Measuring labour productivity and the benefits of interventions for osteoarthritis

Project report for Medicines Australia

Professor Deborah Schofield\textsuperscript{a,b,c}, Dr Rupendra Shrestha\textsuperscript{a}, Dr Michelle Cunich\textsuperscript{a} and Ms Sarah West\textsuperscript{a}

\textsuperscript{a}Faculty of Pharmacy, Charles Perkins Centre, The University of Sydney, NSW 2006, Australia
\textsuperscript{b}Murdoch Childrens Research Institute, Royal Children’s Hospital, Flemington Road, Parkville, Victoria 3052, Australia
\textsuperscript{c}Garvan Institute of Medical Research, Darlinghurst, NSW 2010, Australia

Phone: +61 2 8627 1951
Fax: +61 2 8627 0141
Email: deborah.schofield@sydney.edu.au

2 September 2016
Acknowledgement

This research was supported by funding from Medicines Australia.
Table of Contents

Executive Summary i
1. Context 1
2. Study objectives 3
2.1. Review of productivity measures in osteoarthritis intervention studies 3
2.2. Propose a measure of productive life years gained 3
2.3. Apply the measure of productive life years gained to OA interventions 4
3. A literature review of productivity measures used in intervention and cost of illness studies of OA (Stage 1) 5
3.1. The burden of OA 5
3.2. Evidence on the economic impacts of OA 6
3.3. Materials and methods 9
3.3.1. Selection of studies 9
3.3.2. Quality of the evidence 10
3.4. Literature synthesis 12
3.4.1. Key information extracted 12
3.4.2. Type of synthesis of results 12
3.4.3. Instruments used 12
3.5. Results 13
3.5.1. Identification of relevant articles 13
3.5.2. Judgements about the quality of the OA studies 14
3.5.3. Synthesis of results 15
3.6. Instruments Used 17
3.7. Single questions about productivity 20
3.8. Monetarisation of productivity losses 21
3.9. Statistical analyses 21
3.10. Productivity measures not captured 22
3.11. Discussion 23
3.11.1. Measuring and reporting productivity losses of OA 23
3.12. Conclusion 25
4. Proposed Productivity Measure (Stage 2) 26
4.1. Framework for productivity measurement 26
4.2. Methods for estimating Productive Life Years Gained (Lost) and related costs 29
4.3. Conclusion 30
5. Productive life years gained and the associated economic gains as a result of OA intervention (Stage 3) 31
5.1. Data and methods 32
5.1.1. Association between SF-36 bodily pain scores and labour force participation 33
5.1.2. Labour force impacts 33
5.1.3. Economic benefits 34
5.2. Results 35
5.2.1. Not in the labour force due to OA 35
5.2.2. Association between SF-36 bodily pain domain scores and labour force participation 36
5.2.3. Impacts on labour force participation 36
5.2.4. Economic impacts 37
5.3. Conclusions 38
6. Conclusions and Future Directions 38
Appendix 1: Inclusion and exclusion criteria of the published studies 49
Appendix 2: Search strategy and list of keywords 50
Appendix 3: Summary of Interventions for OA (25 studies) 51
Appendix 4: Summary of cost of illness studies (37 studies; 1 policy document as last entry) 58
Appendix 5: Study characteristics 71
Appendix 6: List of instruments 72
Appendix 7: Scaffolding of the proposed productivity measure 80
Appendix 8: Cost of not being in the labour force due to OA for individuals in 2015 82
Appendix 9: Labour force status of those with OA in 2015 83
Figures

Figure 1. Article selection process 14
Figure 2. Framework for identifying, measuring and valuing productivity changes 27
Tables

Table 1. A framework for judging the state of the evidence for OA-related productivity losses 10
Table 2. Overview of Productivity Instruments 18
Table 3. Estimated numbers of people not in the labour force due to OA, aged 15-64 years, 2015 36
Table 4. Estimated number of people who would have avoided being out of the labour force due to their OA as a result of starting treatment with NSAIDs for the management of their OA, aged 15-64 years, 2015 37
Table 5. Estimated individual economic benefits of being in the labour force because of reduced pain as a result of managing OA, aged 15-64 years, 2015 37
Table 6. Estimated total economic benefits at national level from increased labour force associated with reduced pain as a result of managing their OA, aged 15-64 years (in thousands, in 2015 AU$) 38
Glossary of Common Abbreviations/Key Terms

Control group: Participants in the comparison group in a randomised controlled trial. They are assigned to placebo, no treatment or usual/standard treatment.

Cost: The value of the next best alternative (opportunity) forgone by integrating resources in a particular service.

Cost effectiveness analysis: A type of economic evaluation in which health outcomes of different interventions vary but are measured in identical natural units (e.g. number of falls prevented) and relative inputs are costed. The interventions are compared in terms of the cost per unit of the health outcome.

Economic evaluation: The application of analytical methods to identify, measure and value costs and health outcomes/consequences (benefits) of interventions and assist in making choices regarding resource allocation.

Friction cost method: A method for measuring the productivity losses (indirect costs) for long-term absence, ill-health and mortality mainly only occurring during the time it takes to replace an employee i.e. the friction cost period (usually 3-6 months). Estimates of the price of labour are set close to zero after the friction period and reduced during the friction period.

HILDA: The Household Income and Labour Dynamics in Australia (HILDA) Survey

Human capital cost method: A method for measuring labour productivity losses (indirect costs) as the lost gross income during the time of absence from work.

IGR: Intergenerational report.

Markov health state-transition models: An in-built, flexible, and clear approach of computer-intensive decision-analytic modelling, including Markov model cohort simulation. The particular problem is conceptualised in terms of a set of health states and the type of transitions among these states; it is a commonly used modelling technique in clinical decision analysis, health technology assessment, and economic evaluation.

OA: Osteoarthritis.

Prevalence: The proportion of people with a particular finding (such as chronic disease) in a given population at a given time.

PRODISQ: PROductivity and DISease Questionnaire.

Productive Life Years (PLYs): Number of people in the labour force per year.

RA: Rheumatoid arthritis


RCT: Randomised controlled trial.

WALS: Work Activity Limitations Survey.

WLQ: Work Limitations Questionnaire.

WPAI: Work Productivity and Activity Impairment Questionnaire.
Executive Summary

This project has three aims. They are:

1. To undertake a targeted literature review of OA intervention studies (and cost of illness) which include productivity measures;
2. To propose a framework based measure of productive life years gained; and
3. To apply the new measure and examine the impact on labour productivity of a pharmaceutical health intervention, with the impacts quantified for individuals and the Australian Government.

Aim 1: Literature Review

In the literature review, we identified (a) the type of productivity measures used in the published studies, and how they have been used; (b) the scale of the productivity costs captured (based on the type of questions/instruments used); and (c) the potential productivity measures not captured.

The targeted search yielded 63 full-text articles, with 25 interventions, 37 cost of illness studies, and 1 policy document. A range of methods were used to estimate productivity losses due to osteoarthritis (OA), from a single question about productivity (such as time to return to work, labour force status, hours of work) to established questionnaires (instruments) where several aspects of productivity were evaluated (such as WPAI, WALS, WLQ). Yet many of the productivity measures derived from these questionnaires have only recently been assessed in terms of their sensibility attributes (i.e. face or content soundness, practicability and feasibility). Estimates of the productivity losses also varied considerably across studies using the same questionnaire (instrument), which is likely to be due to between-study differences in the country and the site of pain-related OA examined. Only 8 intervention studies and 4 cost of illness studies valued the productivity losses. Heterogeneity in the methods used (and how results were reported) for estimating productivity losses made a direct comparison between studies (or pooling of results) difficult and, arguably, inappropriate.

As with other instruments, issues surrounding understandability, suitability and sensibility in the population of interest need to be established for OA. For the 250 OA (or RA) workers in Tang et al (2013), the WLQ-25 was considered most comprehensive (endorsed by 92.8%), the WALS performed best in terms of understandability (97.6%) and suitability of response options (97.9%), and the RA-WIS was favoured in terms of length (91.6%). Consistent sensibility performance between OA and RA
was also found. The WALS (32.6%) and WLQ-25 (30.0%) were moderately preferred in the final overall appraisal.\(^1\)

Despite the high disease burden of OA, evidence on the full economic burden for individuals (and government) is only just emerging. Productivity measures that have not been captured in the studies relate to the translation of lost labour force participation into costs for individuals (lost income, income poverty); costs for government (extra welfare payments and lost taxation revenue); and costs for society (lost GDP attributable to OA).

The productivity costs missing from the literature are:

1. A wider range of productivity losses for individuals. Such estimates would include not only lost income from lost labour force participation (or reduced paid working hours) but also the costs of informal care.
2. Costs to government which consist of extra welfare payments and lost taxation revenue due to the lost productivity of people with OA.
3. Societal costs e.g. the percentage of lost GDP due to OA through its impact on work capacity.

**Aim 2: Propose a Productivity Measure**

In this report it is argued that productivity changes due to interventions for health conditions need to be identified, measured and valued – just like other costs (and some benefits) in economic evaluations.\(^2\) The framework for identifying, measuring and valuing productivity is outlined in Section 4. The proposed measure is Productive Life Years (PLYs) gained which is defined as the number of people who were out of the labour force due to their own ill-health (chronic condition) before treatment and subsequently returned to work in each year - takes into account (a) the measures used by government departments in projecting labour demand and supply, and (b) the dual aims of the government to increase productivity and fiscal sustainability.
Aim 3: Modelling the Impacts of a Pharmaceutical Health Intervention on Individuals and the Government

Pharmaceutical and multidisciplinary interventions have the potential to improve labour productivity of those with chronic conditions. We have used NSAIDs interventions for the management of OA as a case study to illustrate how the significant productive life years gained by those with OA can be counted, measured and valued using microsimulation. This has flow-on economic benefits to individuals in terms of increased earnings in addition to the health benefits, and the government in terms of increased income tax revenue, reduced welfare payments and increased GDP. We note that both the type of intervention and chronic disease are illustrative only; what is important is the demonstration of how PLYs can be counted, measured and valued and how they may assist in decision-making by government departments in relation to resource allocation.

Stage 3: The study found that there was about 43 million dollars per year gained in income for individuals, 11 million dollars in additional tax revenue and 16 million dollars in welfare payments savings for the government by an intervention in OA. Importantly, there was also 164 million dollars per year in GDP – a gain from the societal perspective.
1. Context

It is essential to assess both the benefits and the costs of new medical treatments to ensure that scarce resources (public funds) are allocated to treatments representing value for money.\textsuperscript{2} To determine whether a new treatment (or intervention) is effective in treating/managing/preventing a disease or illness, a clinical study of the treatment in a relevant patient population is needed. Randomised controlled trials (RCTs) are known to generate the highest standard of evidence in medicine.\textsuperscript{3,4} An important consideration within the trial is whether the health effects (benefits) justify the costs, which is determined via economic evaluation such as cost-effectiveness analysis.\textsuperscript{5}

When new treatments are developed and the government is requested to fund these treatments, two aspects need to be carefully assessed: (a) whether they are effective in a clinical sense, and (b) whether they provide value for money i.e. are cost-effective. Thus, when new technologies are approved for funding through processes run by bodies such as the Medical Services Advisory Committee (MSAC) or the Pharmaceutical Benefits Advisory Committee (PBAC), there is a requirement for an economic evaluation to demonstrate the cost-effectiveness of the new technologies.\textsuperscript{6,7} This is the main source of information for policymakers in making decisions about resource allocation.

The benefits of introducing new interventions commonly extend beyond the health system,\textsuperscript{8,9} which are apparent as quantifiable improvements in labour productivity and income for individuals, increased income taxation revenue and reduced welfare payments for the government, and increased Gross Domestic Product (GDP). It has been demonstrated that the value of these (indirect) benefits may be greater than the direct (healthcare) benefits for certain interventions.\textsuperscript{10,11}

In addition to MSAC and PBAC, there are several government departments/agencies involved in making decisions about funding new programs, such as the Department of Social Services, Department of Health, Department of Education, Department of Finance and the Treasury. These departments/agencies need to consider the impact of a spending proposal on the entire budget (i.e. health and other portfolios) as well as the impact on individuals. Coinciding with the release of the first Intergenerational Report (IGR) in 2002,\textsuperscript{12} these departments/agencies have been asked to consider the impacts of their choices in relation to long-run trends in economic drivers, and their impact on the budget.
Consecutive government reports have identified population ageing and the related need for increased labour force participation, as serious pressures on the budget; see the IGRs for 2002\textsuperscript{12}, 2007\textsuperscript{13}, 2010\textsuperscript{14}, and 2015\textsuperscript{15}. As a consequence, the Commonwealth Government established the National Commission of Audit to Government – an independent body – to assess and report on the performance, functions, and roles of the government in late 2013. Amongst other things, the Commission reported:

\begin{quote}
Once the impacts of an ageing population and expected lower growth prospects in the longer term are taken into account a growing fiscal gap will emerge at all levels of government across Australia if current expenditure and revenue policies remain unchanged...Today we have five people working for every one retired person, by 2050 we will only have 2.7. This means those 2.7 people working will need to produce as much as the 5 today just to maintain our standard of living.
\end{quote}

(National Commission of Audit, 2014)\textsuperscript{16}

In response to this report, the government implemented various policies to improve labour force participation in the longer-term. For example, the government promoted deferred or gradual retirement in an effort to ease the slow-down in the reduction of the labour force due to population ageing. In the 2014 budget, Treasurer Joe Hockey announced that, in addition to the government raising the age eligibility of the aged pension to 67 years by 2023, it will be increasing eligibility to 70 years by 2035.\textsuperscript{17} This was despite the Productivity Commission (2013) announcing that there was likely to be increased longevity without increased years of healthy life in Australia. The Productivity Commission provided evidence that, over time, there has been an expansion of morbidity, rather than a compression, in Australia.\textsuperscript{18} This has implications for the ability for Australians to participate in the labour force as they age, with an increasing likelihood of living with a medical condition.

The paradox of needing to increase labour force participation among older Australians (aged 45 and over), many of whom suffer from work-limiting health conditions, is not a new challenge and creates a dilemma. Using Health&WealthMOD2030, we have previously estimated that there are 347,000 people aged 45-64 years with lost productive life years (PLYs) due to chronic conditions in 2010 which is projected to increase to 459,000 by 2030 (a 32.3\% increase over 20 years).\textsuperscript{10} The top chronic conditions resulting in lost PLYs among people aged 45-64 years in both years were: back problems (dorsopathies), arthritis, mental and behavioural disorders, depression (excluding postnatal depression), and cardiovascular disease.\textsuperscript{10}

The paradox of trying to increase labour force participation whilst having limiting health conditions is a challenge at the intersection of social, health, and economic policy. It requires a consolidated and consistent approach (through cross-portfolio alignment) by government to address it. Few government reports have taken into account the combined effect of different portfolios on the overall budget, which has led to little significant impact on policy change in Australia with the
notable exception of the IGRs. Prominent government reports that emphasise the essential role of labour productivity in determining fiscal sustainability and national prosperity (such as the IGRs)\textsuperscript{16–18} are important sources that highlight the need for this study.

2. Study objectives

The aim of this study was to examine the potential productivity benefits of investing in a pharmaceutical intervention using osteoarthritis (OA) as an example. The study focuses on OA because of its significant effect on the labour force participation of Australians aged 45-64 years – who are the fastest growing age group in the working-age population.\textsuperscript{19} In addition, early retirement due to OA reduces the income available to individuals, and places a burden on the government due to lost income taxation revenue and increased welfare payments for retired individuals.\textsuperscript{20}

The project was completed in three stages:

1. Literature Review of OA intervention studies which include productivity measures
2. Propose a framework based measure of productive life years gained; and
3. Apply the new measure and examine the impact on labour productivity of a pharmaceutical health intervention

2.1. Review of productivity measures in osteoarthritis intervention studies

In the first stage, recent scientific literature was reviewed to summarise the type of productivity measures that have been included in OA intervention studies. The following issues were covered in the review:

- What productivity measures are included and how they are used;
- The scale of the productivity benefits captured; and
- Potential productivity benefits not captured

Because relatively few intervention studies for OA with measures of productivity were found, we also searched for relevant cost of illness studies to determine the type of productivity measures included in these studies.

2.2. Propose a measure of productive life years gained

In the second stage, we proposed a general measure of productive life years gained taking account of the measures used by the Treasury in estimating GDP projections and the aims of the government to increase productivity and fiscal sustainability.
2.3. **Apply the measure of productive life years gained to OA interventions**

In the third stage, we applied the measure of productive life years (PLYs) gained by modelling one example of the potential productivity impacts of a pharmaceutical intervention to manage OA from the literature. The intervention captured changes in quality of life (using the Short Form-36 or SF-36) occurring in patients with OA during treatment with non-specific non-steroidal anti-inflammatory drugs (NSAIDs). Although two intervention studies were to be used for this stage of the study, where one was on the impact of medicine on OA and the other was to be about the impact of a lifestyle intervention for OA, the literature search undertaken in stage 1 of the project did not reveal any suitable RCTs or other studies. Thus an additional and more specific search was undertaken to identify the best available study: Rabenda et al (2005). This is a naturalistic, prospective follow-up study of 783 patients with OA in whom general practitioners (GPs) decided to start treatment with non-selective NSAIDs. The Short Form-36 (includes SF-36 bodily pain) and the Western Ontario and McMaster Universities OA index (WOMAC) were assessed at baseline and after 3 months. The effect of NSAIDs for the management of OA on the SF-36 bodily pain domain score over 3 months was used in our modelling to demonstrate the productivity benefit associated with these medicines. We note that the PLYs gained per year due to an intervention that improves the health (and thus work capacity) of people with a chronic condition can be applied to any type of intervention (pharmaceutical, lifestyle, workplace and so on) and any type of chronic disease, provided data/modelling needs can be met.

The example provided in Stage 3 shows that there is a clear connection between treatment with medicine and broader benefits that can be associated with productivity gains.
3. A literature review of productivity measures used in intervention and cost of illness studies of OA (Stage 1)

3.1. The burden of OA

Osteoarthritis (OA) is a degenerative disease of the joints or “wear and tear” form of arthritis. OA is the most common chronic health condition of the joints, and one of the most common chronic conditions.\textsuperscript{22-32} OA ensues when the cartilage or cushioning between the joints deteriorates causing pain, stiffness and swelling. The most common parts of the body affected by OA are the feet, knees, hips, wrists.\textsuperscript{28}

The number of people with OA is increasing due to population ageing and the increased prevalence of several risk factors, such as obesity and sedentary behaviours.\textsuperscript{33} Based on the methods used in the Global Burden of Disease study (1997), the population ageing could make OA the 9th cause of disability-adjusted life years (DALYs) in developed countries by the year 2020.\textsuperscript{34} The total number of years lived with disability (YLDs) worldwide caused by OA of the knee and hip increased by 60.2% between 1990 and 2010, which has meant that OA has moved up from 15th to 11th position in the list of the most frequent causes of disability.\textsuperscript{35} Knee and hip joints were responsible for 2.2% of all YLDs due to OA worldwide (and 0.7% of all DALYs) in 2010,\textsuperscript{36} however, these numbers are undervaluations of the OA burden because they do not include the results for all other joints commonly affected by OA.

