



Public Health  
England

Protecting and improving the nation's health

## **Movement Into Employment: Return on Investment Tool**

Estimation of benefits from moving an individual from unemployment into sustainable employment

# About Public Health England

Public Health England exists to protect and improve the nation's health and wellbeing, and reduce health inequalities. We do this through world-class science, knowledge and intelligence, advocacy, partnerships and the delivery of specialist public health services. We are an executive agency of the Department of Health, and are a distinct delivery organisation with operational autonomy to advise and support government, local authorities and the NHS in a professionally independent manner.

Public Health England  
Wellington House  
133-155 Waterloo Road  
London SE1 8UG  
Tel: 020 7654 8000  
[www.gov.uk/phe](http://www.gov.uk/phe)  
Twitter: @PHE\_uk  
Facebook: [www.facebook.com/PublicHealthEngland](https://www.facebook.com/PublicHealthEngland)

© Crown copyright 2017

You may re-use this information (excluding logos) free of charge in any format or medium, under the terms of the Open Government Licence v3.0. To view this licence, visit OGL or email [psi@nationalarchives.gsi.gov.uk](mailto:psi@nationalarchives.gsi.gov.uk). Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned.

Published October 2017  
PHE publications  
gateway number: 2017460

PHE supports the UN  
Sustainable Development Goals



## Authors and acknowledgements

This report was authored by:

Jacque Mallender, Partner, Optimity Advisors  
Rory Tierney, Senior Economist, Optimity Advisors  
Belinda Baah, Economist, Optimity Advisors  
Maria Stavridou, Assistant Economist, Optimity Advisors

Contributors:

Mariana Dates, Evidence Reviews Technical Lead, Optimity Advisors  
Obinna Onwude, Evidence Reviews Technical Lead, Optimity Advisors  
Solveig Bourgeon, Economist, formerly of Optimity Advisors

The contact details for the Optimity Advisors Project Director and Project Manager are listed below:

Project Director

Name: Jacque Mallender

Job title: Partner

Email: <mailto:jacque.mallender@optimityadvisors.com>

Tel: +44 (0)20 7553 4800

Project Management

Name: Rory Tierney

Job title: Senior Economist

Email: <mailto:rory.tierney@optimityadvisors.com>

Tel: +44 (0)20 7553 4838

### Acknowledgements

We would like to thank every member of the Steering Group who guided us in the evidence collection and development of the final deliverables (tool and report): Dame Carol Black (special advisor to PHE), Brian Ferguson (PHE), Shaun Donaghy (PHE), Annalisa Belloni (PHE), Joy Thompson (DWP), Anne Cunningham (Blackburn with Darwen Council), Francis Markus (Greater Manchester Combined Authority), Manuel Ramos (PHE), Justin Varney (PHE), Gina Radford (DH), Chris Anderson (DH/DWP), Matt Hennessey (PHE), Richard Heron (Faculty of Occupational Medicine), Emily Hough (NHSE), Paul Ogden (LGA), Sally Burlington (LGA) and Laura Adelman (DWP). Many thanks to all the stakeholders from PHE, including Nicola Pearce-Smith from PHE, who conducted the literature search that contributed to the project delivery. Finally, many thanks to Sally Wilson (Institute for Employment Studies) for her advice and contribution in the evidence research and final tool and report.

