried out using the Statistical Analysis System	3/385, F =4.9, p <0.03)	

Hughes et al. 2011

Study details	Population and setting	Research aims/objectives; research questions/hypotheses Allocation of individuals to intervention/control	Outcomes[2] and methods of analysis	Results[3]	Notes by review team
Authors: Hughes, S. L., Seymour, R. B., Campbell, R. T., Shaw, J. W., Fabiyi, C., & Sokas, R. **Year:** 2011 **Citation:** Comparison of two health-promotion programmes for older workers. American journal of public health, 101(5), 883. **Country of study:** USA **Aim of study:** To examine the	**Source population/s**: **Country of study:** USA **Setting:** Older (aged 40+) workers at a University **Location**: Urban university (Chicago) **Sample characteristics, including population demographics**: 423 Participants; 150 in COACH intervention group, 135 in RealAge intervention group, 138 in control group. Mean age 51 (range 40-68),	**Research aims/objectives** **Research questions/hypotheses:** To examine the effects of 2 worksite health-promotion interventions (compared with a health-education control) on older workers' healthy behaviours and health outcomes. **Method of allocation**: Randomization sequences determined with custom software designed to achieve balanced allocation of cases to conditions stratified by education and race/ethnicity. **Intervention/s description**: COACH: Contact with a single coach trained in principles of behaviour change and motivational interviewing. Named the COACH programme because it was thought to have greater appeal to a working-age population. The coach was involved in health-risk assessment and	**Outcomes**: Dietary behaviours: changes in per cent energy from fat and in fruit and vegetable intake from baseline to 6 months and 12 months. Physical activity (PA): Changes in vigorous activity, moderate activity, and level of exercise participation from baseline to 6 months and 12 months. Stress: Change from baseline to 6 months and 12 months in 4 measures of stress using scales developed to assess health-related stress, and coping behaviours. Smoking. Smoking cessation was defined as a minimum of 6 months of total abstinence from tobacco use at 6 months and 12 months after baseline, among participants who were current smokers at baseline or 6 months. Body mass index, waist circumference, and	**Report results for all relevant outcomes**: Diet: significant 3-way interaction for % of energy from fat and fruit and vegetable consumption, such that those in the COACH group who were in a SOC coded 1 had greater changes over time. COACH participants reported a borderline significant reduction in % energy from fat at 6 months (P =.063) and a significant reduction at 12 months (P =.027). Participants also reported eating significantly more fruits and vegetables than control group participants at 6 months (P =.026) and 12 months (P <.001). No significant differences were seen on either variable for RealAge participants at either time point. Consistent with the findings from random-effects analyses, COACH participants experienced the largest decrease in % energy from fat, RealAge participants experienced more modest (insignificant) declines, and control participants stayed the same over the 12-month period. Physical activity (PA): COACH participants reported significantly	**Limitations identified by author**: COACH was proactive, RealAge was reactive, relying on the consumer to initiate follow-up; these inherent differences in approach may explain study findings. The interventions were tested with staff at an inner-city university who may have had higher levels of education than do workers in other industries; the generalizability of the findings to workers in other settings requires further testing. Cost/benefit analysis was not undertaken. Both interventions were light interventions in terms of intensity; further study could include more intense interventions. Need to consider whether a dose-response relationship exists between comprehensiveness of services offered, and whether certain programme components, such as incentives, are more effective and more critical than others. Other questions

Study details	Population and setting	Research aims/objectives; research questions/hypotheses Allocation of individuals to intervention/control	Outcomes[2] and methods of analysis	Results[3]	Notes by review team
effects of 2 worksite health-promotion interventions (compared with a health-education control) on older workers' healthy behaviours and health outcomes. **Study design:** Randomised control trial **Quality score: +**	82% female, 62% ethnic minority. **Eligible population**: **Describe how individuals, groups or clusters were recruited:** Broad array of recruitment strategies, including conducting recruitment events at locations frequently visited by the target population, mass e-mails, and campus listservs that sent biweekly messages to staff **Excluded population/s:** Not made explicit	discussion about potential behavioural change. The coach reviewed health-related goals and negotiated an action plan to meet them; this could be altered over time. For a week after the assessment, the coach asked participants about accessing resources needed to implement the plan, and revised the plans for those who were having difficulty doing so. The coach also contacted participants biweekly during months 1-6, and monthly during months 7-12 to keep the goals up to date. RealAge: Participants took the RealAge test online, which reviewed numerous health factors. After participants completed the test, the website generated individual risk profiles and indicated areas that could be improved. Participants used the website to select behaviours and create plans to meet behavioural goals. The site tracked each time a participant used it, and	Weight: For all 3 measures change from baseline at 6 months and 12 months assessed. **Method of analysis**: Mixed-effects regression models involving 2 between subjects variables (group assignment and state of change (SOC)) and 1 within-subject variable (time). Group assignment represented by 2 indicator variables for COACH and RealAge, with the control as the reference group. Time was represented with indicator variables for 6 months and 12 months relative to baseline. SOC indicated pre-contemplation, contemplation, or preparation (coded 1) versus action or maintenance (coded 0). Participants varied with regard to the health behaviour; a respondent might be coded differently on SOC for different outcomes. The 2-way interactions between group and time allowed determination of how patterns of change over time varied by group. 3-way interactions allowed assessment of whether differences over time between groups varied by SOC level. 1-tailed tests of significance used because 7 prior	more minutes of moderate PA than did controls at 6 months (P=.05) and 12 months (P=.013). No significant differences were seen for COACH participants on rapid assessment of physical activity (RAPA) scores or on minutes of vigorous PA at 6 or 12 months, and no significant differences were seen on any of the PA variables at either time point for RealAge participants. Stress: No significant differences. Smoking. At 12 months, 2 of 16 COACH smokers, 4 of 16 RealAge smokers, and 3 of 21 control group smokers achieved maintenance of smoking cessation. There were no significant differences in rates of smoking cessation for smokers in COACH or RealAge relative to those in the control group with or without SOC for smoking at baseline as a covariate. BMI, waist circumference, and weight: When baseline SOC for diet behaviours was used as a covariate, RealAge participants experienced a significant decline in waist circumference at 6 months (P =.05) that was maintained at 12 months (P =.018). Neither COACH nor RealAge	concern the extent to which older workers who adopt healthy behaviours in the workplace are more likely to practice these behaviors in retirement, as well as the attendant impact on retiree health expenditures and functional independence. COACH group participants were almost twice as likely to use their intervention as RealAge, and COACH participants experienced 3 times as many positive outcomes at 6 months and12months for diet and PA, when both interventions were compared with a light health-education control. By contrast, no differences were seen in RealAge with respect to PA, but a significant decrease was seen in waist circumference. **Limitations identified by review team:** Effect size was not reported, confounding factors are not identified, and the generalizability also suffers from the imbalanced gender and ethnic minority

Study details	Population and setting	Research aims/objectives; research questions/hypotheses Allocation of individuals to intervention/control	Outcomes[2] and methods of analysis	Results[3]	Notes by review team
		forwarded this information to the study team at regular intervals.	studies of similar interventions had shown positive effects. Accordingly, a directional hypothesis at the conventional .05 significance level was chosen to detect a null or positive effect	participants experienced significant decreases in BMI or weight at 6 or 12 months compared with control participants.	breakdowns of the study. **Source of funding**: Centre for Disease Control and Prevention and the National Institute On Ageing.