OA is responsible for a high number of primary healthcare visits (GPs) as well as knee and hip replacements, and hospital costs more generally.\textsuperscript{32} For example, a recent United States (US) population-based study of healthcare utilisation found that, on average, the cohort with knee OA had about six more doctor visits and four more non-doctor visits each year than the cohort without OA.\textsuperscript{37} The OA cohort also had a significantly higher number of hospital stays (28% more), with the difference mainly being attributed to total joint replacements. In Australia, OA is one of the top 10 problems managed by GPs. Almost 2.7 million Medicare-paid GP consultations in 2007–2008 included the management of OA.\textsuperscript{38} And in 2004–2005, AU$1.2 billion in health expenditure was attributable to OA. The main costs were from in-patient services, for example, joint replacements.\textsuperscript{38} However, the economic burden of OA consists of direct (healthcare) and indirect (non-healthcare) costs. The latter mainly take the form of productivity losses and informal caregiving because of the limited independence (due to reduced physical function) of people with OA.
3.2. Evidence on the economic impacts of OA

Cost of illness studies generate an estimate of the economic impact (cost) resulting from the illness, including direct (healthcare) costs and indirect costs (such as productivity losses, income losses, carer costs, and travel costs). Thus these studies provide an indication of the potential social benefits that could be achieved if the disease were prevented or treated more effectively. However, certain productivity costs (such as lost income, receipt of welfare payments, lost taxation revenue, informal carer costs) are rarely included in cost of illness studies.

Only one systematic review of the cost of OA (cost of illness studies) and two of the effects of OA on productivity have been published in the last five years. With regard to the latter, Bieleman et al (2011) conducted a systematic literature search for studies on patients with hip or knee OA and various work participation measures (such as work status, sick leave, work disability, reduced productivity, lost days of work, and transitions in and out of work). The quality of the methods used in the studies was assessed using a range of criteria, such as the study population (to assess selection bias), the soundness of how the determinants were assessed (OA and possibly confounding determinants of work outcomes) and reported work measures (to assess information bias), and the quality of the data analysis i.e. statistical analysis (to correct for all factors); and a qualitative data analysis was accomplished. Fifty-three full-text articles were selected out of 1,861 abstracts; and the final data for analysis were extracted from 14 articles. Study design, study populations, definitions, and measurements used in the studies showed large variations; and work outcomes were often only secondary objectives in these studies. The outcomes were summarised as showing a mild negative effect of OA on work participation. However, the authors emphasised that “research on the effect of OA on work participation is scarce and the methodological quality is often insufficient” (p. 1835). Moreover, the “longitudinal course of work participation in OA has not yet been completely described” (p. 1840). The review included only one longitudinal study, Gignac et al (2008), who found that in 4.5 years, 63% of study participants (n=490 with 278 OA, 49 OA+RA, 163 RA; mean age 50.9 years; mean disease duration=9.2 years (sd 8.7), at baseline) remained in paid work. Considering that governments need to increase the number of people working at older age to ensure sufficient future revenue, the impact of chronic diseases (and interventions in improving productivity) require urgent attention in well-designed (RCT) studies.

Palmer’s (2012) systematic review considered the impact of knee OA on work participation and the relationship between work and total knee replacement. In terms of agreement, the authors reported that “probably” knee OA considerably prevents the patient from participating in work (measured as labour force participation, attendance at work, and productivity) but the “evidence is limited” (p. 79). Areas of uncertainty centred around the evidence-base for effective interventions (i.e.}
treatments, work changes or adjustments and work policies) to improve participation in patients with knee OA, and how the type of work they are undertaking affects long-term clinical outcomes (e.g. pain, physical function, and the need for revision through surgery) in patients with total knee replacements.

A recent perspectives study (Hunter et al 2014) on the burden of OA for the individual, the health system and society, draws attention to the magnitude of the current problem with some references to projected figures. The main message is that there is a pressing opportunity to make fundamental changes to the way individuals with OA are cared for, that will affect both direct and indirect costs of the disease. By focusing on the burden of this prevalent, disabling, and costly disease, the authors highlight the opportunity for shifts in health policy towards better prevention and chronic disease management through medicine, lifestyle, workplace and other interventions.

Intervention studies provide assessments of improvements in the condition of an individual or group. RCTs constitute a separate but important class of intervention studies. RCTs are carried out in order to compare the efficacy and effectiveness of two (or more) health/medical treatments, surgeries or combinations of these. The target population constitutes patients with a specific disease or symptom. Well designed RCTs generate the highest quality of evidence after systematic reviews; see NHMRC (2000, 2001) guidelines for assessing scientific evidence in medicine/public health. Only one systematic review on the effectiveness of interventions to improve the productivity of people with musculoskeletal disorders (MSD) (includes OA) has been published between 2010 and 2015: Palmer et al. (2012). They included a total of 42 studies (34 RCTs) on the effectiveness of community and workplace interventions to manage musculoskeletal-related sick leave or other absences and job loss sourced from 35 earlier systematic reviews and searches of Medline and Embase between 1990 and 2010. Among the 42 studies, 27 assessed the issue of return to work, 21 the issue of duration of sick leave or other absences, and 5 job losses. Interventions found in the studies included exercise therapies, behavioural change techniques, workplace adjustments and providing additional services for employees with OA. The studies were typically small (median sample=107 [inter-quartile range (IQR) 77-148]) and limited in quality. Most interventions appeared beneficial, with the median relative risk (RR) for return to work being 1.21 (IQR 1.00-1.60) and for avoiding MSD-related job losses it was 1.25 (IQR 1.06-1.71); the median reduction in sickness absences was 1.11 (IQR 0.32-3.20) days per month. Nevertheless, the work effects were smaller in larger and higher-quality studies, which suggested some publication bias. In this particular study (Palmer et al. (2012)), no intervention was clearly found to be superior among the included set; however, the more effort-intensive interventions were less effective than basic interventions. No cost-benefit analyses had been undertaken to establish whether there were statistically significant
net economic benefits associated with the interventions. Note that the Palmer et al’s (2012)\textsuperscript{46} review was also unable to find any RCTs involving interventions to promote productivity in patients with OA specifically, as well as no cost effectiveness analyses to determine whether they provided value for money.

Palmer and Goodson (2015)\textsuperscript{47} recently reviewed the pattern of musculoskeletal complaints in later middle age (separating out OA) and what is known about their impact on employment; and potential interventions to support older workers with musculoskeletal diseases (mainly using Medline, Google Scholar and two relevant journals). They found that musculoskeletal problems are common in older workers and have a substantial impact on their work capacity. Factors found to influence job retention included the type of musculoskeletal disease, its severity and treatments, the presence of comorbid conditions, people's attitudes to work and retirement, and the work environment's demands on them. The type of interventions that may extend working life were outlined for musculoskeletal diseases separately, with the section on OA having only Palmer et al’s (2012) conclusions noted for the period under consideration (2010-2015). Many gaps in the evidence were found, notably on the health risks and benefits of patients with OA continuing in work and about which interventions actually worked best.

Ideally, an economic evaluation runs alongside an RCT and assesses whether the intervention(s) provides value for money compared to standard care. In the review conducted for this study, we were able to find only three economic evaluations of OA interventions (treatments) to improve productivity (Waimann et al (2014)\textsuperscript{48}, Pinto et al (2013)\textsuperscript{49}, Marra et al (2014)\textsuperscript{50}), but only the last two were a cost-effectiveness analysis running alongside an RCT. Pinto et al (2013)\textsuperscript{49} is a cost-effectiveness analysis of an RCT involving manual therapy, exercise or both options, in addition to usual care, for people with hip or knee OA. Marra et al (2014)\textsuperscript{50} is a cost-utility analysis of a pharmacist-initiated multidisciplinary strategy providing value for money compared to usual care in participants with previously undiagnosed knee OA.
Overall, these published reviews highlight that there is a limited number of intervention and cost of illness studies for OA reporting on productivity and thus they have tended not to take a societal perspective. Consequently, the effects of OA interventions on productivity are rarely identified, measured, and valued in OA studies. The lack of estimates on productivity costs has implications for resource allocation. There is also a lack of synthesis on the productivity measures that have been used in OA studies, and how extensive the valuing of (significant) changes in productivity due to interventions has been, or should be e.g. impacts from the individual’s perspective only (such as increased hours of work, increased income, reduced need for informal care), the individual and the governmental perspectives (such as reduced welfare payments, increased taxation revenue), or the societal perspective (such as a greater pool of healthy workers, GDP growth attributable to improving the health of workers).

The objectives of Stage 1 are (1) to review and identify variations in the methods used to estimate productivity costs in OA studies published between January 2010 and June 2015; and (2) to identify a set of productivity measures that are not captured. This section addresses three research questions:

a) What productivity measures are included in the studies? How have they been used?

b) What is the scale of the productivity benefits captured? and

c) What potential productivity benefits are not captured?

3.3. Materials and methods

A targeted literature search on medical and economics databases (from 1 January 2010 to 11 June 2015) was performed to identify OA studies including productivity costs. Studies were identified by searching five electronic databases: three medical databases (Medline, Embase, and Cochrane Library) and two economics databases (Expanded Academic ASAP and EconLit). Grey literature was further checked using Google Scholar. The inclusion and exclusion criteria for the studies to be included in the literature search are listed in Appendix 1 and the search strategy with the list of keywords are presented in Appendix 2

3.3.1. Selection of studies

Two reviewers (SW and MC) were involved in the selection of studies, where studies were screened for relevance (independently) based on the inclusion criteria. In stage one, duplicated articles were identified and removed from the list in the Excel spreadsheet. Following this, the titles and abstracts of all remaining articles were screened for possible full-text retrieval. In stage two, the full-texts of potentially relevant articles were reviewed again based on the inclusion criteria. Articles meeting the inclusion criteria were accepted into the literature synthesis.
3.3.2. Quality of the evidence

All accepted studies were judged for their methodological quality using the National Health and Medical Research Council (NHMRC) of Australia’s frameworks for evaluating scientific evidence and economic evaluations (NHMRC 2000, 2001)\textsuperscript{4,5}.

<table>
<thead>
<tr>
<th>Evidence</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relevant and replicable</td>
<td>Did the study identify productivity as an appropriate and relevant outcome? Is the study design stated clearly enough so that it could be replicated?</td>
</tr>
<tr>
<td>Theoretical framework</td>
<td>Is there a logical (scientific) reason why the OA intervention will have the desired effect?</td>
</tr>
<tr>
<td>Strength of the evidence</td>
<td>Is it a robust study design?</td>
</tr>
<tr>
<td>Size of the effect</td>
<td>Did the $P$-value or confidence interval reasonably exclude chance? Is the effect size of economically important?</td>
</tr>
<tr>
<td>Transferable and representative</td>
<td>What are the benefits and costs (includes productivity) of the intervention and the comparator? Do they differ between patient groups? Is the study population representative of the population for whom the OA intervention is intended?</td>
</tr>
<tr>
<td>Duration</td>
<td>Are the effects (includes productivity) sustained over a relevant time horizon? And who benefits from the OA intervention – individuals, government, or society?</td>
</tr>
<tr>
<td>Value for money</td>
<td>Is the intervention(s) cost effective (includes productivity outcomes) relative to standard care?</td>
</tr>
</tbody>
</table>

Source: Adapted from NHMRC (2000, 2001)\textsuperscript{4,5} and Schofield et al (2014).\textsuperscript{51}

The NHMRC’s (2000, 2001)\textsuperscript{4,5} criteria for assessing medical intervention and other health-related studies was used as the basis for making judgements about the OA studies in this review. These criteria are summarised in Table 1. The first is whether the study states the type of effects or outcomes (productivity) which may be achieved through the intervention, and whether these outcomes have been measured in appropriate units. For example, whether the OA intervention was intended to improve health and productivity outcomes, or save money but produce no worse health outcomes (non-inferiority). Additionally, the study should have a reasonable (main) hypothesis.
and/or scientific explanation (and perhaps a theoretical model) underpinning it – information that can be used to explain why the effect is expected from the intervention.

There are a range of possible study designs. According to the NHMRC (2000), the highest quality of evidence comes from systematic reviews of randomised controlled trials (RCTs; Level I). Systematic reviews determine whether a treatment (intervention) effect can be replicated and, by pooling the results of RCTs, provides a best estimate of the magnitude of the effect. The second level of evidence comes from well-designed RCTs, implying they have the strength to minimise bias (Level II). Pseudo RCTs, comparative studies, case control studies or cohort studies have the potential to introduce bias and are ranked as Level III. Case series (pre- and post-test studies) suffer from problems related to the lack of randomisation (e.g. non-comparability of control and treatment groups, different outcome measures for the two groups) and are ranked as Level IV. The lowest level of evidence – now excluded by the NHMRC – is expert opinion and consensus from expert committees because these sources do not have a scientific basis (Level V).

Although a study may report a statistically significant effect, it does not necessarily mean that the effect has wider importance i.e. clinical, policy, or economic. For OA treatments with productivity as an outcome, we would want to know what the cut-off was for the productivity scale (instrument) to be regarded as statistically important (i.e. gaining how many points on the instrument would be considered as significant). In this case, it would also be important to know whether the difference in productivity was maintained over time or became insignificant between the intervention and controls.

As public funds become increasingly scarce due to continuous deficits (as the government tries to cope with the ageing population) and external shocks (e.g. recessions, mining boom closing off), ensuring that only health interventions representing efficacy and efficiency are implemented is essential where the latter is determined via economic evaluation.

The main criteria considered in making judgements about the OA studies (and reported in Tables 2 and 3) were the study population, study design, and relevance (and potential usefulness) of the productivity measure(s) used. With regard to the latter, the reviewers considered whether the productivity measure(s) was a single or standalone question specific to the study or well recognised productivity questionnaires (with validated/tested for suitability instruments) suitable for the OA population. Together, these judgements help to inform readers about the quality of OA studies concerning our research questions.
We note that Bieleman et al’s (2011)\textsuperscript{41} systematic review of studies involving patients with hip or knee OA and work participation (various measures) included a comprehensive assessment of the methodological quality of the full-text articles in the final selection.

3.4. Literature synthesis

One reviewer (SW) extracted data from the included studies and a second (MC) checked the accuracy. Any disagreement was resolved by discussion between the two reviewers.

3.4.1. Key information extracted

The current review focused on identifying the methods for estimating productivity losses attributable to OA. Key information was extracted from eligible studies at the study level. The information extracted consisted of the authors; publication details (journal, year, pages); title; abstract; type of study, defined as either “intervention”, “cost of illness”, or “policy document”; study population i.e. type of OA (e.g. knee, hip), average age of participants, sample size, timeframe for the study; measure(s) of productivity; how information on productivity had been collected e.g. a single question or an established productivity questionnaire; the instrument(s) used to measure productivity and any sources containing a description, interpretation, or validation of the instrument; the results for productivity i.e. effect size, statistical significance. In some cases, the two reviewers (SW and MC) also reported the main strengths and weaknesses of the study/productivity measures. Studies were firstly grouped as either intervention, cost of illness, or policy documents and, subsequently, the intervention studies were classified by study design (e.g. RCT, cohort study) to determine the strength of the evidence (as per Table 1).

3.4.2. Type of synthesis of results

A qualitative synthesis of results was undertaken. Quantitative comparisons and pooling of the estimates of productivity losses due to OA (or productivity gains due to OA interventions) were not performed due to recognised between-study heterogeneity (e.g. differences in country, methods for estimating productivity, and the range of productivity measures used) which has been noted in systematic reviews of the costs of OA generally, and of the impact of OA on labour productivity specifically.\textsuperscript{41}

3.4.3. Instruments used

The concept, instrument, or technique used in assembling, measuring or quantifying the productivity benefits (losses) in the articles was our working definition of “scale of productivity benefits”.

12
3.5. Results

3.5.1. Identification of relevant articles

A total of 1,557 records were identified through searches of the five electronic databases:

- Medline = 621
- Embase = 824
- Cochrane Library = 9
- Expanded Academic ASAP = 101
- EconLit = 2

There were 262 duplicates, leaving 1,295 articles for screening of titles and abstracts. After screening, 968 articles were rejected for reasons stated in the exclusion criteria. Specific criteria not being met were noted in the Excel spreadsheet. Full-texts were retrieved for 327 articles, and 209 articles were deemed ineligible. The main reasons for ineligibility were: a) the “article” was a conference presentation/abstract only; (b) the study was not original (e.g. a systematic review); (c) the study had been conducted outside the OECD; d) although the study purported to examine productivity losses (gains) for OA (interventions), it did not have productivity as a principal outcome.

Of the 71 full-texts included for full review, 8 were excluded. Thus 63 were left to be included in the literature synthesis. The final selected articles consisted of 25 interventions, 37 cost of illness, and 1 policy document (Figure 1). The main characteristics of these studies including the productivity measures used in the studies are presented in Appendix 3 (intervention studies) and Appendix 4 (cost of illness).
3.5.2. Judgements about the quality of the OA studies

The studies included in the literature synthesis underwent an assessment for quality by the reviewers. The most common areas of concern were:

- The study population was heterogeneous e.g. consisting of people with any OA (site unknown) or people with various forms (sites) of OA (such as knee, hip or both) grouped together, or a different type of arthritis altogether (RA, inflammatory arthritis (IA)). A study population consisting of people with any OA was more common in the cost of illness studies than in the intervention studies (19 out of 37 (51%) versus 2 out of 25 (8%)). Likewise, a population consisting of people with different sites of OA was more common in the cost of illness studies than in the interventions studies (10 out of 37 (27%) versus 2 out of 25 (8%)). A study population consisting of OA and another type of arthritis was equally common in the two types of studies: 3 studies in each.

- The sample size was relatively small. Nine intervention studies (36%) and 2 cost of illness studies (5%) had a sample size smaller than 50, and an extra 4 intervention and 1 cost of illness studies had a sample size of less than 100 individuals.

- The time horizon for measuring costs was short in most cases. Only one intervention study (Ruiz et al (2013)) projected productivity losses over the lifetime of the individuals with OA using markov health state-transition decision modelling. Two other economic evaluations used a relatively long time horizon. Bedair et al (2014) estimated productivity costs for up to 30 years for a hypothetical 50 year old with end stage knee OA; and Mather et al (2014) estimated costs associated with total knee arthroplasty over a 5-year horizon. Only 5 cost of illness studies were based on longitudinal data and thus capable of tracking changes in
productivity over time, with the longest tracking period being 5 years in Datta and Larsen (2010)\textsuperscript{55} and Piscitelli et al (2012)\textsuperscript{56}.

- There was no control group in all of the RCTs or the logical comparator in cost of illness studies (i.e. “individuals without OA”) was missing. In the intervention studies, Verkleij et al (2010)\textsuperscript{57} whose randomised open-label trial involved pharmaceutical drugs did not have a control group – it only compared costs/outcomes to another medication. Almost all of the cost of illness studies only estimated productivity costs for an OA population, with Tang et al (2011)\textsuperscript{58} one notable exception.