# Contents

|  |    |
|--|----|
| About Public Health England                                    | 2  |
| Executive summary  | 6  |
| 1. Introduction  | 12 |
| 2. Policy Context  | 14 |
| 3. Methods   | 15 |
| 3.1. Research Questions  | 15 |
| 3.2. Project approach  | 16 |
| 3.3. Evidence Review Methodology                               | 16 |
| 3.3.1. Search strategy   | 17 |
| 3.3.2. Screening   | 18 |
| 3.3.3. Extraction of relevant data                             | 18 |
| 3.4. Analysis and tool build                                   | 19 |
| 4. Findings from Evidence Review                               | 21 |
| 4.1. Overview of included studies                              | 21 |
| 4.1.1. Summary of evidence review findings                     | 21 |
| 4.1.2. Summary of health scale findings                        | 22 |
| 5. Economic Analysis   | 23 |
| 5.1. Assessment for modelling and feasibility findings         | 23 |
| 5.2. The SF-36: health outcomes                                | 25 |
| 5.2.1. Health outcomes: conversion to QALYs                    | 26 |
| 5.2.2. Common Mental Health Disorders – threshold approach     | 26 |
| 5.3. Incorporation of financial data                           | 28 |
| 5.4. Overall findings  | 30 |
| 5.5. Additional analysis and elements not included in the tool | 31 |
| 5.5.1. Health outcomes   | 31 |
| 5.5.2. Financial outcomes                                      | 34 |
| 6. Economic Tool   | 37 |
| 6.1. Tool Inputs   | 37 |
| 6.2. Tool Outputs  | 38 |
| 6.3. Indicative Case Study                                     | 40 |

|    |  |     |
|----|--|-----|
| 7. | Conclusion   | 43  |
| 8. | Bibliography   | 45  |
| 9. | Appendix   | 48  |
| A. | Search strategy & database search results                                  | 48  |
| B. | Call for Evidence  | 50  |
| C. | Inclusion / Exclusion Criteria   | 51  |
| D. | Quality Assessment Tool  | 53  |
| E. | Quality Assessment results for included studies                            | 57  |
| F. | High-level overview and quality assessment results for shortlisted studies | 63  |
| G. | Overview of health scales  | 66  |
| H. | Full data extraction for shortlisted and included studies                  | 68  |
| I. | In depth analysis of health scales and ratings                             | 98  |
| J. | Conversion of SF-36 Scores to QALY's                                       | 116 |
| K. | Mental health threshold methodology  | 117 |

## Executive summary

Addressing health-related worklessness has been a key Public Health England (PHE) priority for 2016/17. Work has been shown to improve the wellbeing of individuals, their families, and their communities from both an economic and a quality of life standpoint. Additionally, awareness around the potential detrimental effects of long-term worklessness, on both physical and mental health, is growing (Waddell and Burton, 2006<sup>1</sup>). There is a significant body of evidence specifically on mental health, showing a decline due to unemployment<sup>2</sup>.

This report accompanies the release of an economic tool. Optimity Advisors was commissioned to build the interactive tool to facilitate PHE's aim of bringing health more into the agenda around employment, and vice versa, by helping to demonstrate – quantitatively – that health is an important factor in decisions around employment, and to support more integrated government commissioning of beneficial services. It also feeds directly into another of PHE's key priorities – reducing health inequalities – given the intrinsic links between deprivation, employment status and health.

The tool itself allows local decision-makers to understand the health and financial impacts, for their local population, of getting people back to work, and can be customised in terms of population, mental health condition prevalence, and other variables. As a result of findings from the Rapid Evidence Review (RER), conducted in support of tool development, the tool focuses on the impact of the transition from unemployment to employment on mental health, not physical health. This is a consequence of available evidence predominantly showing the impact of employment on mental, not physical, health.

Note the results of this analysis apply only to those who are able to achieve sustainable employment either at baseline or with support. The analysis does not consider those unable to work due to the severity or circumstances of their health problem.

This tool has been developed to act as an aid to decision-makers, both for those who are considering implementing interventions in order to help people move from unemployment or economic inactivity to sustainable employment, and for those who are attempting to make the case for investment in this area. The model calculates the consequences of a person or person(s) returning to employment, providing the benefit side of the cost-benefit equation, and facilitating comparison against an intervention's cost. The tool itself does not examine specific interventions aimed at getting people into sustainable employment, but allows users to enter the cost of supporting people into work if applicable.

---

<sup>1</sup> Waddell G, Burton AK. Is work good for your health and wellbeing? The Stationery Office; 2006 Sep 6.