Rutanen et al. 2014

Study details	Population and setting	Research aims/objectives; research questions/hypotheses Allocation of individuals to intervention/control	Outcomes and methods of analysis	Results	Notes by review team
Authors: Reetta Rutanen, Clas-Håkan Nygårda, Jaana Moilanena, Tomi Mikkolac, Jani Raitanena,b, Eija Tomasd and Riitta Luotob Year: 2014 Citation: Work 47 (2014) 281–286 Effect of physical exercise on work ability and daily strain in symptomatic menopausal women: A randomized controlled trial	Source population/s: Finland Setting At home and at the research institute. Does not say if rural or urban. Sample characteristics, N = 176 Intervention N = 56 Control N = 53 Females ages 40 – 62 All employed ft or pt All completed primary education Eligible population: Females aged 40 – 62, were symptomatic (daily hot flushes, no current use of postmenopausal hormone therapy, or	Research aims/objectives To investigate the effects of an aerobic physical exercise intervention on perceived work ability, daily strain and mental resources among women reporting menopausal symptoms. Method of allocation: Does not report how groups were allocated Intervention/s description: The intervention included 6 months aerobic exercise training 4 times a week, 50 minutes per session, with a progressive increase in	Outcomes: Include details of all relevant outcome measures and whether measures are objective or subjective otherwise validated The Work Ability Index (WAI). This is a sum of seven items. 1) Work ability in relation to lifetime best (0-10), 2) Work ability in relation to physical and mental work demands (1-5), 3) number of diagnosed diseases, 4) work impairment from diseases (1-6), 5) self reported sick leave, 6) individual prognosis of work ability after 2 years, (1, 4, 7), and 7)	Report results for all relevant outcomes: The report shows physical exercise is associated with self reported improvement in mental resources and a decrease in daily physical work strain. No statically significant changes in work ability were shown. WAI and other work related items were not different at baseline. After 6 months daily perceived work strain was lower among those in the intervention group (2.2 SD = 0.6 vs. 2.5 SD = 0.6, p = 0.025) WAI was higher for intervention group compared to control but was not statistically significant (39.4, SD = 5.4 vs. 38.4, SD = 6.2). Changes in work ability were	Limitations identified by author: Work ability index developed to monitor changes over one year. 6 months may not be long enough to detect changes. Respondents reported relatively high levels of work ability at baseline which could have limited further increases in this. Absence of blinding. Conducted in s small community where subjects

Study details	Population and setting	Research aims/objectives; research questions/hypotheses Allocation of individuals to intervention/control	Outcomes and methods of analysis	Results	Notes by review team
Country of study: Finland Aim of study: To assess the effectiveness of physical exercise on the work ability and daily strain of menopausal women. Study design: RCT Quality score (++, + or -) +	hormonal therapy use withdrawal (wash-out period 3 months), low physical activity (physical exercise less than twice a week) and 6–36 months since last menstruation. Recruited via advertisements in local newspaper. Does not state if eligible population is considered representative of the source population. Selected population: Females aged 40 – 62, Worked between 7 and 60 hours per week A third had a university degree. All non-workers excluded and those undertaking at least 2 sessions of exercise per week.	intensity. At least two sessions were supposed to involve walking or Nordic walking and the other two could be jogging, cycling, swimming, skiing, aerobics or other gymnastic exercise. An aerobics or step aerobics session was also offered at the treatment centre Rating of perceived exertion was taken to check for intensity. The target was 13 – 16 on a 6 – 20 point scale. Authors do not report on this. Control/comparison's description: Control group asked to keep physical activity habits normal. Members of intervention group attended fortnightly	mental resources. Questionnaires on daily physical and mental work strain were filled out in the evenings at baseline and end (0 = very little, 5 = very much). Cardio respiratory fitness assess by UKK walking test. Measures heart rate, walking time and BMI and estimates maximal oxygen consumption. Follow-up periods: 6 months Method of analysis: Differences at baseline assessed by t-tests for continuous variables normally distributed, Mann-Whitney when non-normally distributed and Chi-square for categorical variables.	greater in the intervention group. Between group differences in change of WAI were not significant (Adjusted b = 0.97, 95% CI = 0.33 – 2.26) p value not reported. When adjusted for baseline factors (age and work demand) significant differences were found in mental resources (coefficient 0.58, 95% CI = 0.17-1.00, p < 0.01) and physical work strain (coefficient -0.26, 95% CI = -0.45 - -0.07, p < 0.01) Note any results that detail impact on health inequalities. Total sample: Baseline Intervention N = 63 Control N = 60 End-point Intervention N = 46-57 Control N = 51 - 58 Intervention group(s): Baseline Follow-up (all time points)	likely to meet. Controls were aware of study and so could have engaged in physical exercise. Limitations identified by review team: Source of funding: Not reported

Study details	Population and setting	Research aims/objectives; research questions/hypotheses Allocation of individuals to intervention/control	Outcomes and methods of analysis	Results	Notes by review team
	What % of selected individuals/clusters agreed to participate? 69.9% of total sample eligible at baseline Excluded population/s: (as above)	health/wellbeing talks with research team. Sample sizes at baseline Intervention N = 63 Control N = 60 Baseline comparisons: Report any baseline differences between groups in important confounders. At baseline higher physical work demand in the intervention group was the only difference in groups. Study sufficiently powered: Does not report	Linear, ordinal and multinomial regression tested for effects using SPSS.	End-point Control group(s) Baseline Follow-up (all time points) End-point Attrition details: Indicate the number lost to follow-up and whether the proportion lost to follow-up differed by group (ie invention vs control) 14 participants in intervention group and 8 women in the control group dropped out (ie <20% of total in each group)	

Strijk et al. (2012)