A number of other concerns became apparent when compiling the data. The most common reasons for rejecting studies were: (1) the estimate(s) of productivity was not clearly related to having OA or changes in productivity were not clearly related to the OA intervention and thus there was some ambiguity about whether the estimates of productivity were associated with the disease itself (and not another disease such as chronic pain) or an intervention designed to reduce the severity of OA only; (2) the source for the productivity measure (and where it had undergone suitability, reliability or validity checks; a full-description and interpretation of results) was not adequately described in the study and thus it was not possible to ascertain, from the article, whether the measure was valid and/or suitable for an OA population; and (3) the productivity measure(s) was not stated as a study outcome but mentioned elsewhere in the article (e.g. background, discussion). Ambiguity on any of these areas was checked during Stage 2 of the review where the full-text rather than just title and abstract were appraised.

3.5.3. Synthesis of results

The methods for estimating productivity due to OA in the studies are summarised in the last column of Appendices 3 (interventions) and 4 (cost of illness). The most common productivity measure used in the intervention studies was *Time to return to work post-surgery* (11 studies), followed by a particular productivity questionnaire such as *WALs, Work Limitations Questionnaire* (WLQ) (Lerner et al (2001)\textsuperscript{59}, *PROductivity and DISease Questionnaire* (PRODISQ) (Koopmanschap (2005)\textsuperscript{60}. Single productivity costs were measured as income status and living below the poverty line in Allen et al (2014)\textsuperscript{61}; lost income in Bedair et al (2014)\textsuperscript{53} and Marra et al (2014)\textsuperscript{50}; and transitions from employment to receipt of disability pension because of OA in Svendsen et al (2012)\textsuperscript{62}. Some societal costs (such as household assistance, informal care) were ascertained in Marra et al (2014)\textsuperscript{50}, Pinto et al (2013)\textsuperscript{49} and Waimann et al (2014)\textsuperscript{48}. These studies estimated the costs borne by participants with OA and their friends or family including out-of-pocket costs, transportation costs, lost earnings, informal care; and the productivity costs were estimated using the human capital approach (i.e. the gross earnings of those in employment\textsuperscript{2}) in economic evaluations. Fifteen studies used a mix of productivity measures e.g. Baldwin et al (2010)\textsuperscript{63}, Gignac et al (2015)\textsuperscript{64}, Lerner et al (2012)\textsuperscript{65}, Marra
et al (2014)\textsuperscript{50}, Pinto et al (2013)\textsuperscript{49} and Sankar et al (2013)\textsuperscript{66}. Only eight studies provided a valuation of the productivity losses e.g. Mather et al (2014)\textsuperscript{54} and Waimann et al (2014)\textsuperscript{48}. In terms of study design, only 6 studies were randomised controlled trials and thus capable of generating the highest quality of evidence according to NHMRC (2000, 2001)\textsuperscript{4,5} and only 4 included economic evaluations (Marra et al (2014)\textsuperscript{50}, Mather et al (2014)\textsuperscript{54}, Pinto et al (2013)\textsuperscript{49} and Waimann et al (2014)\textsuperscript{48}). In addition to Mather et al (2014)\textsuperscript{54}, there were 3 other Markov health state-transition models for the treatment of “end-stage” OA. Other study designs were: 6 retrospective studies (no control group, often small sample), 1 retrospective case control study (no randomisation, no control group), 1 retrospective multicentre study, 1 retrospective population cohort, 1 longitudinal study, 2 prospective cohort studies, 1 sub-sample of a prospective longitudinal cohort, and 1 nation-wide cohort study (cross-sectional), and 1 cross-sectional single centre study.

For the cost of illness studies, the WPAI was the most common measure used to estimate productivity losses. Most studies (15) were cohort studies (population-based, prospective cohorts) using individual-level data. Of these, 6 estimated productivity from recognised questionnaires (such as the WPAI, WALS, RA-WIS, \textit{Economic Aspects in Rheumatoid Arthritis Questionnaire} (Verstappen et al (2005)\textsuperscript{67}), Medical Outcomes Study Short Form-36, 4 used specific standalone questions about productivity (such as labour force status, employment status, number of days and frequency of sick leave, work disability, work limitations/satisfaction due to pain from OA, early retirement), and 5 measured productivity in terms of receipt of disability pension due to OA (see Rolfson et al (2012)\textsuperscript{68}, Ropponen et al (2011)\textsuperscript{69}, Saatamoinen et al (2012)\textsuperscript{70}, Van Der Burg et al (2014)\textsuperscript{71}) or total sickness benefit/disability pension payments received (see Hubertsson et al (2013)\textsuperscript{72}). There were 13 cross-sectional studies, with 10 estimating productivity losses due to OA using established productivity questionnaires, and 1 study assessing absenteeism and short-term disability and worker’s compensation (Berger et al (2011)\textsuperscript{73}). One notable cross-sectional study is Conaghan et al (2015)\textsuperscript{74}, who not only examined productivity losses to individuals with OA and their carers, but also valued these and other relevant (indirect) costs (such as travel costs, adjustments to the home, welfare payments/assistance received e.g. benefits, disability parking, disability living allowance) in British pounds. Five were longitudinal studies, with 3 using established productivity questionnaires (see Beaton et al (2010)\textsuperscript{75}, Gignac et al (2011)\textsuperscript{76}, Tang et al (2010)\textsuperscript{77}) and 2 using a unique productivity question: “Are you in the labour market now or have you stopped working?” (Datta and Larsen (2010)\textsuperscript{55}) and, for those under 65 years, number of working days lost due to OA (Piscitelli et al (2012)\textsuperscript{56}). Multiple productivity questionnaires were used in one cohort study (Zhang et al (2010)\textsuperscript{78}) and one longitudinal study (Beaton et al (2010)\textsuperscript{75}).
Although well-designed RCTs (intervention studies) produce the “best quality of scientific evidence”, they have tended to not use the “best” (i.e. reliable, suitable, validated for OA population) productivity questionnaires. In contrast, the cost of illness studies – ranked relatively low in terms of quality of evidence by NHMRC (2000, 2001) – have more often than not used established productivity questionnaires. There is a clear trade-off in terms of quality of study design versus quality of productivity instrument – we suggest this is a feature of the literature on chronic disease and productivity in general.

3.6. Instruments Used

More cost of illness studies used a measure of productivity from a conventional questionnaire (n=16, 42%) than intervention studies (n=6, 24%). The most commonly used productivity questionnaires were the Work Productivity and Activity Impairment Questionnaire (WPAI), Work Activity Limitations Survey (WALS), Work Limitations Questionnaire (WLQ), Work Instability Scale for Rheumatoid Arthritis (RA-WIS), and PROductivity and DISease Questionnaire (PRODISQ). A brief description of the different aspect(s) of productivity measured via the questionnaires used in OA studies, how to interpret the results, and some limitations of the measures is provided in Appendix 6. Table 2 provides an overview of the measures, assessing the following important features:

- Capacity to monetarise productivity impacts
- Whether it includes critical subpopulations e.g. those out of the workforce due to the condition
- Whether it includes obvious major costs e.g. lost earnings,
- Whether it includes other major costs e.g. welfare, tax, worker’s compensation
- Whether it has a direct measure of productivity
- Availability (cost and how easy it is to access).

A tick indicates that the questionnaire has this particular feature, and a cross indicates it does not.
## Table 2. Overview of Productivity Instruments

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Key Item</th>
<th>Availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Workplace Activity Limitations Scale (WALS)</td>
<td>Includes critical subpopulations e.g. those out of the workforce due to the condition</td>
<td>Contact the authors; see Gignac et al (2012).</td>
</tr>
<tr>
<td></td>
<td>Includes obvious major costs e.g. lost income</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Includes other major costs e.g. welfare, tax, worker’s compensation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A direct measure of productivity</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Work Limitations Questionnaire (WLQ)</td>
<td>✓</td>
<td>Copyrighted: Work Limitations Questionnaire 1998 by The Health Institute, Lerner D, Amick B 3rd, GlaxoWellcome. Use of the WLQ requires a fee for commercial applications; free of charge for noncommercial applications.</td>
</tr>
<tr>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PROductivity and DISease Questionnaire (PRODISQ)</td>
<td>✓</td>
<td>See the author for accessing information about PRODISQ: Koopmanschap (2005).</td>
</tr>
<tr>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Work Productivity and Activity Impairment Questionnaire (WPAI)</td>
<td>✓</td>
<td><a href="http://www.reillyassociates.net/WPAI_SHP.html">http://www.reillyassociates.net/WPAI_SHP.html</a>: Permission and charges are not required for using the WPAI.</td>
</tr>
<tr>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Endicott Work Productivity Scale (EWPS)</td>
<td></td>
<td>The EWPS is a copyrighted instrument. There is a basic fee charged per user for commercial use. It is accessed through an unrestricted license arrangement.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Work Instability Scale for Rheumatoid Arthritis (RA-WIS)</td>
<td></td>
<td>Coprertied to the Psychometric Laboratory for Health Sciences, University of Leeds. Go to: <a href="http://www.leeds.ac.uk/medicine/rehabmed/psychometric/Scales3.htm">www.leeds.ac.uk/medicine/rehabmed/psychometric/Scales3.htm</a>.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instrument</td>
<td>Capacity to monetarise productivity impacts</td>
<td>Includes critical subpopulations e.g. those out of the workforce due to the condition</td>
</tr>
<tr>
<td>---------------------------------------------------------------------------</td>
<td>---------------------------------------------</td>
<td>-----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Economic Aspects in Rheumatoid Arthritis Questionnaire</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Medical Expenditure Panel Survey (MEPS)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Helsinki Health Study Baseline Questionnaire</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Health and Labour Questionnaire (HLQ)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>World Health Organization’s Health and Work Performance Questionnaire (HPQ)</td>
<td>✓</td>
<td>×</td>
</tr>
</tbody>
</table>
3.7. Single questions about productivity

In the intervention studies, the productivity questions were about:

- Time to return to work post-surgery
- Job disruptions due to OA related issues (treatment, problems)
- Productivity loss (stated generally)
- Absenteeism
- Reduced work hours
- The amount of work time, absences, and unpaid work
- Current employment status when first report back to work
- Societal costs such as expenses due to knee pain or OA (out-of-pocket costs, lost earnings, income status, and transportation costs), household assistance, informal care, and poverty.

In the cost of illness studies, common one-off (or standalone) productivity questions were about:

- Absenteeism
- Worker’s compensation
- Work force status
- Role limitations/satisfaction
- Sickness benefits and disability pension payments
- Retirement due to ill-health
- Working ability (disability)
- Number of days and frequency of sick leave
- Employment status

The majority of intervention studies considered a single productivity measure, with “time to return to work post surgery” being the usual one (n=8, 32%). Only three studies measured productivity costs from a societal perspective: (1) Marra et al (2014) assessed work changes (work time, absences, and unpaid work) and other societal costs (expenses from knee pain or OA, household help, and informal care); (2) Mather et al (2014) assessed the impacts of treatment for end stage knee OA on lost earnings and productivity, and carer costs; and (3) Pinto et al (2013) assessed participants’ loss of hours and the costs borne by participants’ friends or family (out-of-pocket costs, transportation costs, lost earnings, and informal care). Similarly, Svendsen et al (2012) estimated lost productivity as transitions from employment to disability pension or a flex-job within 2 years of having knee (OA) surgery using the Danish transfer (welfare) payments register.

Most cost of illness studies (and all of the intervention studies) estimated productivity losses for OA only i.e. single chronic condition. However, three cost of illness studies estimated the productivity losses for OA only and OA plus comorbidity (all-cause) (Hubertsson et al (2013), Sayre et al (2010), Van Der Burg et al (2014)). These studies reported statistical differences in odds ratios for having OA versus OA plus comorbidity but did not provide a value (dollars) for the differences. The
OA plus comorbidity method for estimating productivity losses involves counting all productivity losses (costs) for any health-related reason in the estimates. The main benefit of recording all-cause productivity losses is the ability to capture the productivity losses attributable to comorbidities.\textsuperscript{81} As Catay et al (2012) have stated, common coexisting conditions such as cardiovascular disease and obesity could be “consequences of osteoarthritis arthritis”.\textsuperscript{81} Thus, the productivity costs associated with these type of diseases are best reported separately in future OA studies.

3.8. Monetarisation of productivity losses

Only 8 of the 25 intervention studies (32%) provided a valuation of the productivity losses of OA i.e. expressed the productivity losses in monetary terms (dollars). The human capital cost method was used to value these losses, with the exception of Pinto et al (2013)\textsuperscript{49} who used the friction cost method. The human capital method takes the patient’s perspective and counts any hour not worked as an hour lost (based on wages). By contrast, the friction cost method takes the employer’s perspective, and only counts as lost those hours not worked until another employee takes over the person’s work (replacement). Both methods can produce widely different results. Thus authors need to be explicit about the method used for valuation of costs.

Four cost of illness studies (11%) provided valuations of the productivity costs.

3.9. Statistical analyses

The majority of the intervention studies (RCTs) estimated the productivity losses of OA over a relatively short (follow-up) period e.g. 6 months, 9 months, or 2 years in de Carvalho et al (2014)\textsuperscript{82}. Some of these studies were also limited in terms of sample size and missing a control group e.g. Verkleij et al (2010)\textsuperscript{57}. Few intervention studies estimated productivity losses using longitudinal data (Gignac et al (2015)\textsuperscript{64} and Sankar et al (2013)\textsuperscript{66} only). Five intervention studies included economic evaluations/Markov state-transition decision models for the productivity (and health) outcomes of people with OA which generated productivity losses over a longer time horizon (e.g. lifetime). Similar statistical methods to those reported below for the cost of illness studies were used to estimate productivity losses (i.e. logistic (multiple) regression analysis).

The majority of the cost of illness studies aggregated the productivity losses for each study participant with OA, and then computed the mean or median productivity loss incurred by those individuals e.g. Kotlarz et al (2010)\textsuperscript{83}. Hermans et al (2012)\textsuperscript{84} reported productivity losses as median productivity losses (lost paid hours) per patient per month as well as total costs. However, a few studies employed a type of “incremental cost” method where the difference between the productivity losses incurred by people with OA and those without OA was computed e.g. Berger et al (2011)\textsuperscript{73}, DiBonaventura et al (2011)\textsuperscript{85}, Datta and Larsen (2010)\textsuperscript{55}, Tang (2011)\textsuperscript{58}. A few cost of illness
studies used a matched control group. Control groups tended to be matched to the OA group by age, sex, and area of residence. Several studies used regression analysis to compute an “adjusted incremental cost” for productivity in people with OA. The main benefit of estimating adjusted productivity losses is that it enables control for individual heterogeneity (i.e. demographic, socio-economic, and health status and/or comorbidity factors). Pit et al (2010)\(^8\) and Saastamoinen et al (2012)\(^7\) estimated productivity losses for people with OA and without OA taking into account comorbidity). The majority of studies used logistic regression analysis to generate these estimates.

Changing productivity losses over time was a common concern raised by authors of the reviewed studies (e.g. Tang et al (2010)\(^7\)). Productivity costs of OA may change due to updated treatment recommendations (and costs) and new clinical practice. Yet, only five cost of illness studies (14%) examined trends in productivity losses of OA.

### Recommendation 1

That the impacts of interventions on productivity during the study period and in the longer-term be identified, measured, and valued. This process requires productivity to be a central issue that is asked about at baseline and follow-up in trials; however, this may still not be sufficient to identify the full impacts of the intervention from the different perspectives of individuals, government and society due to short follow-up. More sophisticated modelling of productivity impacts is needed e.g. microsimulation models using clinical trial data (for the effects) as inputs linked to productivity data from national surveys.

### 3.10. Productivity measures not captured

The productivity costs missing from the literature are:

1. A wider range of productivity losses for individuals. Such estimates would include not only lost income from lost labour force participation (or reduced paid working hours) but also the costs of informal care.
2. Costs to government which consist of extra welfare payments and lost taxation revenue due to the lost productivity of people with OA.
3. Societal costs e.g. the percentage of lost GDP due to OA through its impact on work capacity.

The above costs would be helpful if reported in dollars, making them comparable across time and countries; and from different perspectives i.e. individual, governmental, societal; and projected several years into the future using reliable data sources and sophisticated econometric or microsimulation models. Providing estimates of far-reaching productivity costs due to OA would be helpful for making decisions about proposed research (includes clinical) and policy.
A range of methods were used to estimate productivity losses due to OA in the studies found, from standalone questions about a particular aspect of productivity (such as labour force status, hours of work, sick leave and time to return to work after an intervention) to more comprehensive workplace activities and limitations surveys (such as WALs, WIS, WLQ, WPAI) and productivity scales (EWPS). However, many of these measures are reported in natural units (such as hours or paid work, number of sick days, or are a scale) and thus without conversion to monetary terms no meaningful comparisons can be made across different OA populations (by gender, country, age). This is important because OA is a gendered disease and available and best treatments differ across countries as well as the costs of accessing healthcare (and the flow-on effects to health and work). As noted earlier, only 12 studies produced an estimate of productivity losses due to OA expressed in dollar terms.

3.11. Discussion

This study is a review of published OA studies measuring productivity losses, either as a primary or secondary outcome, focusing on the methods used to do so. The main result is that the methods for estimating productivity losses of OA vary from a one-off question about a single aspect of productivity to established questionnaires where several components of productivity can be evaluated. Estimates of the productivity losses also varied considerably across studies using the same questionnaire (instrument), which is likely to be due to between-study differences in the country under consideration and the location of OA examined. Only 8 intervention studies (32%) and 4 cost of illness studies (11%) provided a value for the productivity losses (monetary terms) of their study population. Heterogeneity in the methods used (and how results are reported) made a direct comparison between studies (or the pooling of results) difficult and, arguably, inappropriate.

3.11.1. Measuring and reporting productivity losses of OA

Despite the high disease burden of OA, the evidence base underlying the work burden is only just emerging. There is a lack of reliable evidence about the productivity losses of OA in many countries including the UK, US and Australia.

A range of productivity measures have been used to estimate the work burden of OA in the included studies, from standalone or one-off questions about some aspect of productivity (such as time to return to work, labour force status, hours of work) to entire productivity questionnaires (such as WPAI, WALs, WLQ). Yet many of the productivity measures derived from these questionnaires have only recently been assessed in terms of their sensibility features (i.e. face or content soundness and feasibility) from the perspective of patients with OA. For the 250 OA (or RA) workers in Tang et al
(2013)\(^1\), the WLQ-25 was considered to be the most comprehensive work questionnaire (supported by 92.8% of the sample with OA or RA), the WALS was found to perform the best in terms of how well it could be understood by participants (97.6%) and the suitability of the options for responding to specific questions (97.9%), and the RA-WIS was found to be the best in terms of the length of the work questionnaire (91.6%). There was also consistent sensibility between OA and RA. The WALS (32.6%) and WLQ-25 (30.0%) were found to be moderately preferred in the final overall appraisal conducted.