<sup>2</sup> Institute for Work and Health. Unemployment and mental health: Issue briefing. 2009. Available at <https://www.iwh.on.ca/briefings/unemployment-and-mental-health>.

## Tool structure

The tool itself is structured in the following way:

**Table 1: summary of economic tool**

| Section     | Description  |
|-------------|--|
| Welcome     | These sheets provide background information on the project, the aims of the tool and a guide explaining how the tool hopes to support local decision-makers in their understanding of the potential benefits of supporting an individual move from unemployment to employment.   |
| Information |  |
| Inputs      | <p>These sheets provide an opportunity for the user to view, select or override inputs into the tool. This includes selecting a local area (a Clinical Commissioning Group (CCG), a Sustainability &amp; Transformation Partnership (STP) area, a district/unitary authority, a county/unitary authority, combined authority, a region, or England as a whole), after which the tool populates the default numbers for the relevant population metrics: working age population, proportion claiming Job Seekers Allowance (JSA) and Employment and Support Allowance (ESA).<sup>3</sup></p> <p>Users can then select the proportion of people returning to work and the assumed length of benefit. This means the amount of time the average individual spends in work when they would otherwise have been unemployed/economically inactive. An optional user input is intervention cost information. If the user enters one-off fixed and ongoing annual costs of the intervention, the model will then compute Return on Investment (ROI) metrics, such as the benefit-cost ratio.</p> |

---

<sup>3</sup> Note that the decision to use JSA, ESA Work-Related Activity Group (ESA WRAG) & ESA Support Group (ESA SG) claimants as our population of interest as opposed to just the unemployed population as a whole, was due to the fact that those claiming ESA WRAG are deemed to be in a position where they can work towards an eventual returning to work and although this is not the case for the ESA SG, there is recognition that many ESA SG claimants would like gainful employment, and thus may be relevant depending on the extent of support provided by an intervention aimed at help people get back to work, and the timeline over which this can occur. Universal Credit claimants who are required to seek work are included in the tool as JSA claimants, and modelled as if they are claiming JSA – Universal Credit itself is not built into the tool at this stage.

|                |   |
|----------------|---|
|                | <p>Additionally, the user can override default mental health prevalence rates, override the monetary value of a Quality-Adjusted Life Year (QALY), as well as adjust the two discount rates provided, currently set at 3.5% financial costs and benefits, and 1.5% for health benefits, as recommended by PHE.</p>  |
| <p>Results</p> | <p>The tool provides per person and total QALY gains and their respective value, as well as calculating the proportion of people whose mental health status improves due to gaining employment. In addition, the cost savings from a reduction in the number of people facing common mental health disorders (CMDs) is provided, as well as the total health value, ie the value of the QALY gains and the cost savings to the healthcare system.</p> <p>It also provides the financial benefits to the exchequer, the benefits to the programme participant(s), and the benefits to society, broken down by category (eg increase in income tax, increase in wages, reduction of housing benefit payments etc.) as well as savings to health services.</p> <p>In the case where the user has provided an intervention cost, the model will report ROI metrics, namely; the Incremental Cost-Effectiveness Ratio<sup>4</sup> (ICER), the Net Benefit, and the Benefit-Cost Ratio.</p> |
| <p>Charts</p>  | <p>This sheet provides graphical representations of key outputs.</p>  |

## Default results

The economic tool draws its health benefits from the outcomes reported in Schuring et al. (2010), which are changes on the SF-36 scale, a validated, 36-item Short Form questionnaire which measures quality of life across eight domains<sup>5</sup> encompassing mental, physical and general wellbeing metrics. Schuring et al (2010) uses all eight

<sup>4</sup> This is defined as the intervention cost divided by the number of QALYs gained.

<sup>5</sup> Physical functioning; role – physical; bodily pain; general health; vitality; social functioning; role – emotional & mental health

domains of the SF-36 to measure the change in SF-36 scores for people who gained employment after being unemployed.