Study details	Population and setting	Allocation of individuals to intervention/control	Outcomes and methods of analysis	Results	Notes by review team
Authors: Strijk, J. E., Proper, K. I., Van Der Beek, A. J., & Van Mechelen, W. **Year:** 2012 **Citation:** A worksite vitality intervention to improve older workers' lifestyle and vitality-related outcomes: results of a randomised controlled trial. *Journal of epidemiology and community health, jech*-2011. **Country of study:** Netherlands **Aim of study:** To evaluate the effectiveness of a worksite vitality intervention on	**Source population/s**: **Country of study:** Netherlands **Setting:** 2 academic hospitals, relevant participants were those aged 45 and over **Location:** Urban; Amsterdam and Leiden **Sample characteristics, including population demographics** Intervention Group; 74.7% female, mean age 52.5 (SD=4.8) Control Group: 76.3% female, mean age 52.3 (SD=4.9) **Eligible population**: **Describe how individuals, groups or clusters were**	**Research aims/objectives:** To evaluate the effectiveness of a worksite vitality intervention on vigorous physical activity (VPA), fruit intake, aerobic capacity, mental health and need for recovery after work among older hospital workers (i.e., 45 years and older). **Research questions/hypotheses:** As above **Method of allocation**: The workers who consented to participate were, after baseline measurements, individually randomised to the intervention or control group using Random Allocation Software. **Intervention/s description**: 6-month intervention consisting of (1) a Vitality Exercise Program (VEP) with (2) provision of free fruit and combined with (3) three visits to a Personal Vitality Coach (PVC). The VEP consisted of a weekly 45 min: (1) yoga session, (2) workout session and (3) unsupervised aerobic exercise session.	**Outcomes**: Physical activity (PA): measured subjectively through a questionnaire and objectively using accelerometers. Outcome measures were total minutes per week of: (1) sports activities, (2) VPA, and (3) total moderate-to vigorous physical activities (MVPA). Weekly fruit intake: self reported via questionnaire Aerobic capacity (VO2max): estimated using the UKK 2 km walk test. Workers walked briskly for 2 km, with heart rate and performance time monitored, from which VO2 max was estimated. Mental health: Questionnaire that refers to the past 4 weeks: "Did you feel… (1) nervous, (2) down in dumps, (3) peaceful, (4) sad and (5) happy". Need for recovery (NFR) was assessed with a questionnaire consisting of 11 statements (yes/no) concerning the recovery period after a day's	**Report results for all relevant outcomes**: The study showed that the intervention group significantly increased their weekly sports activities and fruit intake, when compared to control group. The intervention also favourably affected the NFR after a day of work. No effects were observed for VPA, aerobic capacity and mental health. Complete cases analyses revealed effectiveness on sports activities (b: 40.4 min/week, 95% CI 13.0 to 67.7). The control group workers increased their sports activities with 35.1 min/week, but when compared to the intervention group, this increase was statistically higher (75.3 min/week). As for the subjectively measured VPA, in total 134 workers (intervention n + 63, control n = 73) completed these measures. It appeared that both the intervention and control group increased their VPA from baseline to 6 months later (+159.5 vs +110.3 min/week, respectively), with no significant differences between groups (b = 48.5 min/week, 95% CI = -81.0 to 178.1). Also based on the accelerometer data, there were no significant differences between groups (b = 8.5 min/week, 95% CI -0.34 to 17.3). No effects were found on total weekly MVPA (SQUASH: b = -1.4	**Limitations identified by author**: Relatively healthy older workers population, mainly consisting of female workers, making it more difficult to generalise the study results. Failed to ensure vigorous intensity physical activity compliance during the guided workout session required to improve aerobic capacity. Also, sensitivity analyses showed similar but smaller estimates of effects, when compared to the complete cases analyses, indicating that the risk of bias is minimal. This is a commonly seen consequent of imputation but could indicate a potentially biased estimation obtained from complete cases. **Limitations identified by review team:** Inclusion/exclusion criteria not listed, making replication potentially difficult **Source of funding**: financially supported by the 'Foundation Institute GAK"

Study details	Population and setting	Allocation of individuals to intervention/control	Outcomes and methods of analysis	Results	Notes by review team
vigorous physical activity (VPA), fruit intake, aerobic capacity, mental health and need for recovery after work among older hospital workers (i.e., 45 years and **older).** **Study design:** Randomised control trial **Quality score: +**	**recruited** Hospital workers aged 45+ were invited to participate. A worker was considered eligible when working at least 16h a week, giving written informed consent and having no risk for developing adverse health effects. Consenting workers were allocated using Random Allocation Software. **Excluded population/s:** Inclusion criteria mentioned, but never made explicit		work. At baseline, data on potential confounders and effect modifiers were assessed by questionnaire including age, gender, education, chronic disease status, smoking, intervention location, type of work and marital status. **Method of analysis**: Differences were tested using independent t test for continuous variables and Pearson's tests for categorical and dichotomous variables. Differences in change over time between the intervention and control group were analysed using linear regression. For the sensitivity analyses, missing data were imputed using multiple imputations based on Multivariate Imputation by Chained Equations.	min/week, 95% CI -126.0 to 123.2; CSA: b = 13.8 min/week, 95% CI -25.9 to 53.5). Regarding fruit intake, the intervention group workers improved their fruit intake significantly more when compared to the control group (+5.7 vs +2.7 pieces/week), resulting in an intervention effect on increasing fruit intake (b = 2.7 pieces/week, 95% CI 0.63 to 4.7). As for the vitality related outcomes, no significant effects were found on aerobic capacity or mental health. As for NFR, the intervention group significantly decreased their NFR more when compared to the control group (-3.2 vs 0.6 points). Hence, the intervention was effective in decreasing workers' NFR (b = -3.5 points, 95% CI -6.4 to -0.54). A significant relationship was found between sports and high compliance to the guided yoga (b=49.6 min, 95% CI 13.9 to 85.2) and workout sessions (b=72.9 min/week, 95% CI 36.1 to 109.8) when compared to the control group. Also for fruit intake, effects were stronger in the high compliance group of both the yoga (b=3.8 pieces, 95% CI 1.1 to 6.4) and the workout sessions (b=4.0 pieces/week, 95% CI 1.1 to 6.4). Sensitivity analyses with imputed data for missing values showed similar findings when compared to the complete cases analyses. However, the effect sizes, derived from the analyses with imputed data, were consistently smaller when compared to the complete cases.	