A range of data were used to estimate the productivity losses in the reviewed studies, from withintrial only data, OA cohort studies and administrative data (such as national welfare payments registers and worker’s compensation data). Only one well designed study (RCT) capable of generating the highest quality of evidence (NHMRC 2000,2001)\(^4,5\) used an established productivity measure (WLQ) (and provided a valuation) to assess the impact of the intervention on patients, including a dollar value (Lerner et al (2012)\(^6,5\)). Likewise, there were only two economic evaluations that included productivity losses from an established questionnaire in their calculations i.e. the OCCQ in Pinto et al (2013)\(^49\), and PRODISQ in Verkleij et al (2010)\(^57\). While the cost of illness studies tended to use a productivity measure sourced from a recognised questionnaire, the issue with these studies is that their study design (cohort, cross-sectional) automatically generates a lower quality of evidence.\(^4,5\)

Estimates of the productivity losses due to OA expressed in monetary terms was lacking, and this was most common in the cost of illness studies (89% did not provide these valuations). Additionally, reliable “flow-on” estimates of the productivity losses of OA expressed in dollar terms as lost income, extra welfare payments, lost taxation revenue and the percentage of GDP attributable to OA was lacking e.g. only Bedair et al (2014)\(^53\), Lerner et al (2012)\(^6,5\), Marra et al (2014)\(^50\), and Ruiz et al (2013)\(^52\) in the intervention group.

Although cost estimations unavoidably vary to a certain extent, particularly across different populations, the variations observed in this review were substantial even between studies from the same country/patient group (e.g. knee OA only), and were mainly due to methodology i.e. the instrument and costing method used (e.g. human capital versus friction cost methods), the statistical method used to estimate the cost (modelling), and how those costs were presented. This may raise suspicions about the consistency, soundness, and appropriateness of the estimates of productivity losses used in economic and policy evaluations at the present. There needs to be standardisation of the productivity measure used in intervention and cost of illness studies to enable more meaningful comparisons to be made across study populations/countries, and assist policymakers in making resource allocation decisions, especially cross-portfolio decisions.
If standardisation cannot be established for all studies focusing on productivity, then perhaps establishing it for particular types of studies (e.g. interventions) would help to address these types of suspicions. This would be helpful in increasing comparability across, at least, groupings of productivity-focused studies and in identifying the factors for variations in the studies.

### 3.12. Conclusion

This review revealed significant variations among studies on the productivity costs of OA in terms of methodology (instrument), cost estimation (statistical analyses), and presentation of these costs, which raises concerns about the reliability, soundness, and appropriateness of productivity cost estimation in economic/medical/policy impact studies.

**Recommendation 2:** To improve transparency, repeatability and comparability in studies involving productivity, there needs to be greater efforts made for standardisation.

Standardisation is needed in order to improve transparency, repeatability and comparability of studies on the productivity losses of OA.
4. Proposed Productivity Measure (Stage 2)

As identified in section 3.10, there were three areas of productivity that were not adequately captured in previous studies:

1. Productivity losses for individuals including lost income and costs of informal care;
2. Costs to government from additional welfare payments and lost tax revenue; and
3. Societal costs as measured through percent and dollar value of GDP lost.

In this section we propose a measure of productivity that captures these three areas. Our proposed general measure of productivity, called Productivity Life Years (PLYs) gained (lost), takes into account measures used by policy makers in Australia. The proposed measure conceptually aligns with the aims of the Commonwealth Government to increase productivity and fiscal sustainability.

The framework for the productivity measure is described in Section 4.1.

4.1. Framework for productivity measurement

As a first step we have developed a framework for identifying, measuring and valuing productivity gains (losses) in intervention studies. This framework consists of “scaffolding” based on compiling an interrelated set of cascading questions. The starting question is about whether a person is in paid employment or not; followed by their usual hours of work per week; and if not full-time, then whether the person’s ill-health has been the main factor for the reduced labour force participation (and OA as the condition causing it) (see Figure 2). Once the nature of the person’s work has been identified, we consider how to extract information that quantifies the productivity loss e.g. current annual salary, sources of income (including disability pension, workers compensation), number of sick leave days due to OA (multiplied by the individual’s hourly wage). The measure of PLY forms part of this framework. A list of questions that have been used as the basis of the measure is found in Appendix 7.

All of this information can be asked within an RCT at baseline and perhaps at 3-month intervals thereafter for the life of the trial (to obtain immediate impacts of the intervention on productivity) and at follow-up (to obtain information about the longevity of the intervention impacts on productivity).

**Productive Life Years (PLYs)** is defined as the number of people in the labour force in any given year. PLYs lost due to a chronic condition is the number of people who were out of the labour force due to that chronic condition. PLYs gained due to an intervention is the number of people who were out of the labour force before treatment and subsequently returned to the labour force in any given year.
Figure 2. Framework for identifying, measuring and valuing productivity changes

The framework questions allow the identification of whether the individual is in paid work, their income from that work (earnings), how any health conditions they have affected their work, and any welfare payments they receive (and thus add to their total income). For those not in paid work, the framework identifies the main reason for being out of the labour force (and whether illness [OA] is the reason), and any welfare payments they receive. In short, it gathers information about the nature of the productivity (labour productivity through number of hours worked) and a valuation of it (because individuals report their income and welfare payments).

Where there is capacity for more complicated modelling, it would be desirable to model additional long-term (and wider-reaching) productivity impacts occurring beyond the study (intervention) period. It is probably not possible to ask participants about their taxation directly – people tend not to remember and it is often not feasible to ask complicated questions in health/medical intervention studies – so an alternative approach would be to estimate this information from established income-welfare payments-taxation microsimulation models based on national population survey data using closely associated variables such as income, age, family type. It is also possible to develop a model of GDP impacts; however, this would be reliant on prevalence data relevant to the trial to be population represented and, from these data, using the hours of work and labour force participation.
variables as parameters in the GDP model. A GDP model requires a particular set of variables (see Intergenerational Reports especially IGR 2007).

Societal costs provide policymakers with the “end result”, including impacts on individuals, the government, and society.

Applying this framework to RCTs or other intervention studies, we note:

1. Large productivity gains are most likely to be seen for interventions preventing a condition or early intervention as a form of secondary prevention reducing the risk of disease progression (and thus reduce the likelihood of exit from the labour force).
2. Capturing productivity gains partly depends on the period of follow-up. The longer the follow-up, the more likely we will observe improvements in productivity due to the intervention (and thus no need to increase the sample size in order to observe an impact on productivity).
3. All the information listed above can be collected by survey in clinical trials and other intervention studies. Where there is no data collected, an estimate of productivity gains can be modelled by using a health outcome variable (such as SF36 or its subscales) which is associated with productivity in survey data (see example, Section 5). One advantage of other types of studies is that their recruitment criteria are often less restrictive and thus are more representative of the population of interest.
4. Productivity losses are commonly also incurred by carers, and surveys should be provided to carers to capture their impacts.
5. Where there is capacity to undertake modelling, it is recommended that additional productivity impacts (such as longer-term labour force participation, income, welfare payments) be projected beyond the study period.
6. Although it is possible to survey participants on their taxation, this is often not feasible so an alternative approach is to use strongly related variables (such as income, age, and family type) with national population survey data through synthetic matching in a microsimulation model.
7. It is also possible to develop a model of GDP impacts; however, this would be reliant on weight in the population trial to be population represented and, from this data, use hours of work and labour force participation variables as parameters in the GDP model. A GDP model requires specific inputs such as number of people in the working-age population, number of people employed.
8. The framework provides a description of data that could be collected within a clinical trial; however, the total impacts will represent those for the trial population only. If national estimates of productivity impacts are desired, then the trial data will need to be reweighted to be representative of the national population and, in particular, the incidence of the condition being treated (and eligible for the trial).
4.2. Methods for estimating Productive Life Years Gained (Lost) and related costs

In Australia, we have a good example of the use of national data and models to capture productivity impacts of illness based on microsimulation.\(^87\) We have previously developed a microsimulation model *Health&WealthMOD* to estimate the economic costs of early retirement due to illness for the older working age (i.e. 45-64 years) Australian population. The development of the model has been described in detail in Schofield et al (2011)\(^88\). The base population of *Health&WealthMOD* was based on the Australian Bureau of Statistics’ Surveys of Disability, Ageing and Carers (SDAC) data.\(^89,90\) The SDACs are nationally representative, large, Australian household surveys conducted by the ABS every 3 or so years. They provide the most comprehensive data about individuals in terms demographics, socioeconomic status, labour force, chronic conditions, and informal care needs and roles. But the SDACs include only limited economic data. More detailed economic information (such as income, welfare payments, income tax paid and the value of wealth assets such as superannuation) needs to be sourced elsewhere. This study has used outputs from other microsimulation model such as *STINMOD* developed for the impact of policy changes on the Australian population. *STINMOD* was developed (and is maintained) by the National Centre for Social and Economic Modelling (NATSEM, University of Canberra) ([www.natsem.canberra.edu.au](http://www.natsem.canberra.edu.au)) and is one of the principal microsimulation models of Australia’s income tax and transfer (welfare) system.\(^91\)

In the SDACs used for this study, respondents were asked to nominate their current labour force status as either:

1. Employed working full-time
2. Employed working part-time
3. Unemployed looking for full-time work
4. Unemployed looking for part-time work
5. Not in the labour force

Respondents who were not in the labour force were also asked to nominate the main reason for not working or looking for work. One of these options is ‘own ill-health or disability’. All respondents were asked whether they had long-term health conditions, and to nominate the type of main condition. The ABS then grouped the diseases based on ICD10 codes. Respondents who reported to be not in the labour force due to their “own ill-health or disability” and their main health condition to be “arthritis and related disorders” (ICD10 code M00-19), for example, were then considered to have ‘arthritis’ as the reason for them being out of the labour force.

We have used the output of *Health&WealthMOD* to estimate productive life years lost due to chronic conditions and the related costs such as lost incomes and savings for individuals and the lost
income tax and the increased welfare payments for the governments. The economic costs of lost productive life years lost for the 45 to 64 year old Australian population have been estimated for a number of chronic conditions such as arthritis, cardiovascular diseases, mental illness, diabetes and back pain.

4.3. Conclusion

The productivity measure proposed in this study is identified and could be measured using relevant questions in intervention studies or from national surveys conducted by the ABS (SDACs), and valued using outputs from microsimulation models commonly accessed by government departments in Australia. A choice will need to be made about how many questions about productivity can be included in an RCT or other intervention study when the entire set is considered to be too time consuming for study participants.

Our expectations of the new productivity measure proposed (PLYs) are that it will be included in intervention studies for all chronic conditions – just as standard questions about healthcare usage are included – and reported separately to cost-effectiveness measures (such as the Incremental Cost Effectiveness Ratio or ICER) to enhance decisions about resource allocation. We note that the framework for the productivity measure includes “yes” and “no” branches for labour force participation (see Figure 2) to help with counting the number of people in important groupings e.g. people with a chronic condition [OA] working full-time versus without a chronic condition [No OA] working full-time, people with a chronic condition who return to work due to an intervention (such as a pharmaceutical medicine) that improved their health versus people in work who did not have it (controls). It also assists in identifying relevant valuations e.g. for people in the work force, their main source of income is likely to be earnings (wages, salaries) whereas, for those not in the work force, it is likely to be a social security payment(s). It is self-reported changes in the labour force participation of people with a particular illness through an intervention that generates a measurable impact that subsequently can be valued for individuals (income), the government (taxes and welfare payments) and society (GDP). As discussed above, there are data sources currently available in Australia that can be used for the productivity measure proposed. The particular questions (branches in Figure 2) may also be added to clinical trials or other intervention studies, although the follow up period may be insufficient to obtain information on the timing and nature of the labour market changes. If the data collected is not extensive or sufficiently detailed at the individual level, the valuation step may need to be completed by bringing in data from other sources (e.g. ABS or outputs from a microsimulation model such as STINMOD).
5. **Productive life years gained and the associated economic gains as a result of OA intervention (Stage 3)**

The costs of OA on individuals are substantial, and Appendix 8 provides estimates of the costs for 4 different individuals (based on age, sex and education).

In this section, we applied the measure of productive life years (PLYs) gained, proposed in the previous section (Stage 2), to estimate the potential productivity benefits of an intervention to manage OA. As discussed earlier in the report, the productivity improvements have significant flow-on economic benefits both for individuals in terms of increases in disposable incomes and savings and for the government in terms of increases in income tax revenue and reductions in welfare payments. We therefore modelled the economic gains associated with the potential PLYs gained as a result of managing OA with an intervention.

A targeted literature review of productivity measures used in OA intervention and other studies in Stage 1 did not reveal any suitable randomised controlled trials or other studies, which have a direct measure of productive life years (PLYs) gained due to an OA intervention. We therefore modelled PLYs gained based on the association between improvements in quality of life, as a result of managing OA with an intervention, and labour force participation.

Our prior work on analysing the association between quality of life measures and labour force participation suggests that the Short Form-36 or SF-36 (specifically the bodily pain domain) score is one of the quality of life measures that has a strong association with labour force participation. Furthermore, there are nationally representative Australian household panel survey data publicly available to model the association between the SF-36 bodily domain score and labour force participation.

A more specific search of the literature to identify the best available study that has measured improvements in SF-36 bodily pain domain scores as a result of managing OA with a pharmaceutical intervention resulted in Rabenda et al (2005). This is a population based naturalistic prospective follow-up study of 783 patients aged 35 and over with OA in whom general practitioners (GPs) decided to start treatment with non-selective non-steroidal anti-inflammatory drugs (NSAIDs) for the symptomatic management of OA. The study measured health related quality of life (QoL) using two QoL instruments: the Medical Outcomes Study Short Form-36 Health Survey (SF-36) and the Western Ontario and McMaster Universities OA index (WOMAC) at baseline and after 3 months. Although the study is not a randomised controlled trial, it provides part of the information needed for modelling PLYs gained due to a pharmaceutical intervention i.e. improvements in SF-36 bodily pain domain scores due to an intervention to manage OA. Being a naturalistic study, it did not have a
control group and thus it is not possible to comment on if the improvement estimated in the study is over the placebo effect (effect in control group). However, the benefit of using results from such study is that sample selection is not too restrictive and thus is more likely to be represent the general population with OA.

We simulated the potential labour force gains and the associated economic gains of managing OA with NSAIDs, as a case study, to demonstrate how interventions to manage OA have the potential to provide economic benefits to both individuals and the government in addition to health benefits. Using the estimated number of Australians aged 15-64 years who were out of the labour force due to their OA in 2015, we estimated how many of them would have remained in the labour force if their OA was managed using NSAIDs.

Any medicines or other interventions clinically proven to be effective may have some productivity benefits. The method described in this section can be extended to other interventions and chronic conditions to estimate the productivity and economic benefits of managing or preventing a condition, provided that relevant outcome data are collected. We chose “the use of NSAIDs” as a case study since our extensive literature search did not reveal any other studies that had analysed a pharmaceutical intervention for the management of OA, which had the relevant information (such as SF-36 scores) required to model the productivity benefit and the associated economic benefit. Rabenda et al (2005)\textsuperscript{21} is the only study that we found, which analysed the effect of a pharmaceutical health intervention (NSAIDs) and had all the relevant data for modelling the potential labour force increase.

### 5.1. Data and methods

Published results from Rabenda et al (2005)\textsuperscript{21} and nationally representative Australian household surveys were used to estimate the potential labour force and related economic benefits associated with pain reduction in patients with OA by managing their OA with the use of NSAIDs. The top five most commonly used NSAIDs reported in the study was Piroxicam, Nimesulide, Meloxicam, Ibuprofen and Diclofenac.\textsuperscript{21} Rabenda et al (2005)\textsuperscript{21} reported an improvement in two health related QoL measures: the SF-36 and the WOMAC over a 3 month period. The SF-36 measures eight health concepts providing scores in eight health domains: physical functioning, role limitations-physical, bodily pain, social functioning, general health, role limitations-emotional, general mental health and...
Each domain score ranges from 0 to 100, with higher scores reflecting better health status. The SF-36 bodily pain domain is used to evaluate overall pain. We chose the SF-36 bodily pain domain score to analyse the association between quality of life and labour force participation because pain is reported to be one of the strongest predictors of labour force participation and is a key symptom of OA.

Rabenda et al (2005) reported that a comparison of SF-36 bodily pain domain scores between baseline and at 3 months after starting treatment showed a statistically significant improvement in SF-36 bodily pain domain scores in patients with OA over the 3 month period. Rabenda et al (2005) estimated that there was an improvement of 6.56 units in average SF-36 bodily pain domain score between baseline and at 3 months.

5.1.1. Association between SF-36 bodily pain scores and labour force participation

The association between the SF-36 bodily pain domain scores and the probability of being in the labour force were analysed using Wave 10 data of The Household Income and Labour Dynamics in Australia (HILDA) Survey which is a nationally representative Australian household panel survey. Analysis was restricted to the surveyed populations aged 20-64 years and those who had SF-36 bodily pain scores in the range between 0 and 70. This would exclude individuals in the HILDA Survey who had lesser or no pain (i.e. higher SF-36 bodily pain scores) and reduce the chance of contaminating the association between the SF-36 bodily pain scores and the probability of being in the labour force.

The association between the SF-36 bodily pain scores and labour force participation was estimated as the prevalence ratio (ratio of the probability of being in the labour force) associated with each unit increase in the SF-36 bodily pain score using a modified Poisson regression model with a robust error variance to estimate the prevalence ratios. Analyses were adjusted for the confounding effects of age and highest level of education, and the models were separately fitted for men and women.

5.1.2. Labour force impacts

We estimated the number of persons, who could have continued to stay in the labour force, among those not in the labour force due to their OA, as a result of pain reduction (for this case study, we chose an example of using NSAIDs). This was estimated as the potential increase in those who were in the labour force and have OA using the equation:

\[ \text{AddLF} = \text{Num}_{\text{int}} * \text{LF}_{\text{OA}} * \text{LF}_{\text{Growth}} \] (1)
where $Num_{int}$ is the estimated number of people not in labour force due to their OA who will start treatment for the management of their OA (for illustration, we used treatment with NSAIDs); $LF_{OA}$ is the estimated labour force participation of those with OA (a rate); and $LF_{Growth}$ is the estimated growth in the labour force participation rates associated with pain reduction as a result of interventions, which is estimated as

$$LF_{Growth} = \exp(Effect \cdot \Delta Pain) - 1$$

(2)

where $Effect$ is the estimated association between the labour force participation and SF-36 bodily pain domain scores, which is estimated as the change in the labour force participation rates for each unit change in SF-36 bodily pain domain score (which was estimated using the Wave 10 data of The HILDA Survey$^{100}$), such that $\exp Effect$ is a prevalence ratio of being in the labour force associated with each unit change in SF-36 bodily pain domain score; and $\Delta Pain$ is the estimated improvement in SF-36 bodily pain score for the patients undergoing an intervention to manage their OA (i.e. starting NSAIDs in this case) based on the published study by Rabenda et al (2005).$^{21}$

5.1.3. Economic benefits

For those individuals who are not in the labour force due to their OA, the potential increase in their incomes as a result of staying in the labour force by managing their OA were estimated as the difference in their incomes (when they are not in the labour force due to their OA) and their expected incomes if they remain in the labour force. The expected incomes were estimated using counterfactual simulation with Monte Carlo methods. For each of those individuals not in the labour force due to their OA, who could have continued to stay in the labour force by managing their OA, a counterfactual was selected at random with replacement from the pool of those in the labour force who had OA, who are in the same age group, same sex and have the same highest level of education as of those who could have continued to stay in the labour force by managing their OA. The mean of the difference in the incomes of those not in the labour force due to OA but who could have continued to stay in the labour force by managing their OA and their counterfactuals were estimated and reported as the expected average increase in incomes. We ran 1,000 simulations and reported the average of the 1,000 simulations as the result in this report.