The raw change in SF-36 scores, reported in Schuring (2010), were used to calculate the potential health benefits, in terms of QALY gains and the reduction in the number of people with CMDs caused by moving from unemployment/economic inactivity<sup>6</sup> to employment.

Two other studies, Butterworth et al. (2011) and Galić et al. (2008), were shortlisted for use in the tool from the RER conducted for this project. Both also reported improvements in mental health, although Butterworth separated out good quality and poor quality jobs, and saw a decline in mental health for poor quality jobs.<sup>7</sup> These were not included in the final tool as the Schuring paper reported for all domains of the SF-36, and because of uncertainty over the size of the evidence base on the effect of job quality at present. Results for Butterworth and Galić are, however, given in the sensitivity analysis below.

Table 2 provides the key health outcomes of moving to employment using the findings from the Schuring et al. (2010) report. Note, that these figures relate to a baseline proportion of people with a CMD of 18.9% (based on findings from the Adult Psychiatric Morbidity Survey).

**Table 2: Key health outcomes**

|   |  |
|---|--|
| Study name                              | Schuring et al. (2010)   |
| Proportion whose mental health improves | 8.8% of the total population returning to work, 46.6% of those with CMD(s) at baseline |
| QALY gain per person                    | 0.0675 (per person returning to work)  |

With regards to financial outcomes, the model provides a breakdown of the net financial benefits accrued to the exchequer, disaggregated by the body to which they accrue (ie to national government, local authority and national health services), net financial benefits to programme participants and net financial benefits to society, which includes every perspective less transfers between parties. The model mainly builds on information provided by the Department for Work and Pensions (DWP) which in turn uses the Policy Simulation Tool (PSM) to derive changes in benefits and indirect tax from moving into work, with the benefit and tax system based on Budget 2016. **Table 3**

<sup>6</sup> It is assumed that these results also apply for those who are economically inactive

<sup>7</sup> The Butterworth study uses several different statistical methods to develop measures of the psychosocial characteristics of jobs using 12 items included in the Household, Income and Labour Dynamics in Australia (HILDA) survey

provides the key financial outcomes per person returning to work (using figures for England rather than a specific local geography).

**Table 3: Key financial outcomes<sup>8</sup>**

|   | Per person returning to work |
|---|------------------------------|
| <b>Financial benefits to the individual(s)</b>                  | <b>£3,500</b>                |
| <b>Financial benefit to society</b>                             | <b>£23,100</b>               |
| <b>Financial benefits to the exchequer, of which accrue to:</b> | <b>£12,000</b>               |
| <b>National Government</b>                                      | £11,400                      |
| <b>Local Authority</b>  | £500                         |
| <b>National Health Services</b>                                 | £85                          |

## Conclusion

This review and analysis has shown that there are significant mental health benefits and consequent financial savings from returning to work, as well as direct financial benefits, and has provided a tool to facilitate the inclusion of these findings in decision making at the local level.

The tool allows users to understand, for a given number of people who return to work, the health benefits and financial benefits to a variety of stakeholders. This can also be compared against an intervention cost to understand if the intervention is cost saving (financial benefits outweigh the cost) and/or cost-effective (the cost per QALY gained is below a given threshold).

The evidence incorporated within the tool was based on a rapid evidence assessment of available research. This picked up an array of recent research on the topic of the health benefits of returning to work but was not as comprehensive as a full systematic review. For instance, papers before 2006, or studies conducted outside of the EU countries, Switzerland, Canada, New Zealand, Australia, Norway or the USA, were not included. Citations from before 2006, such as those mentioned in Waddell & Burton’s overview of the evidence from 2006, were also not included.

Much of the current evidence on the health benefits of employment is found in evaluations of particular interventions. While this is valuable research and should be encouraged, non-intervention specific research can also help separate out the impact of an intervention versus the impact of employment more generally. It is recommended that more research, examining both the quality of life impact to individuals and the actual cost savings resulting from a return to work in the form of longitudinal studies with matched control groups, be conducted.