Wagner et al. 2007

Study details	Population and setting	Research aims/objectives; research questions/hypotheses Allocation of individuals to intervention/control	Outcomes[2] and methods of analysis	Results[3]	Notes by review team
Authors: Wagner, S., Kascehl, R., Paulsen, S., Bleichner, F., Knickenberg, R.H. and Beutel, M.E. Year: 2007 Citation: Wagner, S., Kascehl, R., Paulsen, S., Bleichner, F., Knickenberg, R.H. and Beutel, M.E. (2008). Does a cognitive-training programme improve the performance of middle-aged employees undergoing in-	Source population/s: Report the following Country of study: Germany Setting : Psychosomatic Clinic Neustadt/Saale Location: Not reported Sample characteristics: 345 patients potentially eligible for the study enrolled in the study. Mean age was 53.7 years and mean length of education was 13.7 years. 78% were admitted with depressive disorder, mental disorders. 6% were diagnosed with somatic disorders and adjustment disorders and 5% suffered from an anxiety disorder.	Research aims/objectives To implement and evaluate a cognitive-training programme to improve the cognitive performance of patients with MCI (mild cognitive impairment) Research questions/hypotheses Does the memory performance of training participants improve between intake and discharge, compared with that of members of a control group? Does the self rated memory performance of training participants improve because of the cognitive-training programme? Do the work-related attitudes of the participants change as a result of the training programme?	Outcomes: Giessen Cognitive Screening of the Psychiatric University Clinic formed the basis of the test battery for assessing MCI. Appointment test for prospective memory, revised version of the Wechsler Memory Scale, subsets used: Logical Memory I and II. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD) neuropsychological test battery subsets: learning of word lists, recall. Recognition and construction ability. Achievement Measure System – used the verbal subset. Testbattery for Attentional Performance: used the subset 'Go NotGo'. Mini-Mental-State Test was used to exclude patients with dementia. Participants also completed a Memory Assessments Clinic questionnaire – 49 questions	Report results for all relevant outcomes: The performance of the intervention group on the Appointment test and the Logical Memory II test was significantly improved after training, whereas the performance of the control group was unchanged. ($F=4.95$, $p<.05$ for Logical memory II test and $F=15.06$, $p<0.001$ for Appointment test). There was no significant change in performance on the Logical Memory I test between intake and discharge for either group. No significant improvement in performance was found for either group on any of the other tests. Self-rated memory performance was also examined. In the ability section of the Memory Assessment Clinics questionnaire the self-rated performance of the intervention group improved after training $t=-3.99$, $p<.001$),	Limitations identified by author: Not been able to determine the mechanisms underlying the increased memory performance in the intervention group. Not achieved a long-term outcome, so are unable to comment on actual transfer to daily work and life performance. Limitations identified by review team: Is the battery of tests transferable to other languages? No indication of

Study details	Population and setting	Research aims/objectives; research questions/hypotheses Allocation of individuals to intervention/control	Outcomes[2] and methods of analysis	Results[3]	Notes by review team
patient psychosomatic treatment? *Disability and Rehabilitation, 30 (23),* 1786-1793 Country of study: Germany Aim of study: Study design: To implement and evaluate a cognitive-training programme to improve the cognitive performance of patients with MCI (mild cognitive impairment) Quality score[4] + 	Most frequent somatic disorder was musculoskeltal disorders (22%), cardiovascular system (21%), metabolism (17%) and tinnitus (16%0. 92 individuals (29%) met the criteria of MCI. Eligible population: Referrals were made by general practitioners and health insurance or annuity insurance companies. The study participants did not differ in socio-demographic characteristics of psychological or somatic diagnoses from those who declined to participate. Neither was there any difference between	Method of allocation: Participants who in the first session who showed cognitive impairment or reported memory deficits took part in the cognitive-training programme during the intervention phase but method of allocation to intervention/control group is not stated in the paper. Intervention/s description: The cognitive training programme tool place in interactive, closed groups of 4-8 participants over 7 sessions. An introductory session (60 minutes) a group therapist gave out essential information about memory processes which were illustrated by exercises. The subsequent 6 training sessions (90 minutes), were based on behaviour analysis, and two topics found to be	on a Likert type scale about memory of daily life situations. Beck Depression Inventory and State-Trait Anxiety Inventory. The work-related behaviour pattern surveyed the various aspects of work related experiences and behaviour patterns. Follow-up periods: A 1 year follow up examination is planned, although at discharge neuropsychological tests were repeated with the participants with cognitive impairments. Method of analysis: The raw scores of the Achievement Measures System were transformed into corrected t-values for education. Testbattery for Attentional Performance were generated as education-corrected and age-corrected t-values. MCI was defined as below average performance in at least two of the five areas of functioning as compared	and they also felt less concerned about everyday memory function than they had been before training (t=-2.83, p<.01). Control group also rated their memory at discharge as significantly better than at intake (t=-2.98, p<.01), but they were still very worried about their memory performance. Work-related attitudes: after training the intervention group reduced their pursuit for perfection and their level of exhaustion significantly (t=3.23, p<0.1 and t=-4.17, p<.001). The control group experienced no significant changes in their work-related attitudes between intake and discharge. According to responses on a subjective questionnaire at the end of the training 82% of the intervention group felt able to apply the problem-solving strategies acquired during training at their workplace, 74% felt able to analyse difficulties	work roles of the sample, and so will results be generalizable across industries? Source of funding: Not reported

Study details	Population and setting	Research aims/objectives; research questions/hypotheses Allocation of individuals to intervention/control	Outcomes[2] and methods of analysis	Results[3]	Notes by review team
	the intervention and control groups in this regard. Selected population: 92 participants fulfilled the criteria for MCI – 27 with a memory disorder, 31 suffered from additional impairments in other cognitive areas of functioning and 34 had cognitive impairments not involving memory. Excluded population/s: Those with dementia as diagnosed by the Mini-Mental-State Test.	important for this age group were discussed: prospective memory and structures processing of new information. From this analysis strategies are worked out that might improve performance in the future. On the basis of a detailed description of the problem (behavioural problem relevant to the cognitive problem), problem solving strategies are worked out in the group. Over the course of the training, options and difficulties in implementing the modifications were discussed and problem-solving strategies refined. Participants asked to develop ways to facilitate transfer into everyday lives, and expected to practise and consolidate skills acquired during their daily routine at the clinic and homework assignments.	with the mean of each age group. Impaired participants were divided into three subgroups: memory disorders, memory deficits and additional deficits in other areas of cognitive functioning; and those with impairments in several areas of cognitive functioning only. Data were analysed using SPSS with parametric and non parametric procedures (t-test and ANCOVA).	with cognitive demands themselves. Majority of the intervention group learned that strategies that helped them to remember appointments and to structure new information in a useful manner (71 and 70%). 68% felt better able to accept variations in job performance, as a consequence of training, and 70% of the intervention group rated their overall cognitive ability as improved. Total sample: Follow up period not mentioned in this paper, therefore total sample as reported in sample characteristics. No control or intervention sample size of characteristics given.	