Economic benefits were also estimated for the government in terms of the expected average increase in income tax payment and the average reduction in welfare payments for those who are not in the labour force due to their OA but who could have continued to stay in the labour force by starting NSAIDs. The economic benefits were estimated at the national level and are reported here as total increases in income and reduction in welfare expenditure. The benefits for individuals were estimated in terms of an increase in total accumulated income for all additional people in the labour
force and, for the government, in terms of total savings in welfare payments and additional income tax revenue.

We also estimated the impacts on gross domestic product (GDP) resulting from the additional people who could have continued to stay in the labour force by managing their OA. The impact on GDP was estimated on a pro-rata basis based on the published GDP data and the number of people in the labour force.

The labour force impacts and the associated economic benefits of using NSAIDs for the management of OA were estimated at the population level for Australians aged 15-64 years (the official working age population) who were out of the labour force due to their OA in 2015. The estimate of the potential labour force impacts was based on the assumption that all those not in the labour force due to their OA would take NSAIDs for the management of their OA. The estimated number of Australians aged 15-64 years who were not in the labour force due to their OA and economic data including incomes, income tax paid and welfare payments received are based on our microsimulation models Health&WealthMOD and Health&WealthMOD2030\(^7,8\) that we developed to analyse the economic impacts of early retirement due to chronic conditions. The development of these models are described in detail in Schofield et al (2011) and Schofield et al (2014).\(^7,8\) We used a modified version of our models, updated with more recent data to estimate the results for 2015 and to include all working age individuals (15-64 years), for this analysis.

5.2. Results

5.2.1. Not in the labour force due to OA

The prevalence data and the labour force information were not available for those with OA in the main data sources for our microsimulation models, the ABS Surveys of Disability, Ageing and Carers, which only have information for those with overall arthritis. Thus, we estimated the labour force information for those with OA based on a previous study which estimated the age-specific proportion of OA among overall arthritis.\(^10\) Based on the assumption that all individual labour force groups would have the same proportion of OA among those with arthritis, it was estimated that there would be 9,910 men and 19,422 women not in the labour force due to their OA in 2015 (Table 3). The estimated numbers of people in other labour force categories by age group and gender are listed in Appendix 9.
Table 3. Estimated numbers of people not in the labour force due to OA, aged 15-64 years, 2015

<table>
<thead>
<tr>
<th>Age group</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>15-44 years</td>
<td>409</td>
<td>986</td>
</tr>
<tr>
<td>45-64 years</td>
<td>9,501</td>
<td>18,436</td>
</tr>
<tr>
<td>Total</td>
<td>9,910</td>
<td>19,422</td>
</tr>
</tbody>
</table>

5.2.2. Association between SF-36 bodily pain domain scores and labour force participation

The association between SF-36 bodily pain domain scores and the labour force participation rates was estimated using Wave 10 data of the HILDA Survey. For those aged between 20-64 years and with bodily pain domain scores between 0 and 70, the estimated prevalence ratios of being in the labour force associated with each unit increase in SF-36 bodily pain domain score was 1.0112 for men and 1.0103 for women.

The labour force participation rates of those with OA were 63.1% for men and 52.5% for women in 2015. The Rabenda et al (2005) estimates of the improvement in SF-36 bodily pain domain score as a result of starting treatment with NSAIDs for the management of OA from baseline to 3 months follow-up (a 6.56 units improvement) were then used as an input into the equation (2) (page 35) to estimate the increase in labour force participation. It was found that there would be an estimated increase of 7.6% points in the labour force participation rate of men and 6.9% points for women with OA by follow-up.

5.2.3. Impacts on labour force participation

We estimated how many people not in the labour force due to their OA would have continued to stay in the labour force if they had started treatment with NSAIDs for the management of OA. If we assume that all those not in the labour force due to their OA (Table 3) had started managing their OA with NSAIDs, there would be an additional 472 men and 705 women aged 15-64 years in the labour force, who otherwise would be out of the labour force due to their OA in 2015 (Table 4).
Table 4. Estimated number of people who would have avoided being out of the labour force due to their OA as a result of starting treatment with NSAIDs for the management of their OA, aged 15-64 years, 2015

<table>
<thead>
<tr>
<th>Age group</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>15-44 years</td>
<td>19</td>
<td>36</td>
</tr>
<tr>
<td>45-64 years</td>
<td>453</td>
<td>669</td>
</tr>
<tr>
<td>Total</td>
<td>472</td>
<td>705</td>
</tr>
</tbody>
</table>

5.2.4. Economic impacts

The potential increase in the labour force participation of those with OA as a result of reduced pain by managing their OA with NSAIDs has follow-on economic benefits both to individuals in terms of an increase in their disposable income and for the government in terms of savings in welfare payments and additional income tax revenue.

Men aged 45-64 years, who were not in the labour force due to their OA, would have an estimated increase of $51,888.00 in their annual income if they continued to stay in the labour force by managing their OA. They would have, on average, paid an additional income tax of $14,144.00 and received $14,524.00 less in welfare payment annually (Table 5). For women of this age group, the estimated increase in their annual income was $25,834.00 if they stayed in the labour force by managing their OA. They would have, on average, paid an additional income tax of $6,395.00 and received $13,940.00 less in welfare payment annually (Table 5).

Table 5. Estimated individual economic benefits of being in the labour force because of reduced pain as a result of managing OA, aged 15-64 years, 2015

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>15-44 years</td>
<td>45-64 years</td>
</tr>
<tr>
<td>Increase in annual incomes</td>
<td>40,614.00</td>
<td>51,888.00</td>
</tr>
<tr>
<td>Increase in annual income tax payment</td>
<td>12,077.00</td>
<td>14,114.00</td>
</tr>
<tr>
<td>Reduction in annual welfare payment</td>
<td>12,680.00</td>
<td>12,524.00</td>
</tr>
</tbody>
</table>

It was estimated that there would be a total increase of $42.7 million in income per year associated with people avoiding being out of the labour force due to their OA if they had started NSAIDs early on for the management of their OA (Table 6). Similarly, it was estimated that there would be a saving of about $15.6 million per year in welfare payments and an increase of $11.2 million in
income tax revenue because of those individuals, who would otherwise be out of the labour force due to their OA, staying in the labour force as a result of pain reduction (Table 6).

As a result of the estimated 1,177 additional individuals aged 15-64 years in the labour force who would have avoided being out of the labour force due to their OA by starting treatment with NSAIDs, there would be an estimated increase in Australian GDP by approximately 164 million dollars in 2015. In our previous work we measured the impacts of chronic diseases on GDP using Treasury’s GDP equation, and this equation was used here to measure the impacts of the intervention on GDP.

Table 6. Estimated total economic benefits at national level from increased labour force associated with reduced pain as a result of managing their OA, aged 15-64 years (in thousands, in 2015 AU$)

<table>
<thead>
<tr>
<th>Benefits</th>
<th>AU$ (’000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Increase in annual incomes</td>
<td>42,702</td>
</tr>
<tr>
<td>Total Increase in annual income tax payment</td>
<td>11,164</td>
</tr>
<tr>
<td>Total Reduction in annual welfare payment</td>
<td>15,597</td>
</tr>
<tr>
<td>Total Increase in GDP</td>
<td>164,000</td>
</tr>
</tbody>
</table>

5.3. Conclusions

People with OA who experience less bodily pain are more likely to be in the labour force. Effective interventions that reduce bodily pain may thus help patients to remain in the workforce longer. It has been estimated that there would be an additional 1,177 people in the workforce associated with reduced pain as a result of managing their OA by starting early NSAIDs treatment. The cumulative economic benefit of this increased labour force participation would be an estimated increase of $42.7 million in their annual income in 2015. As a result of this estimated increased labour force, Australian GDP in 2015 would have increased by approximately 164 million dollars. The benefit of increased labour force participation associated with pain reduction by managing OA extends beyond individuals, and provides economic benefits to government as well.

6. Conclusions and Future Directions

The targeted search of OA studies revealed that relatively few studies include productivity as an outcome measure. The majority of intervention studies used a single, stand-alone question such as labour force status, hours of work, sick leave and time to return to work after surgery; whereas the cost of illness studies tended to use more comprehensive (and usually validated/tested for suitability with OA population) productivity questionnaires/instruments such as the WALs, WIS, WLQ, WPAI, and EWPS. Many of these measures are reported in natural units (such as hours or paid work,
number of sick days, or are a scale) and thus without a conversion into dollars, no meaningful comparisons can be made across different OA populations, interventions or competing calls on scarce resources.

The productivity costs missing from the literature are: (1) a wider range of productivity losses for individuals, such as lost income from lost labour force participation (or reduced paid working hours) and the costs of informal care; (2) costs to government, such as extra welfare payments and lost taxation revenue due to lost productivity of people with OA; and (3) societal costs, such as the dollar value of lost GDP due to OA through its impact on employment of people with ill-health.

Although RCTs generate the highest quality of evidence, there are practical limitations to trying to measure productivity within an RCT, such as the short time frame of trials (sometimes as short as 4-12 weeks) which limits the potential to capture longer term productivity impacts.

The productivity measure proposed in this study – whole productive life years gained in a given year – is based on a series of questions designed to elicit information about the labour force participation of individuals with and without ill-health, their income, receipt of welfare payments, estimates of current income (earnings plus welfare payments), and taxes. Consideration should be given to how to incorporate the new measure of PLYs into intervention studies, assuming there is a desire to maintain internal consistency and validity of the new measure by including all of the questions proposed in the framework (Figure 2). In other words, if there is a desire to identify, measure and value PLYs in an intervention study then all questions (branches) should be included in questionnaires for the trial; however, if there are limitations in terms of what is feasible to collect, then consideration should be given to how many of these questions can be included and thus how the outcome measure will be reported, for example, only a count of the number of people in the workforce before and after having the intervention but don’t express it in monetary terms (no income data collected).

Given the lack of productivity measures in published studies, and the short time frame of most intervention studies, other alternatives to measuring productivity in trials (such as our modelled approach in Stage 3) or whether other types of intervention studies with longer follow-up are suitable need to be considered. There is therefore a need for further data to show the significant association between the intervention and quality of life measure (e.g. SF-36), and then the quality of life measure and productivity (e.g. labour force participation). We addressed this issue in the modelling in Stage 3, by using HILDA data to examine the link between quality of life outcomes (SF-36 bodily pain) and labour force participation. Our previous work demonstrated that the SF-36 bodily pain component was strongly related to having OA symptoms and, through that link,
differences in productivity. There is a need to collect data on other commonly used pain scales and associations with productivity, preferably longitudinally.

The results shown in this study indicate that there is a need to further examine and integrate productivity based measures in clinical trials and studies for health interventions. By doing so, there will be an opportunity to better understand the broader economic impacts, and assist in accurately understanding the link between better health outcomes and broader economic contributions that result from pharmaceutical interventions.
References:


140. Lerner D, Rogers WH, and Chang H, *Scoring the work limitations questionnaire (WLQ) and the WLQ index for estimating work productivity loss*. Technical Report.


Appendix 1: Inclusion and exclusion criteria of the published studies

The inclusion criteria consisted of:

1. Study participants: men and/or women aged 18 years or older who were living in the community.

2. Study outcome measure: productivity of people with OA or changes in productivity due to interventions to manage OA (expressed in natural or monetary units, can be a primary or secondary outcome, can be expressed as a cost or benefit); see the search in section 3.3. for specific productivity measures.

3. Study design: studies of the costs of OA; studies on the direct and indirect costs of OA; productivity costs of OA from any perspective, including that of the patient, the health system or government, third party payers (such as health insurance companies), or society; the disease was explicitly stated as OA in the methods (but could have been self-reported doctor diagnosed or clinically confirmed OA); studies using a prevalence-based approach (cost of illness); OA intervention studies (including RCTs, retrospective cohort studies, prospective cohort studies); studies conducted in developed countries (defined, for the purpose of this review, as countries in the Organisation for Economic Co-operation and Development, OECD); any intervention (i.e. pharmaceutical, medical – surgery, diet and lifestyle) in a study that included productivity as an outcome could be considered.

4. Years considered: studies published between 1 January 2010 to 11 June 2015.

5. Language: published in English.

6. Publication type: original research only.

The exclusion criteria consisted of:

1. Study participants: study population was children only, or a group with special conditions not representative of the general population; or the studied subjects were non-human (e.g. mouse) and/or non-living such as cadavers, and human specimens (e.g. genes).

2. Study design: costs of OA did not include, or have a specific breakdown for, productivity losses; or did not explicitly state that the productivity costs were attributable to OA (e.g. “musculoskeletal conditions” instead of “osteoarthritis”) i.e. not an outcome measure.

3. Years considered: studies published before 1 January 2010.

4. Publication types: duplicate articles or non-original studies (such as comments, editorials or letters, and review articles); conference proceedings and other reports where a detailed study could not be retrieved were excluded.
Appendix 2: Search strategy and list of keywords

The search was run by one reviewer (SW). The search strategy combined one search for OA studies with a second on studies of productivity. In the selected databases the following keywords, search strategy and limits were included:

Keywords:

1. osteoarthritis
2. labour
3. retire
4. financial
5. societal
6. time off work
7. work
8. carer
9. caregiver
10. presenteeism
11. absenteeism
12. sick leave
13. employment
14. job
15. workers compensation
16. child care
17. income
18. pension
19. productivity

We combined the keywords in our search as follows:

a) All the productivity keywords (2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19)
b) 1 and (a)
c) Limit to (English language and humans and year=“2010-Current”)

All search results were assigned into a Microsoft Excel spreadsheet under key headings (such as author (year), type of study, study population, productivity measure(s), results), and the full-texts undergoing further checks for inclusion were compiled.
## Appendix 3: Summary of Interventions for OA (25 studies)