---

<sup>8</sup> Figure rounded to the nearest £100 if greater than £200

The evidence search clearly highlighted the scarcity of evidence around the potential benefits of employment on physical health. Of the few reports that included the impact of transitioning from unemployment to employment on physical health, a number showed some level of (non-significant) improvement. However, none of the studies focused primarily on physical health, with it being seen as a secondary issue. Consequently, due to the low level of robust evidence around physical health impacts, these potential benefits were omitted from the tool. More research examining the effects of employment, including the transition from unemployment to employment is advocated.

In addition, it would be interesting to explore how various variables impact on the health benefits of returning to work, such as those highlighted in the original research questions that were not possible to answer on current evidence. These include duration of unemployment, health conditions suffered by the individual, and reason for unemployment (eg whether or not it is health-related).

Further research to build on the Butterworth study is also encouraged with regard to investigating whether health benefits are dependent on the type or quality of job. A more comprehensive evidence base could help determine whether or not, and the extent to which type of quality of job affects wellbeing. Additionally, if job quality was better defined in the literature, this would help policy makers focus efforts on creating more of those types of jobs and/or steering those with health conditions towards them.

There are other benefits to employment, such as a possible reduction in crime, which was not included due to uncertainty around the estimates but which further research could develop.

Ultimately, the tool, combined with other available evidence around the benefits of employment, equips decision-makers with the evidence needed to support the case for better investment in areas aimed at alleviating worklessness. Better investment will likely lead to better outcomes for the individuals transitioning into employment, their families and their wider local communities, leading to an overall more healthy and inclusive society.

# 1. Introduction

Work has been shown to improve the wellbeing of individuals, their families and their communities from not only an economic but also a quality of life standpoint. The awareness around the potential detrimental effects of long-term worklessness, on both physical and mental health, is also growing (Waddell and Burton, 2006<sup>9</sup>). There is, too, a significant body of evidence specifically on mental health, showing a decline due to unemployment.<sup>10</sup>

Optimty Advisors was commissioned to build an interactive tool to facilitate PHE's aim of bringing health more into the agenda around employment, and vice versa, by helping to demonstrate – quantitatively – that health is an important factor in decisions around employment, and to support more integrated government commissioning of beneficial services. It also feeds directly into another of PHE's key priorities – reducing health inequalities – given the intrinsic links between deprivation, employment status and health.

The tool itself, which this report accompanies, allows local decision-makers to understand the health and financial impacts, for their local population, of getting people back to work, and can be customised in terms of population, mental health condition prevalence, and other variables.

Section 2 of this report describes the wider policy context of tackling worklessness and its relationship to health outcomes.

Section 3 describes the methodology and criteria used to conduct an evidence review, an explanation of how data was synthesised and extracted and a high-level overview of data analysis and the tool build, with greater detail provided in Section five.

Section 4 describes the findings from the evidence review, including quality assessments of key papers found during the evidence search and a summary of the evidence review findings.

Section 5 describes in detail, the economic analyses conducted in order to convert these findings into health outcomes, in terms of potential Quality Adjusted Life Years (QALYs) gained and healthcare cost savings, associated with an improvement in mental health. In addition, this section describes analysis of financial data used to express the financial outcomes achieved from moving a person or persons from unemployment to employment, in terms of savings to the exchequer, disaggregated by agency to which the savings accrue; savings to the individual, and savings to wider society.

---

<sup>9</sup> Waddell G, Burton AK. Is work good for your health and wellbeing? The Stationery Office; 2006 Sep 6.

<sup>10</sup> Institute for Work and Health. Unemployment and mental health: Issue briefing. 2009. Available at [https://www.iwh.on.ca/system/files/documents/iwh\\_briefing\\_mental\\_health\\_2009.pdf](https://www.iwh.on.ca/system/files/documents/iwh_briefing_mental_health_2009.pdf).