Study details	Population and setting	Research aims/objectives; research questions/hypotheses Allocation of individuals to intervention/control	Outcomes[2] and methods of analysis	Results[3]	Notes by review team
		Control/comparison's description: (as above) Sample sizes at baseline: 92 meeting MCI criteria Baseline comparisons: Not reported Report any baseline differences between groups in important confounders. Study sufficiently powered: Not reported			

Wegner et al. 2011

Study details	Population and setting	Research aims/objectives; research questions/hypotheses Allocation of individuals to intervention/control	Outcomes and methods of analysis	Results	Notes by review team
Authors: Wegner, R., Berger, P., Poschadel, B., Manuwald, U and Baur, X. Year: 2011 Citation: Wegner, R., Berger, P., Poschadel, B., Manuwald, U and Baur, X. (2011). Burnout hazard in teachers results of a clinical-psychological intervention study. *Journal of Occupational Medicine and Toxicology, 6,* 37-42. Country of study: Germany Aim of study: To investigate whether an established	Source population/s: Report the following Country of study Germany Setting Psychotherapeutically oriented clinic Location Rural Germany Sample characteristics, including population demographics Group to be tested consisted of 200 teachers (civil servants) from Germany, aged between 27-64 years who voluntarily underwent inpatient treatment for emotional exhaustion between 2001-2007. 134 females and 66 males. 34 were high-school teachers (18	Research aims/objectives To investigate whether an established psychotherapeutically oriented inpatient treatment supplemented by a job-specific intervention shows long-acting success and whether gender gaps and differences between high school teachers and teachers of other levels exist. Research questions/hypotheses Not reported Method of allocation: NA Intervention/s description: Treatment programme averages about 7 weeks a pre-history based on depth psychology was conducted as well as a	Outcomes: Questionnaire of 60 questions of varying complexities covering demographic data, questions about working hours, work organisation, professional history the duration of the incapacity due to illness in the last quarter and the Maslach Burnout Inventory (MBI) in its German translation. MBI consists of 22 statements about feelings and attitudes that assess the three aspects of	Report results for all relevant outcomes Of the 200 inpatient treated teachers 63.5% had depressive disorders, 23.5% had neurotic disorders and 11.5% had personality disorders. The remaining 1.5% had a.o. somatoform disorders. Gender differences existed in the frequency of personality disorders (males 18.2%. females 8.2%; $p < .05$). The teachers who did not return the follow-up questionnaire showed no significant differences from those who participated twice, except for men with the diagnosis of personality disorders (20% vs. 9%, not returning vs. returning questionnaire). The percentage of teachers with burnout risk (EE >26) was in the group tested	Limitations identified by author: Number of participating high school teachers compared to those from other school types is relatively narrow. The lack of control group restricts the results. For all participants an inpatient clinical treatment was indicated. Thus due to ethical considerations the acute symptoms of the treated teachers would have complicated the realization of a control group setting. Lack information on a comparable extensive intervention study which could present an appropriate inpatient control group. Studies comparing results of treatment and control groups were mostly

Study details	Population and setting	Research aims/objectives; research questions/hypotheses Allocation of individuals to intervention/control	Outcomes and methods of analysis	Results	Notes by review team
psychotherapeutically oriented inpatient treatment supplemented by a job-specific intervention shows long-acting success and whether gender gaps and differences between high school teachers and teachers of other levels exist. Study design: Quality score -	men and 16 women) 166 worked at other schools (48 men and 118 women Eligible population: Patients were referred to the clinic by external medical specialists who considered outpatient therapy insufficient. State if eligible population is considered by the study authors as representative of the source population. Selected population: Patients with an acute psychosis or a florid addiction were excluded from the investigation, But of the admitted teachers 100% participated in the study Include potential	medical examination. Based on these results a psychodynamic treatment was developed and performed by a team of physicians, psychologists, kinesiotherapists, gestalt therapists and nurses. The approach was holistic and included the concept of combining all areas of the clinic as a therapy location. In addition to the discussion therapy group meeting twice a week, these patients had three sessions of gestalt therapy and concentrative kinesiotherapy. The participants were able to symbolise their conflicts and problems. Once per week patients came together in a burnout group to discuss	burnout: Emotional exhaustion, depersonalisation and personal accomplishment. Items measured on a 7 point Likert scale. Questions about weekly working time (including working hours at home) were for the last working week and were evaluated for comparable conditions (full-time work, uninterrupted by holidays or illness in the last week) Follow-up periods: Follow up mail survey was conducted on	first between 72.3% (neurotic disorders) and 82.6% (personality disorders, $p > .05$; depressive disorders 80.3%, total group 78.5%). There was neither an age difference at the first examination nor a distinction of MBI results between responders and non-responders ($p > .05$). Out of 150 teachers, who had participated in the follow- up survey 112 (74.7%, males 76.1%, females 74.0%) were still active or had resumed teaching. The percentage of retired or no longer teaching participants increased with age. The weekly working hours of teachers who had resumed work decreased slightly from 38.1 to 35.5 hours (non-significant). The percentage of those who were not ill in the last	obtained with outpatients, internet based or by investigations at work. Limitations identified by review team: May not be generalizable across other countries or sectors Source of funding: No funding was obtained for this study

Study details	Population and setting	Research aims/objectives; research questions/hypotheses Allocation of individuals to intervention/control	Outcomes and methods of analysis	Results	Notes by review team
	sources of bias. NR Excluded population/s: Patients with acute psychosis or florid addiction were excluded from the investigation.	concrete everyday problems or the organisation of work to be performed at home with colleagues, school management, pupils and pupils' parents. The members defined the topics to be dealt with. The programme of stress management for teachers elaborated on by Kretschmann in 2001formed the basis of behaviour therapy in a group setting. Each session was 100 minutes in duration. In addition, two one-on-one depth psychology conversations lasting 50 minutes took place. The topics of these discussions were conflicts, interpretation and work on behaviour and reactions shown during group psychotherapy. The main emphasis of this	average 2 years after treatment termination, one year at the earliest – using the same questions as at the first examination. Method of analysis: Missing data of the MBI (1.4% of all items; incomplete data sets in 7.5% of first and 9.3% of follow-up surveys) were replaced by calculated personal mean values of the corresponding MBI factor if only one value of the corresponding factor was missing. The answers were	quarter increased from 29.5 to 51.8% ($p < .001$), the number of days off due to illness (all employees) in the last quarter decreased to less than one-third. There was also an essential improvement of MBI scores of EE. High school teachers showed a statistically significant higher score of emotional exhaustion compared to teachers of other levels ($p < .05$). The difference disappeared after treatment ($p = .599$). At the first survey, males had higher EE scores ($p < .0001$) and DP scores ($p < .05$) than females and lower PA scores ($p < .05$); however, males only had higher EE scores ($p < .05$) at the follow-up survey. In the follow-up survey, female and male participants demonstrated improvements in the	