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Population</th>
<th>Study Design</th>
<th>Intervention</th>
<th>Measure of Productivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allen et al. 2014&lt;sup&gt;61&lt;/sup&gt;</td>
<td>USA</td>
<td>Knee/hip OA. Age 60-66. N=280</td>
<td>Randomised Controlled Trial</td>
<td>Examined clinic variation in recruitment metrics, patient characteristics and treatment use in an RCT of osteoarthritis. Clinic variations: 1) Patient Intervention Only, 2) Provider Intervention Only, 3) Patient Intervention plus Provider Intervention and, 4) Usual Care (Control).</td>
<td>Self-reported whether the household participants belonged to was below country (household) poverty line; and income status (low, middle, high) was reported for those with OA who have intervention and who do not. Note: Low income was defined as “just meet basic expenses” or “don’t have enough to meet basic expenses.” (Also self-reported general health on a 5-point scale, from excellent to poor.)</td>
</tr>
<tr>
<td>Baldwin et al. 2012&lt;sup&gt;63&lt;/sup&gt;</td>
<td>USA</td>
<td>RA or OA. Mean age 51 years. N=89. 87% women, 38% with RA, 62% with OA.</td>
<td>Randomised Controlled Trial (prospective study)</td>
<td>Workplace ergonomic intervention</td>
<td>Role scores (i.e. impact of arthritis on employment); Job Satisfaction Survey (JSS).</td>
</tr>
<tr>
<td>Bedair et al. 2014&lt;sup&gt;53&lt;/sup&gt;</td>
<td>USA</td>
<td>Knee OA. Hypothetical (patient) model N=1.</td>
<td>Markov model for the economic benefits to society of total knee arthroplasty in younger patients..</td>
<td>Total Knee Arthroplasty</td>
<td>A Markov state-transition decision model was assembled to compare the overall average cost over 30 years of total knee arthroplasty with the average 30 year cost of non-operative treatment for a 50-year-old patient with end-stage OA. Earned income, lost wages and direct medical costs related to non-operative treatment and to total knee arthroplasty, including revisions and complications, were considered.</td>
</tr>
<tr>
<td>Study</td>
<td>Country</td>
<td>Population</td>
<td>Study Design</td>
<td>Intervention</td>
<td>Measure of Productivity</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---------</td>
<td>------------</td>
<td>---------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Cowie et al. 2013&lt;sup&gt;103&lt;/sup&gt;</td>
<td>UK</td>
<td>OA. Age &lt;65. N=238 (OA).</td>
<td>Retrospective study</td>
<td>total hip replacement</td>
<td>Indirect costs (productivity losses): lost income due to work disability, absenteeism, decreased work hours, job change, and unemployment which were estimated by an equivalent total number of days not worked in one year multiplied by the daily wage.</td>
</tr>
<tr>
<td>Daruwalla et al. 2013&lt;sup&gt;104&lt;/sup&gt;</td>
<td>UK</td>
<td>Wrist OA. Age 26-71. N=10</td>
<td>Prospective study (no control group, small sample)</td>
<td>Pyrocarbon Adaptive Proximal Scaphoid Implant (APSI) – an alternative treatment option for scaphoid nonunion advanced collapse (SNAC) and radioscaphoid OA.</td>
<td>Time to return to work post-surgery.</td>
</tr>
<tr>
<td>de Carvalho et al. 2014&lt;sup&gt;105&lt;/sup&gt;</td>
<td>Brazil</td>
<td>Knee OA. Age 21-65. N=26</td>
<td>Retrospective study (small sample size, relatively short average follow-up period – 48 months)</td>
<td>Distal femur osteotomy for the treatment of lateral compartment knee OA.</td>
<td>Interviews based on a questionnaire relating to the level of satisfaction with the surgery; functional and physical activities were scored according to published ratings. Productivity measure: Resumption of normal work duties after surgery at the individual's preoperative functional level without limitation or decline in performance.</td>
</tr>
<tr>
<td>Faschingbauer et al. 2015&lt;sup&gt;106&lt;/sup&gt;</td>
<td>Germany</td>
<td>Knee OA. Age &gt;30. N=51 (OA=37).</td>
<td>Retrospective study (relatively short follow-up period)</td>
<td>High tibial osteotomy</td>
<td>Time to return to work post-surgery.</td>
</tr>
<tr>
<td>Foote et al.</td>
<td>UK</td>
<td>Knee OA. Age &lt;60.</td>
<td>Retrospective study</td>
<td>Knee arthroplasty</td>
<td>Time and details on return to work</td>
</tr>
<tr>
<td>Study</td>
<td>Country</td>
<td>Population</td>
<td>Study Design</td>
<td>Intervention</td>
<td>Measure of Productivity</td>
</tr>
<tr>
<td>-------</td>
<td>---------</td>
<td>------------</td>
<td>--------------</td>
<td>--------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>2010&lt;sup&gt;106&lt;/sup&gt;</td>
<td>Canada</td>
<td>N=109.</td>
<td></td>
<td></td>
<td>post-surgery i.e. patients provided information on their employment status pre-surgery, details of their return to work post-surgery and their subjective opinion as to what effect their knee operation had had on their ability to return to work. This opinion was graded from 1 (significant worsening) to 5 (significant improvement). All patients returning to work were also asked to rate their post-surgery occupation as ‘less physically intense’, ‘same physical intensity’ or ‘more physically intense’ as compared to their pre-surgery occupation.</td>
</tr>
<tr>
<td>Gignac et al. 2015&lt;sup&gt;64&lt;/sup&gt;</td>
<td>Canada</td>
<td>OA or IA (e.g. rheumatoid arthritis, psoriatic arthritis), or both OA and IA. Age ≥25. N=219.</td>
<td>Longitudinal study</td>
<td>Workplace accommodations (flexible hours, modified schedules, special equipment/adaptations, work-at-home arrangements) and benefits (extended health, short-term leaves) for people with arthritis.</td>
<td>12-item Workplace Activity Limitations Scale (WALS) was used to measure arthritis-related activity limitations at work. Additionally, information was ascertained about job disruptions, productivity losses, absenteeism, and reduced work hours (via telephone screening and self-administered questionnaires).</td>
</tr>
<tr>
<td>Hungerer et al. 2011&lt;sup&gt;107&lt;/sup&gt;</td>
<td>Germany</td>
<td>Foot OA. Age &gt;31. N=108.</td>
<td>Retrospective case control (no randomisation or control group)</td>
<td>Primary arthrodesis and secondary arthrodesis after revision surgery.</td>
<td>Productivity measure: regaining ability to work after fusion of the subtalar joint and revision surgery (if necessary). (Main outcome was health status measured by SF-36 (36-item Short Form Health Survey)).</td>
</tr>
<tr>
<td>Husby et al. 2010&lt;sup&gt;108&lt;/sup&gt;</td>
<td>Norway</td>
<td>Hip OA. Age &lt;60. N=22.</td>
<td>Randomisation of eligible patients in a clinic to either maximal strength training intervention in the early postoperative phase after OA-</td>
<td>Productivity measure: Work efficiency (but based on time on treadmill, oxygen levels).</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Country</td>
<td>Population</td>
<td>Study Design</td>
<td>Intervention</td>
<td>Measure of Productivity</td>
</tr>
<tr>
<td>------------------</td>
<td>------------------------------</td>
<td>----------------------------------------------------------------------------</td>
<td>--------------------------------------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Klein et al. 2013</td>
<td>Germany</td>
<td>Trapeziometacarpal joint (TMJ) OA. Age 38-78. N=59.</td>
<td>Cross-sectional, single-centre study design</td>
<td>Modified Epping procedure for trapeziometacarpal OA.</td>
<td>Time and details on return to work post-surgery e.g. whether have full or partial (manual) work performance.</td>
</tr>
<tr>
<td>Lerner et al. 2012</td>
<td>Australia (4 sites), Canada (15 sites), New Zealand (6 sites), USA (87 sites)</td>
<td>Knee OA. Age 40-65. N=758.</td>
<td>Randomised Controlled Trial (randomised, double-blind, active- and placebo-controlled, parallelarm, multicentre, phase III trial).</td>
<td>Tapentadol extended release versus oxycodone controlled release for OApain.</td>
<td>Work Limitations Questionnaire (WLQ) – a self-report questionnaire that measures the frequency with which health problems prevent the performance of job tasks and reduces productivity at work. The WLQ provided the Productivity Loss Score for this study. Also imputed the change or difference in WLQ Productivity Loss Scores based on health variables representing the RCT's endpoints (such as pain intensity). Valuation (dollars): Expressed the improvement in productivity loss from baseline to week 12 of the maintenance period as a productivity cost-savings, by multiplying the imputed change in at-work productivity within each group by an assumed annual income of $100,000 per person.</td>
</tr>
<tr>
<td>Marra et al. 2014</td>
<td>Canada</td>
<td>Knee OA. N=73.</td>
<td>Economic evaluation of a cluster Randomised Controlled Trial</td>
<td>Pharmacist-initiated multidisciplinary strategy to manage knee OA.</td>
<td>Questions on work (work time, absences, and unpaid work) and other societal costs (expenses from knee pain or OA, household help, and informal care) were asked at 3 and 6 months</td>
</tr>
</tbody>
</table>
after randomisation. Productivity losses were captured using the human capital approach, where average age- and sex-adjusted earnings were utilised to determine the daily loss of income when reported by participants. Markov health state-transition decision model was developed for the treatment of end-stage knee OA. Indirect costs included lost earnings and productivity, caregiver costs (expressed in dollars). Patients were assumed to recover 80% of the indirect costs of end-stage knee OA after successful primary total knee arthroplasty. The model incorporated indirect costs into the model, and ICERs were produced with this information included.

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Diagnosis</th>
<th>Methodology</th>
<th>Treatment</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nunley et al. 2011&lt;sup&gt;110&lt;/sup&gt;</td>
<td>USA</td>
<td>Knee OA. Age 18-60. N=661.</td>
<td>Retrospective, multicentre study</td>
<td>Total knee arthroplasty</td>
<td>Retrospective, multicentre study</td>
</tr>
<tr>
<td>Pauchard et al. 2014&lt;sup&gt;111&lt;/sup&gt;</td>
<td>France</td>
<td>Wrist OA. Age=mean 50. N=30</td>
<td>Retrospective (no control group, small sample size)</td>
<td>Dorsal locking plates versus staples in four-corner fusion.</td>
<td>Retrospective (no control group, small sample size)</td>
</tr>
</tbody>
</table>


Nunley et al. 2011<sup>110</sup> USA Knee OA. Age 18-60. N=661. Retrospective, multicentre study Total knee arthroplasty

Pauchard et al. 2014<sup>111</sup> France Wrist OA. Age=mean 50. N=30 Retrospective (no control group, small sample size) Dorsal locking plates versus staples in four-corner fusion.
<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Country</th>
<th>Condition</th>
<th>Age</th>
<th>Sample Size</th>
<th>Study Design</th>
<th>Intervention</th>
<th>Outcome Measures</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pinto et al. 2013</td>
<td>New Zealand</td>
<td>Hip, Knee or both OA.</td>
<td>Adult population, N=209.</td>
<td>Randomised Controlled Trial with Economic Evaluation.</td>
<td>Manual therapy, exercise therapy, or both, in addition to usual care for hip or knee OA.</td>
<td>Questions about loss of hours (productivity measure) and information about costs borne by participants' friends or family (out-of-pocket costs, transportation costs, lost earnings, and informal care). Instrument used: Osteoarthritis Costs and Consequences Questionnaire (OCCQ) – freely accessible online.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ruiz et al. 2013</td>
<td>USA</td>
<td>Knee OA.</td>
<td>Hypothetical N=1.</td>
<td>Markov model with three primary components: (1) quality of life; (2) direct medical costs for the total knee arthroplasty, complications of surgery, and revision arthroplasty; and (3) indirect costs involving employment status, earnings, time missed from work (or absenteeism), and disability payments.</td>
<td>Total knee arthroplasty compared with nonsurgical treatment for knee OA.</td>
<td>Markov model results: Employment status, earnings, time missed from work (or absenteeism), and disability payments.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sankar et al. 2013</td>
<td>Canada</td>
<td>OA.</td>
<td>Age 18-85. N=310.</td>
<td>Sub-sample of a prospective longitudinal cohort</td>
<td>Total hip or knee replacement.</td>
<td>Primary outcome was return to work based on the patient’s response to the question: ‘which of the following [responses] best describes your current employment status’ on the first report of a return to work. Secondary outcomes from the Workplace Activity Limitations Scale (WALS).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schroter et al. 2013</td>
<td>Germany</td>
<td>Knee OA.</td>
<td>Age=mean 47. N=32.</td>
<td>Retrospective study (variations in length of Open wedge High Tibial Osteotomy.</td>
<td>Time and details on return to work post-surgery i.e. duration of incapacity to</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Country</td>
<td>Condition</td>
<td>Study Design</td>
<td>Intervention/Procedure</td>
<td>Follow-up</td>
<td>Control Group</td>
<td>Analysis</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>---------------</td>
<td>----------------</td>
<td>-------------------------------</td>
<td>-------------------------------------------------------------</td>
<td>--------------------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Svendsen et al. 2012&lt;sup&gt;62&lt;/sup&gt;</td>
<td>Denmark</td>
<td>OA of the AC joint. Age 18-63.</td>
<td>Nation-wide cohort study</td>
<td>Surgery for rotator cuff-related shoulder disorders, frozen shoulder, and OA.</td>
<td>follow-up, no control)</td>
<td>calculated the risk of post-operative permanent work disability which was defined as making a transition to disability pension (leaving the labour market due to permanent work disability) or flex-job within two years after the date of admission for surgery. Linked with data from the Danish National Register on Public Transfer Payments (DREAM, a Danish acronym for ‘the Register Based Evaluation of Marginalization’).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>van Bergen et al. 2013&lt;sup&gt;113&lt;/sup&gt;</td>
<td>Netherlands</td>
<td>Ankle OA. Age 13-52. N=50.</td>
<td>Prospective cohort study (relatively long follow-up – 8-20 years, no control)</td>
<td>Arthroscopic debridement and bone marrow stimulation.</td>
<td>time to return to work post-surgery.</td>
<td>calculated the risk of post-operative permanent work disability which was defined as making a transition to disability pension (leaving the labour market due to permanent work disability) or flex-job within two years after the date of admission for surgery. Linked with data from the Danish National Register on Public Transfer Payments (DREAM, a Danish acronym for ‘the Register Based Evaluation of Marginalization’).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verkleij et al. 2010&lt;sup&gt;57&lt;/sup&gt;</td>
<td>Netherlands</td>
<td>Knee OA. Age ≥45 years.</td>
<td>Randomised open-label trial (no control group – only comparing to another medication, study incomplete)</td>
<td>Diclofenac versus acetaminophen for primary care patients with knee OA.</td>
<td>time to return to work post-surgery.</td>
<td>calculated the risk of post-operative permanent work disability which was defined as making a transition to disability pension (leaving the labour market due to permanent work disability) or flex-job within two years after the date of admission for surgery. Linked with data from the Danish National Register on Public Transfer Payments (DREAM, a Danish acronym for ‘the Register Based Evaluation of Marginalization’).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Waimann et al. 2014&lt;sup&gt;48&lt;/sup&gt;</td>
<td>USA</td>
<td>Knee OA. Age=mean 65. N=212.</td>
<td>Economic evaluation (CEA) of total knee replacement. Data from a prospective cohort study.</td>
<td>Total knee replacement in OA patients versus hypothetical non-surgery strategy.</td>
<td>time to return to work post-surgery.</td>
<td>calculated the risk of post-operative permanent work disability which was defined as making a transition to disability pension (leaving the labour market due to permanent work disability) or flex-job within two years after the date of admission for surgery. Linked with data from the Danish National Register on Public Transfer Payments (DREAM, a Danish acronym for ‘the Register Based Evaluation of Marginalization’).</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Appendix 4: Summary of cost of illness studies (37 studies; 1 policy document as last entry)

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Population</th>
<th>Study Design</th>
<th>Objective</th>
<th>Measure of Productivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ackerman et al. 2013&lt;sup&gt;114&lt;/sup&gt;</td>
<td>Australia</td>
<td>hip/knee joint disease (arthritis or OA) Age 56-74. N=237 (N=1157?)</td>
<td>Cross sectional (nation-wide)</td>
<td>Examined health-related quality of life (HRQoL), work status, and healthcare utilisation (and costs) according to severity of hip and knee joint disease (OA) (measured using WOMAC) and compared to asymptomatic group.</td>
<td>Self-reported employment information covering paid and unpaid work status, premature exit from the workforce because of hip or knee arthritis or OA, and changes to work because of hip or knee arthritis or OA. (Also used the Assessment of Quality of Life (AQoL) instrument which is the Australian generic measure of HRQoL; and self-reported demographic and health service information).</td>
</tr>
<tr>
<td>Agaliotis et al. 2013&lt;sup&gt;115&lt;/sup&gt;</td>
<td>Australia</td>
<td>Knee OA. Age 45-75. N=289</td>
<td>A cohort nested within a randomised controlled trial. Longitudinal data because of the 12 month follow-up. Great accuracy in evaluation (used a day-by-day evaluation of lost work productivity specifically related to knee problems for 7 days every 2 months over a 12-month follow-up period) and thus likely to have produced more reliable estimates of presenteeism.</td>
<td>To determine risk factors associated with reduced work productivity among people with chronic knee pain.</td>
<td>Absenteeism (number of days off work due to knee problems) and presenteeism (% of reduced work productivity while at work due to knee problems) (productivity questions) were derived from the Work Productivity and Activity Impairment Questionnaire (WPAI) via face-to-face assessments. Also assessed changes to work e.g. changing occupations, working hours, retired, lost job.</td>
</tr>
<tr>
<td>Beaton et al. 2010&lt;sup&gt;75&lt;/sup&gt;</td>
<td>Canada</td>
<td>RA or OA. Age 19-65. N=250 in total; OA=130, RA=120. No breakdown of</td>
<td>Longitudinal design (12-months?)</td>
<td>Evaluated the relative strengths and weaknesses of 5 measures for quantifying health-related at-work productivity losses in patients with RA or OA</td>
<td>Workplace Activity Limitations Scale (WALS), 6-item Stanford Presenteeism Scale (SPS-6), Endicott Work Productivity Scale (EWPS), RA</td>
</tr>
<tr>
<td>Study</td>
<td>Country</td>
<td>Population</td>
<td>Study Design</td>
<td>Objective</td>
<td>Measure of Productivity</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------------</td>
<td>-------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Berger et al. 2011</td>
<td>USA</td>
<td>OA. Age 18 or older. N=2399 (OA)</td>
<td>Cross sectional study. Private sector employees in the Health and Productivity Management (HPM) Database with evidence of OA (based on hospital visits) in 2007</td>
<td>To estimate direct and indirect economic costs among private sector employees with OA.</td>
<td>Absenteeism; short-term disability; worker’s compensation were the indirect costs (measured and valued in dollars; compared to those without OA). All information was on the Medstat MarketScan Health and Productivity Management (HPM) Database i.e. administrative data.</td>
</tr>
<tr>
<td>Bieleman et al. 2010</td>
<td>Netherlands</td>
<td>Hip/knee OA. Age 45-61. N=275</td>
<td>Cohort study i.e. cohort hip and cohort knee (CHECK) versus healthy ageing workers (cross sectional)</td>
<td>(1) To compare self-reported health status and functional capacity of subjects with early OA of hip and/or knee to reference data of healthy working subjects, and (2) to assess whether this capacity is sufficient to meet physical job demands.</td>
<td>Short-Form 36 Health Survey used to assess health status. Six tests of the WorkWell Systems Functional Capacity Evaluation used to assess productivity = lifting low/overhead, carrying, bending, repeated side reaching.</td>
</tr>
<tr>
<td>Bieleman et al. 2010</td>
<td>Netherlands</td>
<td>Hip/knee OA. Age 45-61. N=970</td>
<td>Cross sectional study design, using data from the Cohort Hip and Cohort Knee (CHECK) study</td>
<td>To examine the work participation of Dutch people with early OA in hips or knees and compare this with data from the American Osteoarthritis Initiative (OAI) cohort.</td>
<td>Economic Aspects in Rheumatoid Arthritis questionnaire (validated) used to assess various aspects of work participation (present or last job, work hours, working history, present working status, and sick leave). Participants in paid employment were asked about their present condition and whether they had adapted or would like to adapt their work via tasks, hours, and work place. Subjects without paid work</td>
</tr>
<tr>
<td>Study</td>
<td>Country</td>
<td>Population</td>
<td>Study Design</td>
<td>Objective</td>
<td>Measure of Productivity</td>
</tr>
<tr>
<td>------------------------------</td>
<td>--------------------------</td>
<td>-----------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Bieleman et al. 2013</td>
<td>Netherlands</td>
<td>Early OA of hips or knee. Age=45-65.</td>
<td>Prospective cohort study – 2 years of follow-up in the Cohort Hip and Cohort Knee (CHECK).</td>
<td>To identify prognostic factors for the 2-year course of work participation in early OA of hips or knees.</td>
<td>were asked for reasons for not having a job. Used the Economic Aspects in Rheumatoid Arthritis questionnaire (validated) to assess several aspects of work participation: present or last job, work hours, working history, present working status, and sick leave, and physical work demands. Labour force participation was defined as having a paid job for 8 hours or more per week. Participants with paid employment were asked if they had been on sick leave, and if so, if this was because of hip/knee complaints or other health reasons. Another question was whether they had adapted or would like to adapt their work (hours, tasks, workplace). Subjects without paid work were asked for reasons for not having a job. (SR health was measured using SF36).</td>
</tr>
<tr>
<td>Bushmakin et al. 2011</td>
<td>France, Germany, Italy, Spain and UK</td>
<td>OA. Age &gt;50. N=1739.</td>
<td>Cross sectional, observational study</td>
<td>To determine how patient-rated osteoarthritis (OA) severity correlated with other patient-reported and clinical outcomes in the European clinical setting.</td>
<td>Work Productivity and Activity Impairment (WPAI). The first question is on employment status, and the remaining five, referenced to the past 7 days, pertain to hours worked; hours missed from work because of OA; impact of OA on productivity (rating scale from 0% no effect to 10% completely</td>
</tr>
</tbody>
</table>
prevented from working); and degree OA affected regular activities (rating scale 0% no effect to 10% completely prevented daily activities). Responses to the questions were used to calculate absenteeism (i.e. the percentage of work time missed due to OA), presenteeism (i.e. the percentage of impairment on the job due to OA), percentage of overall work impairment due to OA; and percentage of activity impairment due to OA. These percentages were then used to estimate the annual costs of lost productivity in US dollars based on average annual wages estimated by the World Bank for 2008.

From the seeding statement where participants were asked to form statements about how OA affects their life (and people around them), two relevant impacts were identified: financial and productivity impacts.