Section 6 provides a step-by-step description of tool inputs and outputs.

Finally, Section seven of the report presents conclusions, key findings, gaps in the evidence base around worklessness and its links to health outcomes as well as recommendations for future research.

## 2. Policy context

There are economic and social arguments that work is one of the most effective ways to improve the wellbeing of individuals and their communities. However, assessments of the health benefits that result from a return to work, and the associated financial benefits that result, have rarely been fully quantified.

Addressing health-related worklessness has been a key PHE priority for 2016/17. This is in alignment with national government objectives (eg governmental commitment to halve the disability employment gap, and a forthcoming green paper on health, work and disability) and other key stakeholders, such as the NHS (eg workplace health is prioritised by NHS England in the Five Year Forward View) and wider government. The Department for Communities and Local Government is specifically tasked with providing greater freedom and flexibility to local government so that it has the resources to genuinely lead its communities and drive local growth.

It is known that health problems are a key cause of worklessness, whether work-related or not: having a long-term condition (LTC) is associated with unemployment and worklessness, with an employment rate of only 60% – typically 11-15 percentage points lower than the general population.<sup>11</sup> A key aim of the Department of Health's NHS Outcomes framework is to improve functional ability for people with LTCs via gainful employment.<sup>12</sup> Moreover, the disability employment gap – the difference between the employment rates of disabled and nondisabled people – stood in mid-2016 at 32 percentage points.<sup>13</sup> The government has pledged to halve this gap, and has introduced a new Work and Health Programme to support those in long-term unemployment due to health conditions or disabilities. According to a DWP Green Paper, “too many people are falling into a downward spiral of declining health and worklessness, not reaping the benefits of employment and creating pressures on the healthcare system”.

However, this project aims to understand the potential benefits of returning to work, including its impact on health status. Specifically, the primary aim of this project is to assess the benefits that come from moving from unemployment (or economic inactivity) to sustainable employment, and the cost savings that result to the public purse from employment itself, and improved health outcomes.

While this analysis will not look at interventions to get people into work, it will attempt to show the benefits that can result beyond individual remuneration and productivity, ie health outcomes, and be used to facilitate the discussions around work-based interventions.

---

<sup>11</sup> Employment of people with long term conditions (source: NHS Outcomes Framework; NHS Digital). 2015/2016. Available at <http://www.qualitywatch.org.uk/indicator/employment-people-long-term-conditions>

<sup>12</sup> NHS Outcomes Framework: at a glance. 2016. Available at [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/513157/NHSOF\\_at\\_a\\_glance.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/513157/NHSOF_at_a_glance.pdf)

<sup>13</sup> Key Statistics on People with Disabilities in Employment. 2016 Available at <researchbriefings.files.parliament.uk/documents/CBP-7540/CBP-7540.pdf>

## 3. Methods

The tool focuses on the impact of the transition from unemployment to employment on mental health, not physical health. This is a consequence of available evidence; the majority of currently available evidence is around the impact of employment on mental, not physical health. There is a wide range of robust, useable evidence, regarding the positive impact of employment on mental health but the evidence on benefits of employment to physical health is extremely limited. Chapter 4 explains in detail the research findings, which support the rationale behind focusing the tool on the impact of employment on mental health.

### 3.1. Research questions

Given the context described above, PHE commissioned Optimity Advisors to develop an interactive tool, the aim of which was to use analysis of existing data to answer the following research questions (which were slightly modified in discussions with the project Steering Group):