Study details	Population and setting	Research aims/objectives; research questions/hypotheses Allocation of individuals to intervention/control	Outcomes and methods of analysis	Results	Notes by review team
		connection was to obtain information about techniques of work and time management as well as the analysis of subjective feelings justifying teachers' behaviour at work, with the objective of changing their attitudes. This procedure was directed towards the patients' resources and to their sound personal characteristics thus therapeutically improving their competence. The initiation of further outpatient therapy and/or the discussion of possibilities of supervision to perform a professional self-reflection at home was another element of the inpatient treatment.	evaluated statistically (t-tests with paired random samples to compare the results of survey periods, t-tests with unpaired random samples for group comparisons as well as corresponding Chi square tests to check frequency differences) using the programme Statistica 7 (Statsoft Inc., Tulsa, Oklahoma, USA).	subscale values of high EE, high DP, and low PA. Total sample: Baseline Follow-up (all time points) End-point Intervention group(s): Baseline Follow-up (all time points) End-point Control group(s) Baseline Follow-up (all time points) End-point Attrition details: Indicate the number lost to follow-up and whether the proportion lost to follow-up differed by group (ie invention vs control) There was a 25% non-response rate to the follow-up questionnaire.	

Midtsundstad and Nielsen (2014)

Study details	Population and setting	Research aims/objectives; research questions/hypotheses Allocation of individuals to intervention/control	Outcomes and methods of analysis	Results	Notes by review team
Authors: Tove I Midtsundstad and Roy A Nielsen Year: 2014 Citation: Scand J Public Health 2014 42: 207 originally published online 21 November 2013 Country of study: Norway Aim of study: The study aimed to examine the effect of preventative measures and	Source population/s: Country of study Norway Setting 713 Workplaces with at least 10 employees, and 1 employee over the age of 60 in 2005. Location (urban, rural) Unknown Sample characteristics, including population demographics 2001 Treatment population: 41.3% female Mean age 56 (SD 4.4) 2001 non-treatment 43.6% female Mean age 56.2 (SD 4.5) 2007 Treatment 40.5% female Mean age 56.9 (SD 4.7) 2007 non treatment 40.8% female Mean age 56.7 (SD 4.7)	Research aims/objectives The study aimed to examine the effect of preventative measures and work adjustment, initiated and financed by the establishments, on the probability of sickness absence among workers aged over 50 years. Research questions/hypotheses Do work place prevention measures affect the probability of sickness absence amongst older workers? Method of allocation: Data taken from random sample of Norwegian establishments / employers and cross section of employees from register data. No researcher allocation. 41% of establishments had some form of preventive workplace measure in place in 2007. It is unknown when these were put in place. Intervention/s description: Describe intervention in detail	Outcomes: Include details of all relevant outcome measures and whether measures are objective or subjective otherwise validated The only objective outcome measure was sickness absence certified by a physician and lasting for more than 16 days. Coded into binary variable indicating is individual had a sickness spell lasting over 16 days. Follow-up periods: 6 years between measures (2001 – 2007) Method of analysis: indicate if ITT or completer	Report results for all relevant outcomes: TBC Examples include mean values, confidence intervals, p values, standard deviations, standard errors, effect size, odds ratios, relative risks or any other relevant statistical detail reported by the original study authors. Note any results that detail impact on health inequalities. Model 1 In the unadjusted model, individuals in establishments without measures saw no change in sickness absence levels. Employees in establishments with measures saw a 10% drop in OR in the period 2001 to 2007. Health measures OR 1.20 CI 95% = 1.12-1.28. Change 2001 – 2007 OR = 0.97, CI 95% = 0.9101.03) Measure x change OR = 0.89,	Limitations identified by author: Does not include detail of what the preventive measures are only if they were in place It is unknown when measures were put in place. Authors state that it is unlikely establishments, with and without measures are similar is all manners. Also unlikely respective employees are similar. This is addressed by adjustments in the models (see methods of analysis) Limitations identified by review team: Does not identify whether individuals actually access an intervention.

Study details	Population and setting	Research aims/objectives; research questions/hypotheses Allocation of individuals to intervention/control	Outcomes and methods of analysis	Results	Notes by review team
work adjustment, initiated and financed by the establishments, on the probability of sickness absence among workers aged over 50 years. Study design: Difference-in-difference approach using representative survey data and register data on demographic variables. Quality score[4] (++, + or -) -	Eligible population: Describe how individuals, groups or clusters were recruited Data on establishment's taken from a (random sample) survey carried out amongst 713 establishments, 73% response rate. Sectors included: manufacturing, construction, retail, hotels and restaurants, public administration, education, health and social services, and "other" industries. Data on individual characteristics, work and sickness absence taken from 'Statistics Norway's' registries. A cross section was taken from 2001 (14,261) and 2007 (18,960). In sum 33221 individuals' record were taken. State if eligible	including: "Preventive measures" was whether establishments had arrangements to facilitate work among employees with health problems or reduced capacity. Initiated and financed by the establishment Control/comparison's description: (as above) Employees who do not work for a company with preventive measures in place. Sample sizes at baseline Total sample N=Intervention group(s) 2001 N = 5885 2007 N = 7957 N= Control group(s) 2001 N=8376 2007 N= 11,003 Baseline comparisons: Report any baseline differences between groups in important confounders. Authors state that it is unlikely establishments, with and without measures are similar is all manners. Also unlikely respective employees are	analysis was used and if adjustments were made for any baseline differences in important confounders. Difference-in-differences approach where changes in likelihood of sickness absence over time between with and those without access to preventive work place measures. Logistic regression – reporting odds ratio with 95% CI As a control, linear probability models used to substantiate reported estimates. Models were run, to adjust for employee	CI 95% 0.81-0.97 Model 2 Adjusting for individual characteristics does not alter this finding. (Health measures 1.18, CI 95% = 1.10-1.27. Change 2001 – 2007 OR = 1.00 CI 95% = 0.94-1.06. Measure x change OR = 0.90, CI 95% = 0.82-0.96) Model 3 and 4 Adjusting for establishment characteristics and establishment characteristics and individual characteristics finds unadjusted effects can be accounted for by establishment characteristics and not by workplace measures themselves. (model 3: Health Measures OR = 1.17, CI 95% = 1.09-1.26. Change 2001 – 2007 OR = 1.17, CI 95% 1.09-1.26, Measure x change OR = 0.91, CI 95% o.84 – 1.00. Model 4: Health Measures OR = 1.15, CI 95% 1.07-1.23. Change	Does not control for differences in individuals health status, working environment and working conditions. Evidence gaps and /or recommendations for future research noted by study author. Addressing the above limitation was suggested as further areas for study. Source of funding: For example, government (NHS), voluntary/charity, pharmaceutical company and the role of funding organisations Funded by FARVE – the Norwegian Labour and Welfare Administration