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>OA Location</th>
<th>Methodology</th>
<th>Study Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Busija et al. 2013&lt;sup&gt;120&lt;/sup&gt;</td>
<td>Australia</td>
<td>Hip/knee, hand, generalised OA. Age 51-80. N=26.</td>
<td>Integrated theoretical model (based on qualitative and quantitative data)</td>
<td>To develop a conceptual model to describe the burden of OA in individuals with this condition and on the broader community.</td>
</tr>
<tr>
<td>Conaghan et al. 2015&lt;sup&gt;74&lt;/sup&gt;</td>
<td>UK</td>
<td>OA. Age 19-91. N=2001</td>
<td>Cross sectional study, using the Arthritis Care OA Nation 2012 survey.</td>
<td>To understand the impact of OA on individuals and to explore current treatment strategies.</td>
</tr>
<tr>
<td>Study</td>
<td>Country</td>
<td>OA Type</td>
<td>Study Design</td>
<td>Objective</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>-----------</td>
<td>------------------------------</td>
<td>-----------------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Cunha-Miranda et al. 2015</td>
<td>Portugal</td>
<td>OA. Age 45-87. N=103</td>
<td>Observational, cross-sectional (the MOVES study)</td>
<td>To estimate the prevalence of self-reported OA and quality of life in Portuguese adults aged 45 years or older.</td>
</tr>
<tr>
<td>DiBonaventura et al. 2011</td>
<td>USA</td>
<td>OA (specific site of OA pain unknown). Age ≥20. N=2173</td>
<td>Cohort study, using the National Health and Wellness Survey (large sample size and population level analysis based on a weighted assessment to reflect the demographic composition of the US population)</td>
<td>To evaluate the impact of OA pain on healthcare resource utilisation, productivity and costs in employed individuals.</td>
</tr>
<tr>
<td>DiBonaventura et al. 2012</td>
<td>USA</td>
<td>OA (severity provided but not joint affected) vs non OA. Age ≥20. N=39772</td>
<td>Cross sectional analysis, using National Health and Wellness Survey 2009.</td>
<td>To evaluate the impact of self-rated OA severity on quality of life, healthcare resource utilisation, productivity and costs in an employed population with OA relative to</td>
</tr>
<tr>
<td>Study</td>
<td>Country</td>
<td>Diagnosis</td>
<td>Study Design/Study Description</td>
<td>Aims</td>
</tr>
<tr>
<td>------------------------------</td>
<td>-------------</td>
<td>-----------</td>
<td>-----------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>Fearon et al. 2014</td>
<td>Australia</td>
<td>Hip OA.</td>
<td>Observational case control study</td>
<td>To evaluate work participation, health related quality of life,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Age=mean 62 years. N=20</td>
<td></td>
<td>pain and function, and activity participation, in people with a broad spectrum of Greater Trochanteric Pain Syndrome (GTPS) severity.</td>
</tr>
<tr>
<td>Gignac et al. 2011</td>
<td>Canada</td>
<td>OA (57%) or IA.</td>
<td>Longitudinal (18 months; working population only)</td>
<td>To examine the type, degree, and episodic nature of arthritis-related work place activity limitations and the consistency of the relationship of activity limitations to job modifications and work place outcomes.</td>
</tr>
<tr>
<td>Gignac et al. 2014</td>
<td>Canada</td>
<td>OA or IA. Age&gt;40 years.</td>
<td>Cross-sectional study (convenience sample of employed individuals)</td>
<td>To examine men and women’s perceptions of inter-role balance/imbalance in work, arthritis, and personal roles and its association with demographic, health and employment factors, including job stress, career satisfaction, job disruptions, absenteeism and perceived productivity losses.</td>
</tr>
<tr>
<td>Study</td>
<td>Country</td>
<td>Sample Description</td>
<td>Study Design</td>
<td>Objective</td>
</tr>
<tr>
<td>-------</td>
<td>---------</td>
<td>---------------------</td>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Datta and Larsen 2010 (^{25})</td>
<td>Denmark</td>
<td>Various conditions, including self-reported OA i.e. ‘doctor told you that you have – or within the last year have had OA’). N=approximately 3,000 person-wave observations.</td>
<td>Longitudinal study (older workers drawn from a Danish panel survey from 1997 and 2002, merged to longitudinal register data)</td>
<td>To investigate whether, even for narrowly defined health measures, a divergence exists in the impacts of health on retirement between self-reported health and objective physician-reported health.</td>
</tr>
<tr>
<td>Gignac et al. 2013 (^{125})</td>
<td>Canada</td>
<td>Knee, hip, and/or groin pain or other joints affected by OA. Age &gt;40. N=177 (OA)</td>
<td>Cross-sectional study</td>
<td>Examined role salience (i.e. importance), role limitations, and role satisfaction among middle- and older-aged adults with and without OA and its relationship to depression, stress, role conflict, healthcare utilisation, and coping behaviours.</td>
</tr>
<tr>
<td>Study</td>
<td>Country</td>
<td>Condition</td>
<td>Age/Region</td>
<td>Study Design</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>--------------------------</td>
<td>-----------</td>
<td>----------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>Hermans et al. 2012[84]</td>
<td>Netherlands</td>
<td>Knee OA.</td>
<td>Age=18-65 years. N=117.</td>
<td>Randomised controlled trial (all subjects in employment; BMI measurement taken too)</td>
</tr>
<tr>
<td>Hubertsson et al. 2013[71]</td>
<td>Sweden</td>
<td>Knee OA.</td>
<td>Age &gt;46. N=15345.</td>
<td>Population-based cohort study</td>
</tr>
<tr>
<td>Kingsbury et al. 2014[126]</td>
<td>France, Germany, Italy, Spain</td>
<td>OA. Age &gt;18. N=3750.</td>
<td></td>
<td>Cross-sectional study</td>
</tr>
</tbody>
</table>

Kotlarz et al. 2010 USA OA. Age 18-64. National survey (cross-sectional) To quantify the effects of OA on the cost of absenteeism from work. Medical Expenditure Panel Survey (MEPS). Absenteeism costs were valued in US dollars (and reported for men and women).

Pit et al. 2010 Australia OA. Age 45-64. N=3160. Cross-sectional study To examine which health problems are associated with early retirement due to ill health among Australians aged 45–64 years. Participants who reported being retired were asked why they had retired, and were allowed multiple responses including: ‘reached usual retirement age’, ‘to care for family member/friend’, ‘made redundant’, ‘lifestyle reasons’, ‘ill-health’ (and specific condition e.g. OA), ‘could not find a job’, or ‘other’.

Rolfson et al. 2012 Sweden Hip OA. Age 26-95. N=2635. Cohort study To estimate direct and indirect costs incurred by patients with hip disease and eligible for total hip arthroplasty. Patients listed for THA answered a questionnaire concerning the use of resources and loss of productivity because of their hip disease (asked whether the patient was currently in work, retired, unemployed, on sick leave, or on disability pension). (Informal carer costs and costs for productivity loss because of sick leave and disability pension were measured and valued in SEK.)
<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>N or Study Design</th>
<th>Study Methods</th>
<th>Study Objectives</th>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roppone et al. 2011</td>
<td>Finland</td>
<td>N=24043.</td>
<td>Cohort study</td>
<td>To investigate the incidence of and risk factors for disability pensions due to any musculoskeletal diagnoses and specifically due to OA.</td>
<td>Receipt of disability pension (main outcome of the study) due to OA.</td>
</tr>
<tr>
<td>Saastamoinen et al. 2012</td>
<td>Finland</td>
<td>OA. N=6258.</td>
<td>A prospective cohort study</td>
<td>To examine the association of pain with subsequent disability retirement due to all causes as well as musculoskeletal diseases, mental disorders, and a heterogeneous group of other diseases.</td>
<td>Linked Helsinki Health Study baseline questionnaire and national pension register data used to ascertain disability retirement due to pain caused by OA.</td>
</tr>
<tr>
<td>Sadosky et al. 2010</td>
<td>USA</td>
<td>OA (mainly knee and hip). Age&gt;50. N=714.</td>
<td>Cross sectional observational study</td>
<td>To determine how patient-reported OA severity correlates with patient-reported outcomes including pain, function and work productivity.</td>
<td>WPAI. Valuation provided = cost of work impairment for employed persons (cost per patient per year resulting from lost productivity). For patients aged &lt;65 years, loss of productivity in terms of working days lost was estimated.</td>
</tr>
<tr>
<td>Piscitelli et al. 2012</td>
<td>Italy</td>
<td>Hip &amp; Knee OA. Age &gt;25.</td>
<td>Longitudinal study (5 years of national hospitalisation records)</td>
<td>To assess the socioeconomic burden of total joint arthroplasties (TJAs) performed for symptomatic hip and knee OA in the Italian population.</td>
<td></td>
</tr>
<tr>
<td>Sayre et al. 2010</td>
<td>Canada</td>
<td>OA (knee, hip, hand, foot, lower back or neck). Age &gt;19. N=688.</td>
<td>Cross sectional study</td>
<td>To investigate the effects of the site of OA (knee, hip, hand, foot, lower back or neck) on employment reduction due to OA (ERoa).</td>
<td>Current and past employment status, and Employment reduction due to OA (ERoa) collected via mailed out questionnaire.</td>
</tr>
<tr>
<td>Tang et al. 2010</td>
<td>Canada</td>
<td>OA or RA. Age 54 + 6.7. N=130 with OA.</td>
<td>Longitudinal study (data collected at baseline, 3 months, 6 months and 12 months)</td>
<td>To evaluate the ability of the Work Instability Scale for Rheumatoid Arthritis (RA-WIS) to predict arthritis-related work transitions within a 12-month period (primary outcome) and perceived decreases in work capacity within same time period (secondary outcome).</td>
<td>Used the Work Instability Scale for Rheumatoid Arthritis (RA-WIS) to (1) predict arthritis-related work transitions within 12 months (any, change in job/occupation, disability leave of absence, reduction in work hours); and (2)</td>
</tr>
</tbody>
</table>
within 12 months, perceived decreases in work capacity (decreased ability to work, increased illness interference, decreased work productivity). For the secondary outcome, 3 self-reported global indicators to examine perceived changes in work capacity 1) change in ability scale, 2) change in degree of illness based on illness Intrusiveness Ratings Scale, and 3) change in work productivity scale.

Tang et al. 2011\(^{58}\) Canada OA. N=130. Cohort To examine disease-related differential item functioning (DIF) in the RA-WIS.

Van Der Burg et al. 2014\(^{71}\) Netherlands OA. Age 18-65. N=12140. Prospective cohort study among working individuals To determine the risk of sick leave and work disability in relation to rheumatic diseases and cardiovascular comorbidities among working individuals. Number of days and frequency of sick leave; work disability (problems at work due to health issues, perceived physical demands of the job); receipt of a full or partial work disability pension over 10 year period.
<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Age Range</th>
<th>Study Design</th>
<th>Objective</th>
<th>Methodology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wilkie et al. 2014 (a)</td>
<td>England</td>
<td>OA 50-65</td>
<td>Population-based prospective cohort study</td>
<td>To describe the prevalence of expected work limitations (EWL) prior to future retirement age in OA consulters, and the associated health, sociodemographic, and workplace factors.</td>
<td>&quot;Do you think joint pain will limit your ability to work before you reach 69 years old?&quot; (will limit or stop me/ don’t know/ won’t limit) - a single question.</td>
</tr>
<tr>
<td>Wilkie et al. 2014 (b)</td>
<td>England</td>
<td>OA &gt;50</td>
<td>Population-based prospective cohort study</td>
<td>The objective of this study was to describe the extent of premature work loss (PWL) in OA consulters across a 6-year observation period, and associated factors.</td>
<td>PWL was defined as moving from employment to retirement prior to state retirement age or transition from employment to being off work due to health or unemployment between baseline and 3 years or between 3 and 6 years using the participant’s self-report of employment status [employed (including self-employed), off work due to illness, unemployed, homemaker, retired].</td>
</tr>
<tr>
<td>Wilkie et al. 2015</td>
<td>England</td>
<td>Age &lt;65</td>
<td>Population-based prospective Cohort study – primary care consulters for OA</td>
<td>To provide an in-depth examination of how pain leads to the onset of work productivity loss and identify potential new intervention opportunities.</td>
<td>Work productivity was measured using a single item from the Medical Outcomes Study Short Form-36. Employment status (and early retirement, homemaker) was also ascertained over a 3 year period.</td>
</tr>
<tr>
<td>Zhang et al. 2010</td>
<td>Canada</td>
<td>18-65</td>
<td>Prospective cohort</td>
<td>To estimate and compare lost work hours attributable to presenteeism,</td>
<td>Lost hours due to presenteeism (reduced productivity while</td>
</tr>
</tbody>
</table>
defined as reduced productivity while working, in individuals with OA or RA), according to 4 instruments. Working) was estimated using several questionnaires: (a) Health and Labor Questionnaire (HLQ), (b) Work Limitations Questionnaire (WLQ), (c) the World Health Organization’s Health and Work Performance Questionnaire (HPQ), and (d) Work Productivity and Activity Impairment Questionnaire (WPAI). Instruments were administered at baseline, 3, 6 and 12 months. Estimated lost hours and associated monetary valuation of productivity loss due to presenteeism among workers with OA (dollars).

Tang et al. 2014 (POLICY) Canada OA Policy document Worker productivity outcome measures: OMERACT filter evidence and agenda for future research. Discusses different productivity measures. OMERACT group are described, and discuss WPAI, WPS-RA, QQ, HPQ, and WAI.
Appendix 5: Study characteristics

The main characteristics of the 63 studies and the productivity measures used in OA studies are presented in Appendix 1 (intervention studies) and Appendix 2 (cost of illness). Together, these studies were published in 2015 (n=6, 10%), 2014 (n=14, 22%), 2013 (n=14 (includes one policy document)), 2012 (n=6, 10%), 2011 (n=8, 13%), and 2010 (n=13, 21%). They reported the productivity losses of OA in 17 countries, with 12 from Canada, 11 from the USA, and 6 from Australia. Twenty-two studies (9 intervention and 13 cost of illness) included one type of productivity measure (i.e. a single question/instrument), 36 included multiple productivity measures (15 intervention and 21 cost of illness), and 9 included measures related to productivity (such as quality of life using SF-36) but did not link these results to productivity (3 intervention and 6 cost of illness). In terms of the productivity measures, 11 studies (intervention) measured time to return to work post-surgery either on its own or in addition to other measures, 22 studies (3 intervention, 18 cost of illness and 1 policy) measured work absences (e.g. absenteeism, sick leave) through either single questions or the Work Productivity and Activity Impairment (WPAI) questionnaire (Reilly et al (1993), and see http://www.reillyassociates.net/WPAI_SHP.html), 14 (4 intervention and 10 cost of illness) measured workplace activity limitations due to OA using the Workplace Activity Limitations Scale (WALs) (Gignac et al (2004)), 3 (cost of illness) studies measured productivity through the Work Instability Scale for Rheumatoid Arthritis (RA-WIS) (Gilworth et al (2003)), and 2 (intervention studies) measured job satisfaction either post surgery or when working with OA. Seven studies (3 intervention and 4 cost of illness) assessed quality of life or used markov health state-transition decision models for tracing changes in productive lives. Of these, 3 used the SF-36 to assess quality of life with OA/after surgery, 1 used the AQoL, and 3 used Markov state-transition decision models. Note that none of the studies linked the individual with OA's quality of life assessment to their productivity – which had implications for the modelling in Stage 3. Fifty-six studies (23 intervention and 33 cost of illness) produced productivity outcomes for people with OA only, while the remaining studies producing these outcomes for individuals with OA or another form of arthritis (OA or RA, or OA or IA). The majority of the studies assessed productivity benefits (losses) for the entire working-age population (e.g. 18-60 years), as opposed a subset of the population.
Appendix 6: List of instruments

The intervention studies have only used questionnaires (1)-(4) below; the cost of illness studies have used all of the questionnaires listed below, except (5).

1. **Workplace Activity Limitations Scale (WALS)**

   The WALS consists of 12 items intended to measure arthritis-related workplace activity limitations\(^{76,79,135}\). The instrument items assess the degree of difficulty people with arthritis have with various job-related tasks affecting upper/lower limb function, and their difficulties with travelling, scheduling, concentration, and the pace of the work. Respondents are asked about the difficulties they encounter “in general” or “typically” but without a defined period for which to recall occurrences. Response options are based on a four-point Likert scale moving from “no difficulty (a score of 0)” to “not able to do (a score of 3)” plus two other options: “difficulty unrelated to arthritis” and “not applicable” (both are assigned a score of 0). The WALS has a score range of 0 to 36.

   **Limitations of the instrument:**
   - No capacity to estimate the monetary value of the productivity losses
   - Excludes people not working
   - Does not include obvious major costs (e.g. lost earnings for workers, income/profit for employers)
   - Does not include other major costs (e.g. welfare, tax, worker’s compensation)
   - The instrument asks: “As a result of your arthritis, how much difficulty do you have concentrating or keeping your mind on your work?”\(^{135,136}\) but this is not a direct measure of presenteeism (i.e. not quantifying the degree of difficulty a person has at work due to their illness in terms of lost work hours or lost earnings/income).

2. **Work Limitations Questionnaire (WLQ)**

   The WLQ-25 was developed to measure the impact of chronic diseases and treatments on work performance\(^{59,137}\) and has been validated in both OA\(^{138}\) and RA\(^{139}\) populations. Questionnaire items ask about the proportion of time (moving from “none of the time” to “all the time”) over the last 2 weeks, in which the respondent experienced some difficulty in four subscales (each is scored from 0 to 100): time management, physical demands, mental interpersonal, and output demands. The content and format of the WLQ-25 were originally derived from focus groups and cognitive interviewing of workers with gastrointestinal, psychiatric or respiratory disorders, or epilepsy.\(^{59}\). Lerner et al. (2003)\(^{140}\) provides a formula (WLQ Index) that converts WLQ-25 subscale scores into an
estimate of percentage of productivity lost compared to healthy controls (i.e. without OA, RA), which could be used to measure the economic impacts.

Limitations of the instrument:

- Consists of people working only (i.e. excludes those out of the workforce)
- Does not include obvious major costs (e.g. lost earnings for workers, lost income for employers)
- Does not include other major costs (e.g. welfare, tax, worker’s compensation)

Strengths:

- The instrument asks about the amount of health-related productivity loss experienced in a fixed time period (presenteeism) and thus a direct productivity measure is provided.
- It has the capacity to estimate the monetary value of the productivity loss.

3. PROductivity and DISease Questionnaire (PRODISQ)

The Productivity and Disease Questionnaire was developed to measure productivity costs due to knee symptoms. The questionnaire covers various aspects on the relationship between productivity and health, including the number of knee-related absences from work (sick leave) in the last 3 months. It includes the measurement of lost productivity due to knee symptoms while at work using the quality and quantity method.

Limitations of the instrument:

- Questions are for people working only
- Does not include obvious major costs (e.g. lost income)
- Does not include other major costs (e.g. welfare, tax, worker’s compensation)

Strengths:

- The instrument asks about the amount of absences due to ill-health in the last 3 months which could be used to quantify the productivity losses.

4. Osteoarthritis Costs and Consequences Questionnaire (OCC-Q) (used in Pinto et al (2013))

The (OCC-Q) consists of questions about the costs and consequences of rheumatic diseases (hip and knee OA) incurred by those with the diseases aged 40-85 years, their friends and families (e.g. carers). The OCC-Q has been validated for use in the New Zealand OA population. Respondents are asked to recall their visits to health professionals, use of public and private hospitals, medications, transport costs, aids and adjustments that were needed, and the use of community services over the last 3 months. Information about costs incurred by respondents’ friends or family (such as out-of-pocket costs, transport costs, lost earnings, and informal caregiving) were also
collected. With regard to productivity losses related to arthritis, respondents are asked (a) whether they are in paid work, and if not, if this is because of their arthritis; and (b) how their work has been affected by their arthritis in the last month (i.e. how many days off work, reduced hours, restrictions at work). For those in paid work, there are also questions about the hours they worked and hourly pay rate or annual salary.