1. What are the cost-savings for health and social care services from moving an individual from unemployment (and, if possible, inactivity) into sustainable employment?
2. What are the cost-savings for the wider government, such as reduction in welfare benefits, pressure on local services etc. from moving an individual from unemployment (and, if possible, inactivity) into sustainable employment?
3. What are the benefits to the individual and to the local economy from moving them from unemployment (and, if possible, inactivity) into sustainable employment?
4. Does this cost saving vary based on the duration of unemployment (and, if possible, inactivity)?
5. Does this cost saving vary based on the gender, age or other demographic of the unemployed (and, if possible, inactive) individual?
6. How does the nature or duration of health issues contributing to unemployment (and, if possible, inactivity) affect the cost saving? If possible, differentiate between cases where health is a key factor in becoming unemployed and cases where health problems emerge after the onset of unemployment.
7. Does this cost saving vary in national-local variation? Local level to be defined as local authority, CCG or STP (eg the unemployment rate in the North East is 7.7% while in the South East it is 4.5%)

A Rapid Evidence Review (RER) was conducted in order to capture the available evidence base around the questions above and gather the necessary data to build an interactive economic tool. Given this, the search strategy was restricted to papers that provided some level of economic analysis and that specifically looked at the transition of moving from unemployment (or economic inactivity) to employment.

After completing the RER and analysing all available evidence, it was clear that we would be unable to answer all of the questions. The revised focus of the tool was thus to show:

1. The cost-savings for health and social care services from moving an individual from unemployment into sustainable employment.
2. The cost-savings for the wider government, such as reduction in welfare benefits, from moving an individual from unemployment into sustainable employment.
3. The health benefits to the individual from moving from unemployment to sustainable employment.

Consequently, an additional outcome of this review will be an assessment of any gaps in the evidence base, which will form part of our conclusions, described in Section 7.

### 3.2. Project approach

Throughout the project, Optimity Advisors met regularly with a Steering Group consisting of project and topic experts and consulted with external subject matter experts. In addition, a representative user group has also been utilised to test the usability, reliability and relevance of the tool for its target user group.

The project approach took place in two phases. The approach adopted in Phase I consisted of:

- a RER to identify evidence on the health outcomes of employment, focusing specifically on the transition from unemployment to employment;
- full extraction of data from the papers in the review that could help populate an economic model

On completion of Phase I, Phase II of the project commenced. Phase II included:

- further analysis and parameter searching to establish health and financial outcomes
- build of economic tool

The rest of Section 3 and Sections 4, 5 and 6, provide greater detail with regards to both Phases I and II.

### 3.3. Evidence review methodology

A systematic approach was adopted to identify, assess and synthesise relevant information. This ensured that findings were comprehensive as well as reliable, and that the modelling and resultant economic tool is based on an accurate understanding of the evidence. The review was conducted using a RER, the methodology for which is summarised in the diagram below.

**Figure 1: Evidence review approach**



### 3.3.1. Search strategy

The search for evidence took three forms:

- a formal literature search of databases of published literature, using a specific search strategy
- a search for grey literature, through examining key websites and Google searches
- a call for evidence issued to the members of the project Steering Group and other stakeholders, and published online

Using an agreed strategy, a simplified version of which is provided in Figure 2 below, a formal literature search of the following electronic databases was conducted for studies published between 2006 and 2016 inclusive:<sup>14</sup>

- Medline
- Embase
- PsycINFO

---

<sup>14</sup> Studies conducted before 2006 that were referred to in papers from 2006-2016 were not reviewed. For instance, it was not possible within the scope of this study to assess the papers referenced by Waddell and Burton (2006), in their study: Waddell, G., Burton, A.K. 2006 [Accessed 18 Aug 2017] Is work good for your health and wellbeing? Department for Work and Pensions.

[https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/214326/hwwb-is-work-good-for-you.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/214326/hwwb-is-work-good-for-you.pdf)

## Figure 2: Simplified search strategy

Simplified search strategy:

- terms for employment
- terms for worklessness
- Terms for health, mental health, government, benefits, taxation, income, insurance, socioeconomic factors, financial outcomes, social outcomes, crime, housing, smoking, physical activity, caregiving, lifestyle factors

Or:

- Terms for employment AND health in title

The full search strategy and the results of the searches can be found in Appendix A.