Study details	Population and setting	Research aims/objectives; research questions/hypotheses Allocation of individuals to intervention/control	Outcomes and methods of analysis	Results	Notes by review team
	population is considered by the study authors as representative of the source population. Yes – two pooled cross sections taken in 2001 and 2007. At both time points establishments with and without preventive measures were similar regarding the distribution of gender, employees mean age and SD, education level and percentage being partly disabled. 41% of employees in 2007 worked in establishments with arrangements to facilitate work among employees with health problems or reduced work capacity. Selected population: Employers with at least 10 employees, 1 aged 60 plus in 2005. Employees aged 50 or	similar. This is addressed by adjustments in the models (see methods of analysis) Study sufficiently powered: Not reported	characteristics and to adjusted for establishment characteristics. Models were also run for each sector.	2001 – 2007 OR = 1.15, CI 95% 1.07-1.23, Measure x change OR = 0.92, CI 95% 0.84 – 1.01.) Measures were shown to effect public sector employees. Health Measures OR =1.70, CI 95% = 1.37–2.11 Change 2001 – 2007 OR = 1.27 CI 95% = 1.06–1.52 Measure x change OR = 0.60, CI 95% = 0.45–0.79 Total sample: Baseline Follow-up (all time points) End-point Intervention group(s): Baseline Follow-up (all time points) End-point Control group(s) Baseline Follow-up (all time points) End-point Attrition details: Indicate the number lost to follow-up and whether the proportion lost to follow-up differed by	

Study details	Population and setting	Research aims/objectives; research questions/hypotheses Allocation of individuals to intervention/control	Outcomes and methods of analysis	Results	Notes by review team
	older in 2001 and 2007 respectively Include potential sources of bias. They do not know exactly when different measures were introduced. They assume that most came after 2001 when the IW agreement began. Excluded population/s: (as above) Non employees Aged below 50			group (ie invention vs control) None lost	

Appendix 6: Bibliography – Included Studies

Härmä, M., Tarja, H., Irja, K., Mikael, S., Jussi, V., Anne, B., Pertti, M. (2006). A controlled intervention study on the effects if a very rapidly forward rotating shift system on sleep-wakefulness and well-being among young and elderly shift workers. *International Journal of Psychophysiology*, 59. 70-79.

Midtsundstad, T. and Nielsen, R. (2014) Do work-place initiated measures reduce sickness absence? Preventive measures and sickness absence among older workers in Norway? *Scandinavian Journal of Public Health*, 42: 207

Reetta Rutanen, Clas-Håkan Nygårda, Jaana Moilanena, Tomi Mikkolac, Jani Raitanena, Eija Tomasd and Riitta Luotob (2014). Effect of physical exercise on work ability and daily strain in symptomatic menopausal women: A randomized controlled trial. *Work* 47, 281–286

Wagner, S., Kascehl, R., Paulsen, S., Bleichner, F., Knickenberg, R.H. and Beutel, M.E. (2008). Does a cognitive-training programme improve the performance of middle-aged employees undergoing in-patient psychosomatic treatment? *Disability and Rehabilitation*, 30 (23), 1786-1793

Wegner, R., Berger, P., Poschadel, B., Manuwald, U and Baur, X. (2011). Burnout hazard in teachers results of a clinical-psychological intervention study. *Journal of Occupational Medicine and Toxicology*, 6, 37-42.

Appendix 7: Bibliography – Excluded Studies and Reasons for Exclusion

Population

Publication Details

Setting

Not set in one of the designated countries (see Appendix 2) or did not focus on workers or a workplace

Relevance

Boockmann, B., Fries, J., & Göbel, C. (2012). *Specific measures for older employees and late career employment* (No. 12-059). ZEW Discussion Papers.

Iweins, C., Desmette, D., & Yzerbyt, V. (2012). Ageism at Work: What Happens to Older Workers Who Benefit from Preferential Treatment? *Psychologica Belgica*, *52*(4), 327-349.

Niessen, C., Swarowsky, C., & Leiz, M. (2010). Age and adaptation to changes in the workplace. *Journal of Managerial Psychology*, *25*(4), 356-383.

Nowalk, M. P. et al. (2010). Improving influenza vaccination rates in the workplace: a randomized trial. *American journal of preventive medicine*, *38*(3), 237-246.

Stenberg, A., X. Luna, et al. (2012). "Can adult education delay retirement from the labour market?" Journal of Population Economics 25(2): 677-696.

Terry, P. E. et al. (2013). Analyzing best practices in employee health management: how age, sex, and program components relate to employee engagement and health outcomes. *Journal of Occupational and Environmental Medicine*, *55*(4), 378-392.

Wallen, E. S., & Mulloy, K. B. (2006). Computer-based training for safety: comparing methods with older and younger workers. *Journal of safety research*, *37*(5), 461-467.

Focus

Study focussed on chronic illnesses

Alcover, C. M., Topa, G., & Fernández, J. J. (2014). Organizational management of older workers and the processes of maintaining, extending and leaving employment. *Papeles del Psicólogo*, *35*(2), 91-98.

van Oostrom, S. H. et al. , (2010). Economic evaluation of a workplace intervention for sick-listed employees with distress. *Occupational and environmental medicine*, *67*(9), 603-610.

Shaw, W. S. et al. (2014). Manage at work: a randomized, controlled trial of a self-management group intervention to overcome workplace challenges associated with chronic physical health conditions. *BMC public health*, *14*(1), 515.

Intervention

Did not study workers over fifty (at least 51% of the population), or the impact of interventions on workers over fifty

Boedeker, W. et al. (2008). The impact of work on morbidity-related early retirement. *Journal of Public Health*, *16*(2), 97-105.

Carr, D. C. and B. L. Kail (2013). "The influence of unpaid work on the transition out of full-time paid work." Gerontologist 53(1): 92-101.

Chen JC, Linnan L, Callahan LF, Yelin EH, Renner JB. (2007) *Workplace policies and prevalence of knee osteoarthritis: the Johnston County Osteoarthritis Project.* Occupational and Environmental Medicine, 64(12): 798-805.