**Main limitation of the instrument:**

- Questions about productivity are for those who are working (i.e. excludes those who are not working). Everyone is asked about carer costs.

**Strengths:**

- The instrument asks about a number of productivity losses over a fixed time period and current income, which could be used to quantify (monetarise) the losses.

5. Work Productivity and Activity Impairment Questionnaire (WPAI)

The **WPAI Specific Health Problem V2.0 (WPAI:SHP)** consists of the following questions about the effect of a person’s health problem on their work capacity and to perform common tasks. It begins with asking people whether or not they are currently employed (Yes/No) and then about how their health problem affected their work in the last seven days, including the number of hours missed because of problems associated with their health problem, how the problem affected their productivity while working. In terms of productivity limitations, respondents are asked to:

> Think about days you were limited in the amount or kind of work you could do, days you accomplished less than you would like, or days you could not do your work as carefully as usual.

And they are asked to circle a number from 0 (health problem had no effect on their work) to 10 (health problem completely prevented them from working) to indicate how much the problem affected their productivity.

**Limitations of the instrument:**

- Excludes people out of the workforce
- Does not include obvious major costs (e.g. lost earnings for workers, lost income for employers)
- Does not include other major costs (e.g. welfare, tax, worker’s compensation)

**Strengths:**

- The WPAI directly measures productivity: current employment status, number of hours missed due to a health problem (absenteeism), the extent to which their health problem affected productivity in hours (presenteeism). Consequently, it has the capacity to estimate the monetary value of the productivity losses.
6. Stanford Presenteeism Scale (SPS-6)

The Stanford Presenteeism Scale (SPS-6) is a short derivative of the original SPS-32. It was designed for non-specific clinical and employee populations. The SPS-6 measures a worker’s ability to focus on and complete work tasks regardless of health-related distractions, in the last month. Scale content of the original version (SPS-32) was formulated through a thorough literature review and the developers’ experiences. The SPS-6 consists of six questions with options being on a five-point Likert scale moving from “strongly agree” (a score of 1) to “strongly disagree” (a score of 5) to provide a scale score from 6 to 30 (with 30 indicating “most presenteeism”).

Limitations of the instrument:

- No capacity to estimate the monetary value of the productivity losses
- Excludes people not working
- Does not include obvious major costs (e.g. lost income)
- Does not include other major costs (e.g. welfare, tax, worker’s compensation, carer costs)
- The instrument includes questions about “how hard” people find different aspects of work (e.g. completing hard tasks) but this is not a direct productivity measurement.

7. Endicott Work Productivity Scale (EWPS)

The Endicott Work Productivity Scale (EWPS) was formulated to quantify the frequency of attitudes and behaviours that are not productive in the last week, and is relevant for a range of diseases and occupations. It consists of four domains: attendance (absenteeism), quality of the work performed, the worker’s performance capacity, and important personal factors within 25 items. Five response options are offered for each item: “never” (a score of 0) to “almost always” (a score of 4), and the scale is scored from 0 to 100 with “lowest productivity” having a score of 100.

Limitations of the instrument:

- No capacity to estimate the monetary value of the productivity losses
- Excludes people not working – an important subpopulation
- Does not include obvious major costs (e.g. lost income)
- Does not include other major costs (e.g. welfare, tax, worker’s compensation)
- The instrument includes several questions relating to productivity, but people self-rate how they have performed in these areas in the last week (and thus no quantifiable direct productivity measurement).

8. Work Instability Scale for Rheumatoid Arthritis (RA-WIS)

The RA-WIS was developed for the RA population to assess the degree of incongruity between the functional abilities of the worker and the demands of their job, and to identify workers needing workplace adjustments in order to sustain their employment. The scale items were initially
generated from interviews with workers with early diagnosed RA. Respondents are asked to respond according to their personal experiences “at the moment”. The scale consists of 23 items where “yes” has a score of 1, and “no” has a score of 0. A total score of < 10 signifies “low work instability”, 10-17 signifies “moderate work instability”, and >17 signifies “high work instability”. Higher work instability has been shown to be predictive of transitions in and out of work among people with arthritis.

Limitations of the instrument:

- No capacity to estimate the monetary value of productivity losses
- Excludes people who are not working
- Does not include obvious major costs (e.g. lost earnings for workers, lost income for employers)
- Does not include other major costs (e.g. welfare, tax, worker’s compensation)
- A range of specific work-related issues are covered (e.g. ways of controlling arthritis symptoms, time management, difficulties completing tasks, distress caused by concerns about being able to maintain employment in the future) that could signify functional disability or an inability to undertake all the demands of the job; but these are not direct measures of productivity losses (i.e. not quantifying how the instability converts into productivity losses due to illness as lost work hours or lost earnings/income).


The Economic Aspects in Rheumatoid Arthritis questionnaire enquires about several aspects of work participation: present or last job, work hours, work history, present work status, sick leave, and physical demands of the job.\textsuperscript{67} Respondents in paid work are asked if they have been on sick leave, and if so, if this was because of hip/knee complaints or other health reasons. Another relevant question is whether they have altered or would like to alter their work due to their illness. Respondents not in paid work are asked to provide reasons for not having a job. Household work is also measured (and valued) through this questionnaire.

Main limitation of the instrument:

- Includes household productivity losses, defined as the (value of) household tasks that have to be carried out by formal (a paid housekeeper) or informal carers (friends, family), if the person was unable to perform these tasks because of their illness (OA). But it does not include other major costs (e.g. welfare, tax, worker’s compensation).

Strengths:

- The instrument has the capacity to estimate the monetary value of productivity losses for those in paid work (such as cost of sick leave days due to illness) and unpaid work (household productivity).
• It enables obvious major costs (e.g. lost income for workers and cost of sick leave for employers) to be calculated.

10. Medical Expenditure Panel Survey (MEPS) (used in Koltarz et al (2010)\textsuperscript{83})

The Medical Expenditure Panel Survey (MEPS) is a set of large-scale surveys of individuals and families, health care providers, and employers across the USA. The MEPS began in 1996 and data is collected every year. There are two major components to the survey: the Household Component and the Insurance Component. The former collects data from a sample of individuals and families across the country, drawn from a nationally representative subsample of households that participated in the previous year’s National Health Interview Survey (conducted by the National Center for Health Statistics). During the household interviews, MEPS collects detailed information for each person in the household on the following: demographic characteristics, health conditions, health status, use of medical services, charges and source of payments, access to care, satisfaction with care, health insurance coverage, income, and employment (see http://meps.ahrq.gov/mepsweb/about_meps/survey_back.jsp.).

Main limitation of the instrument:

- Does not include other major costs (e.g. welfare, tax, worker’s compensation)

Strengths:

- The MEPS includes people who are working and not working.
- From the MEPS, it is possible to ascertain the number of lost (paid) work days due to illness and a valuation of these using the income data.

11. Helsinki Health Study Baseline Questionnaire (used in Saatamoinen et al (2012)\textsuperscript{70})

The purpose of collecting ongoing information on the Helsinki Health Study cohort is to facilitate longitudinal studies on the social and work related factors of health and well-being using self-reported and independent register data. The target population was employees in Helsinki, Finland. Baseline data for the cohort were derived from questionnaires conducted in 2000, 2001 and 2002 among employees reaching 40, 45, 50, 55 or 60 years in each year. At baseline there were 8,960 respondents (80% women, and a 67% response rate). Age-relevant health examination data were also collected. A follow up survey was conducted in 2007, generating 7,332 respondents and a boost in the response rate to 83% response rate. A range of health measures were collected such as self-rated health, mental health conditions, functioning, pain, and major diseases. Social determinants include: socio-demographics, socio-economic measures, working conditions, social support, and how they balance work and family commitments. Further register linkages include sickness absences, hospitalisations, prescribed drugs, and retirement.\textsuperscript{146}
Main limitation of the instrument:

- Does not include other major costs (e.g. welfare, tax, worker’s compensation)

Strengths:

- There is the capacity to estimate the monetary value of the productivity losses because there is information on sickness absence (lost work days) and disability retirement due to main (diagnosed) illness (includes musculoskeletal diseases), as well as income data.

12. Health and Labour Questionnaire (HLQ) (used in Zhang et al (2010)\textsuperscript{78})

The HLQ was designed to collect data on the connections between illness, treatment, and work performance.\textsuperscript{147,148} The instrument can provide estimates of productivity losses (costs) in both paid and unpaid work. Reduced productivity at paid work is quantified by a single question asking how many extra hours individuals would have to work to catch up on tasks they were unable to complete in normal working hours due to health problems in the last 2 weeks. There are also questions about how work was reorganised to take account of health problems (e.g. working at a slower pace, colleagues taking on some of their tasks), income, sick leave, pension and early retirement.

Strengths of the instrument:

- The HLQ includes people who performed paid work, unpaid work (e.g. family members doing household tasks), and were unable to perform paid work due to their illness and other reasons (i.e. important subpopulations).
- The HLQ permits the estimation of productivity losses: an estimate of on-the-job production losses based on the mean hourly wage (income).
- The HLQ permits the estimation of changes in paid work (e.g. household work completed by family members of the person with the illness).

13. World Health Organization’s Health and Work Performance Questionnaire (HPQ) (used in Zhang et al (2010)\textsuperscript{78})

The HPQ is intended to assess the workplace costs of health-related reduced performance in jobs, sickness absences, and injuries from accidents at work.\textsuperscript{149,150} Respondents are asked to rate their overall performance on the days they worked during the last 7 days on a scale of 0 (a total lack of performance) to 10 (no lack of performance).\textsuperscript{149-151} If ideal work performance is set at 100%, then presenteeism can be measured as reduced work performance corresponding to a percentage below 100%; and the implied loss in hours of work can be estimated by multiplying this percentage by the number of hours worked. There are two versions of the HPQ: one for employees and another for clinical trials. The recall period is either the last 7 or 28 days of work, and short versions of the questionnaire (and scoring algorithms) are available (http://www.hcp.med.harvard.edu/hpq/info.php).
Limitations of the instrument:

- Excludes people who are not working
- Does not include obvious major costs (e.g. lost earnings for workers, lost income for employers)
- Does not include other major costs (e.g. welfare, tax, worker’s compensation)
- A scale is used to (self) assess presenteeism – not a direct measure of productivity losses that could be used by agencies interested in productivity such as Treasury.

Strengths:

- There is capacity to estimate the monetary value of the productivity losses because there is information on the (usual) hours of work and the amount of work time lost due to illness (self-rated) (work section) which could be used with the information on how much a person is paid per hour (wage) and annual salary (before tax) (demographics section).

In terms of the review undertaken for this report, one study (Wilkie et al (2015)\(^\text{131}\)) used a single question from the Medical Outcomes Study Short Form-36 to assess productivity; however, this form is normally used to derive quality of life. Other instruments used in the cost of illness studies were: the Workwell Systems Functional Capacity Evaluation, which assesses how OA impacts on the physical demands of the job (used in Bieleman et al (2010a)\(^\text{116}\)); the Social Role Participation Questionnaire (SRPQ), which assesses role satisfaction (used in Gignac et al (2013)\(^\text{125}\)). The policy document (i.e. Tang et al (2014)\(^\text{132}\)) and review by Zhang et al (2010)\(^\text{78}\) compare productivity outcomes from several instruments, including the WPAI, HLQ, WLQ, HPQ.
Appendix 7: Scaffolding of the proposed productivity measure

The starting point for the proposed productivity measure is:

**Are you in paid work (employment)?**

- **If Yes,**
  - What are your usual paid work hours per week?
    - If you are working part-time, what is your main reason for this?
      Options would include ill-health (including the condition relevant to the study e.g. osteoarthritis), illness severity (or pain severity if the intervention is designed to treat arthritis), other health conditions that prevent me from working.
  - What is your current hourly wage or annual salary?
  - Do you receive any social security payments? Common wording for this type of question includes “Are you receiving any Centrelink support or income payments?” or “Are you receiving any government income or support payments?” For example, the ABS’ Survey of Disability, Ageing and Carers 2003 and 2009 includes questions on receipt of welfare payments, as well as lists of specific payments (e.g. Aged Pension, Disability Support Pension, Carer’s Payment, Carer’s Allowance, Newstart). \(^{89,90}\)
  - Have you taken a lower paid job due to your illness [osteoarthritis]?
    - If Yes, what was the size of the pay cut? (Specify the new hourly wage rate or annual salary)
  - Have you taken any time off due to your illness [osteoarthritis] in the last X (e.g. last month) [and duration of the study]? 
    - How many hours did you take off in total?
    - How was the cost of these hours covered?
      - Paid sick leave, unpaid sick or other leave, workers comp, other forms of insurance, other (please specify).

- **If No,**
  - Are you seeking paid work (employment)?
    - If No, what is the main reason for this? Options to include: ill-health (including specifically the condition of interest in the study), illness severity, other health conditions prevented me from working (and nominate the conditions).
  - How long has it been since you last worked?
  - What was your previous annual income in your last paid job?
  - What was your previous annual income before you developed your illness [osteoarthritis]?
  - Are you receiving any social security payments?
    - If Yes,
- What type of payments are you receiving? (e.g. Aged Pension, Newstart Allowance, Disability Support Pension, Sickness Allowance, Youth Allowance, Carer Payment, Carer Allowance).
- How much are your total social security payments per fortnight?
Appendix 8: Cost of not being in the labour force due to OA for individuals in 2015

<table>
<thead>
<tr>
<th>Individual person characteristics</th>
<th>Average Total weekly income</th>
<th>Average Tax paid</th>
<th>Average Transfer payment received</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male, aged 45-49 years, no university qualification, <em>not in the labour force due to OA</em></td>
<td>$169.00</td>
<td>$0.00</td>
<td>$152.00</td>
</tr>
<tr>
<td>Male, aged 45-49 years, no university qualification, <em>in the labour force and no OA</em></td>
<td>$1,432.00</td>
<td>$310.00</td>
<td>$0.00</td>
</tr>
<tr>
<td>Female, aged 45-49 years, no university qualification, <em>not in the labour force due to OA</em></td>
<td>$464.00</td>
<td>$0.00</td>
<td>$437.00</td>
</tr>
<tr>
<td>Male, aged 45-49 years, no university qualification, <em>in the labour force and no OA</em></td>
<td>$970.00</td>
<td>$121.00</td>
<td>$0.00</td>
</tr>
</tbody>
</table>

Note: “In the labour force” means could be “full-time employed” or “part-time employed” or “unemployed”. 
### Appendix 9: Labour force status of those with OA in 2015

<table>
<thead>
<tr>
<th>Gender</th>
<th>Age group</th>
<th>Estimated number of people</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Employed full-time with OA (i.e. anyone with OA)</td>
</tr>
<tr>
<td>Male</td>
<td>15-24 years</td>
<td>1,554</td>
</tr>
<tr>
<td>Male</td>
<td>25-34 years</td>
<td>5,010</td>
</tr>
<tr>
<td>Male</td>
<td>35-44 years</td>
<td>12,977</td>
</tr>
<tr>
<td>Male</td>
<td>45-54 years</td>
<td>34,214</td>
</tr>
<tr>
<td>Male</td>
<td>55-64 years</td>
<td>53,425</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>107,180</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender</th>
<th>Age group</th>
<th>Estimated number of people</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>15-24 years</td>
<td>325</td>
</tr>
<tr>
<td>Male</td>
<td>25-34 years</td>
<td>2,652</td>
</tr>
<tr>
<td>Male</td>
<td>35-44 years</td>
<td>8,889</td>
</tr>
<tr>
<td>Male</td>
<td>45-54 years</td>
<td>26,110</td>
</tr>
<tr>
<td>Male</td>
<td>55-64 years</td>
<td>40,148</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>78,124</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender</th>
<th>Age group</th>
<th>Estimated number of people</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>15-24 years</td>
<td>140</td>
</tr>
<tr>
<td>Male</td>
<td>25-34 years</td>
<td>988</td>
</tr>
<tr>
<td>Male</td>
<td>35-44 years</td>
<td>1,865</td>
</tr>
<tr>
<td>Male</td>
<td>45-54 years</td>
<td>4,102</td>
</tr>
<tr>
<td>Male</td>
<td>55-64 years</td>
<td>12,291</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>19,386</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender</th>
<th>Age group</th>
<th>Estimated number of people</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>15-24 years</td>
<td>107</td>
</tr>
<tr>
<td>Male</td>
<td>25-34 years</td>
<td>1,012</td>
</tr>
<tr>
<td>Male</td>
<td>35-44 years</td>
<td>8,371</td>
</tr>
<tr>
<td>Male</td>
<td>45-54 years</td>
<td>27,130</td>
</tr>
<tr>
<td>Male</td>
<td>55-64 years</td>
<td>49,109</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>85,729</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender</th>
<th>Age group</th>
<th>Estimated number of people</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>15-24 years</td>
<td>52</td>
</tr>
<tr>
<td>Male</td>
<td>25-34 years</td>
<td>305</td>
</tr>
<tr>
<td>Male</td>
<td>35-44 years</td>
<td>1,254</td>
</tr>
<tr>
<td>Male</td>
<td>45-54 years</td>
<td>2,630</td>
</tr>
<tr>
<td>Male</td>
<td>55-64 years</td>
<td>3,657</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>7,898</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender</th>
<th>Age group</th>
<th>Estimated number of people</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>15-24 years</td>
<td>124</td>
</tr>
<tr>
<td>Female</td>
<td>25-34 years</td>
<td>476</td>
</tr>
<tr>
<td>Female</td>
<td>35-44 years</td>
<td>1,183</td>
</tr>
<tr>
<td>Female</td>
<td>45-54 years</td>
<td>2,368</td>
</tr>
<tr>
<td>Female</td>
<td>55-64 years</td>
<td>1,563</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>5,714</strong></td>
</tr>
<tr>
<td>Gender</td>
<td>Age group</td>
<td>Number of people</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td>------------------</td>
</tr>
<tr>
<td><strong>Not in the labour force due to OA as a main condition</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>15-24 years</td>
<td>44</td>
</tr>
<tr>
<td>Male</td>
<td>25-34 years</td>
<td>115</td>
</tr>
<tr>
<td>Male</td>
<td>35-44 years</td>
<td>250</td>
</tr>
<tr>
<td>Male</td>
<td>45-54 years</td>
<td>2,722</td>
</tr>
<tr>
<td>Male</td>
<td>55-64 years</td>
<td>6,779</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>9,910</strong></td>
</tr>
<tr>
<td>Female</td>
<td>15-24 years</td>
<td>5</td>
</tr>
<tr>
<td>Female</td>
<td>25-34 years</td>
<td>0</td>
</tr>
<tr>
<td>Female</td>
<td>35-44 years</td>
<td>981</td>
</tr>
<tr>
<td>Female</td>
<td>45-54 years</td>
<td>4,942</td>
</tr>
<tr>
<td>Female</td>
<td>55-64 years</td>
<td>13,494</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>19,422</strong></td>
</tr>
</tbody>
</table>