Leijten FRM et al. (2014). Associations of work-related factors and work engagement with mental and physical health: a 1-year follow-up study among older workers. *Journal of Occupational Rehabilitation* doi:10.1007/s10926-014-9525-6

Ng, Thomas W.H. and Feldman, Daniel. (2010) Idiosyncratic deals and organizational commitment.*Journal of Vocational Behavior,* Volume 76, Issue 3, June 2010, Pages 419–427

Terry, P. E., Fowles, J. B., & Harvey, L. (2010). Employee engagement factors that affect enrolment compared with retention in two coaching programs—the ACTIVATE study. *Population health management*, *13*(3), 115-122.

Virtanen, M. et al. (2010). Overtime work and incident coronary heart disease: the Whitehall II prospective cohort study. *European Heart Journal*, *31*(14), 1737-1744.

Virtanen, M., T. Oksanen, et al. (2014). "Extending Employment beyond the Pensionable Age: A Cohort Study of the Influence of Chronic Diseases, Health Risk Factors, and Working Conditions." PLoS ONE 9(2): 1-8.

Wagenaar, A. F. et al. (2014). Who gets fired, who gets re-hired: the role of workers' contract, age, health, work ability, performance, work satisfaction and employee investments. *International archives of occupational and environmental health*, 1-14.

Williams, L. C., & Day, B. T. (2011). Medical cost savings for web-based wellness program participants from employers engaged in health promotion activities. *American Journal of Health Promotion*, *25*(4), 272-280.

Methodology

Qualitative study, did not explicitly measure health and wellbeing, unclear methodology, no control group or longitudinal element

Bal, P. Matthijs, De Jong, Simon B., Jansen, Paul G. W. and Bakker, A. (2012) Motivating Employees to Work Beyond Retirement: A Multi-Level Study of the Role of I-Deals and Unit Climate, *Journal of Management Studies,* Volume 49, Issue 2, pages 306–331

Cau-Bareille, D., Gaudart, C., & Delgoulet, C. (2012). Training, age and technological change: Difficulties associated with age, the design of tools, and the organization of work. *Work: A Journal of Prevention, Assessment and Rehabilitation*, *41*(2), 127-141.

Midtsundstad, T. and H. Bogen (2014). "Active Aging Policies between Individual Needs and Collective Goods. A Study of Active Aging Policies and Practices in Norway." Nordic Journal of Working Life Studies 4(2): 139-158.

Skoglund, B., & Skoglund, C. (2005). Can age management promote work ability among older workers?. In *International Congress Series* (Vol. 1280, pp. 392-396). Elsevier.

Strijk, J. E. et al. (2009). The Vital@ Work Study. The systematic development of a lifestyle intervention to improve older workers' vitality and the design of a randomised controlled trial evaluating this intervention. *BMC Public Health*, *9*(1), 408.

Yen, L. et al. (2006). Participation in employer-sponsored wellness programs before and after retirement. *American journal of health behavior*, *30*(1), 27-38.

Appendix 8: References

Bevan S (April 2010). The Business Case for Employees Health and Wellbeing. A report prepared for Investors in People UK. The Work Foundation.

Black C (March 2008). Working for a Healthier Tomorrow. Dame Carol Black's Review of the health of Britain's working age population. Department of Work and Pensions.

Bloomer E (2014) *Local action on health inequalities: Increasing employment opportunities and retention for older people*, Public Health England.Cacioppo, J. and Patrick, W. (2008) Loneliness: Human Nature and the Need for Social Connection. New York: Norton & Company.

Clark, R. Morill, M.S. and Allen, S.G. (2009). *Employer-provided retirement planning programs.* Working paper : North Carolina State University.

Crawford J, Graveling R, Cowie H, Dixon K, MacCalman L (2009), The health, safety, and health promotion needs of older workers. An evidence-based review and guidance. Report submitted to the IOSH Research Committee. IOSH.

Equality and Human Rights Commission (EHRC) (2010). Working Better: The Over 50s, the New Work Generation, Manchester: Equality and Human Rights Commission.

Foresight Mental Capital and Wellbeing Project (2008). *Final Project Report*. The Government Office for Science. London

Hershey, D.A., Mowen, J.C. and Jacobs-Lawson, J.M. (2003). An experimental comparison of retirement planning intervention seminars. *Educational Gerontology*, 29, 339-359.

Hsu Y-S (2013). Training Older Workers –A Review. In: Aging, Work and Society (Field J, Burke RJ and Cooper CL (edit). The Sage Publications, London

Kunze F and Boehm SA (2013). Research on Age Diversity in the Workforce: Current Trends and Future Research Directions. In: Field J, Burke R and Cooper CL (edit.): *Ageing Work and Society*. Sage Publications.

Macleod A., Worman D., Wilton P., Woodman P. and Hutchings P. (2010) 'Managing an ageing workforce – How employers are adapting to an older labour market'. Chartered Institute of Personnel and Development/ Chartered Management Institute Research Report, London: CIPD.

Marmot M (2010) *Fair Society, Healthy Lives: The Marmot Review,* HYPERLINK "http://www.ucl.ac.uk/marmotreview" www.ucl.ac.uk/marmotreview, February 2010

NICE, (2009), Promoting mental wellbeing through productive and healthy working conditions: guidance for employers. NICE public health guidance 22, National Institute for Health and Clinical Excellence, November 2009.

NICE, (2012) Methods for the development of NICE public health guidance (third edition), National Institute for Health and Clinical Excellence, September 2012

Noone, J.H., Stephens, C. and Alpass, F.M. (2009). Preretirement planning and well-being in later life: A prospective study. *Research on Aging,* 31(3), 295-317

ONS (2014), Labour Market Statistics, July 2014

Parry E and Harris L (Dec 2011). The Employment Relations: Challenges of an Ageing Workforce. Acas Future of Workplace Relations, Discussion paper seriesSinclair D, Watson J, Beach B (2013), Working Longer: An EU perspective. ILC-UK

Smeaton D., Vegeris S. and Shain-Dikmen M. (2009) 'Older workers: Employment preferences, barriers and solution's. Manchester: Equalities and Human Rights Commission.

Taylor P (2007), Employment and labour market policies for an ageing workforce and initiatives at the workplace. National overview report: United Kingdom. European Foundation for the Improvement of Living and Working Conditions.

Topa, G., Moriano, J.A., Depolo, M., Alcover, C., Morales, J.F. (2009). Antecedents and consequences of retirement planning and decision-making. *Journal of Vocational Behavior, 75*(1), 38-55.

Vaughan-Jones H, Barham L (2010). *Healthy Work; Evidence into Action.* The Oxford Health Alliance, The Work Foundation and RAND Europe, available at HYPERLINK "http://www.bupa.com"http://www.bupa.com

Waddell, G., & Burton, A. K. (2006). *Is work good for your health and wellbeing*? London: Department for Work and Pensions.