Grey literature was retrieved using search terms such as 're-employment benefits. in Google Scholar, and also via specific websites, such as those of the National Institute for Health and Care Excellence (NICE), PHE, NHS, the Department for Work and Pensions (DWP), the Department of Health (DH), and Mind.

Finally, a call for evidence was issued in December 2016, and literature received before the mid-January deadline was included. The call for evidence is provided in Appendix B.

### 3.3.2. Screening

All records identified by the searches were uploaded into a database and duplicate records were removed. Inclusion criteria were developed to identify relevant studies for the estimation of benefits of moving from unemployment to sustainable employment. Initially, the records were screened on title and abstract. Where no abstract was available, a web search was first undertaken to locate one; if no abstract could be found, records were screened on title alone. The full inclusion / exclusion criteria checklist can be found in Appendix C.

### 3.3.3. Extraction of relevant data

Given the remit of the project was to assess the health benefits of moving from unemployment to sustainable employment, papers were excluded if they:

- a) Reported only the differing health characteristics of those who are employed in comparison to those who are unemployed
- b) Reported the health decrement from becoming unemployed; and/or
- c) Reported the health benefits of employment achieved via a programme likely to or shown to improve one's health in and of itself.

The last of these, intervention studies, were specifically excluded in order to avoid capturing health benefits related to this type of intervention rather than the potential health gains from return to work itself.

Each included study underwent a thorough data extraction exercise where key information was collected, analysed for suitability for the analysis and assessed in terms of quality (assessment tool can be found in Appendix D) and applicability. Examples of the key information collected is listed below, with the full data extraction table presented in Appendix H.

The key data extracted from each included full text paper included:

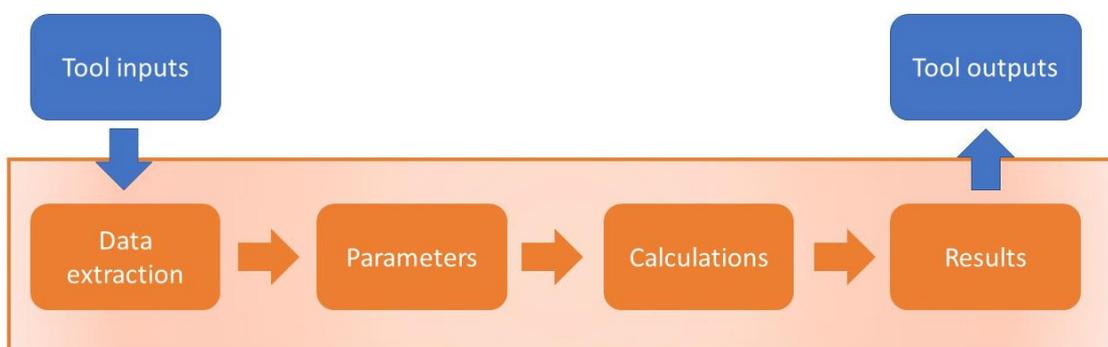
- study population
- study type (eg systematic review)
- return to work definition
- length of follow-up
- duration of employment/unemployment
- outcome type – health, financial or other
- outcome metric (eg overall health measured by the SF-36)
- key findings

### 3.4 Analysis and tool build

Following extraction of data on the health benefits of a return to work from the evidence review, further analysis, searching for model parameters, and model design was required in order to produce a functioning economic tool.

Model design, or conceptualisation, involved understanding what the tool would look like and include. The aim was to include information on the health and financial outcomes for a given local population, allowing users to enter inputs and see results, with an underlying model carrying out the calculations. This high-level schema is shown below.

**Figure 3: High-level tool design**



The health and financial data required included QALY gains and cost savings to the health service and to wider government. In order to calculate this, the health gains reported in the papers extracted from the review were converted into QALY gains and healthcare cost savings where possible, and the approach taken for this is detailed under 'Economic Analysis' in Section 5.

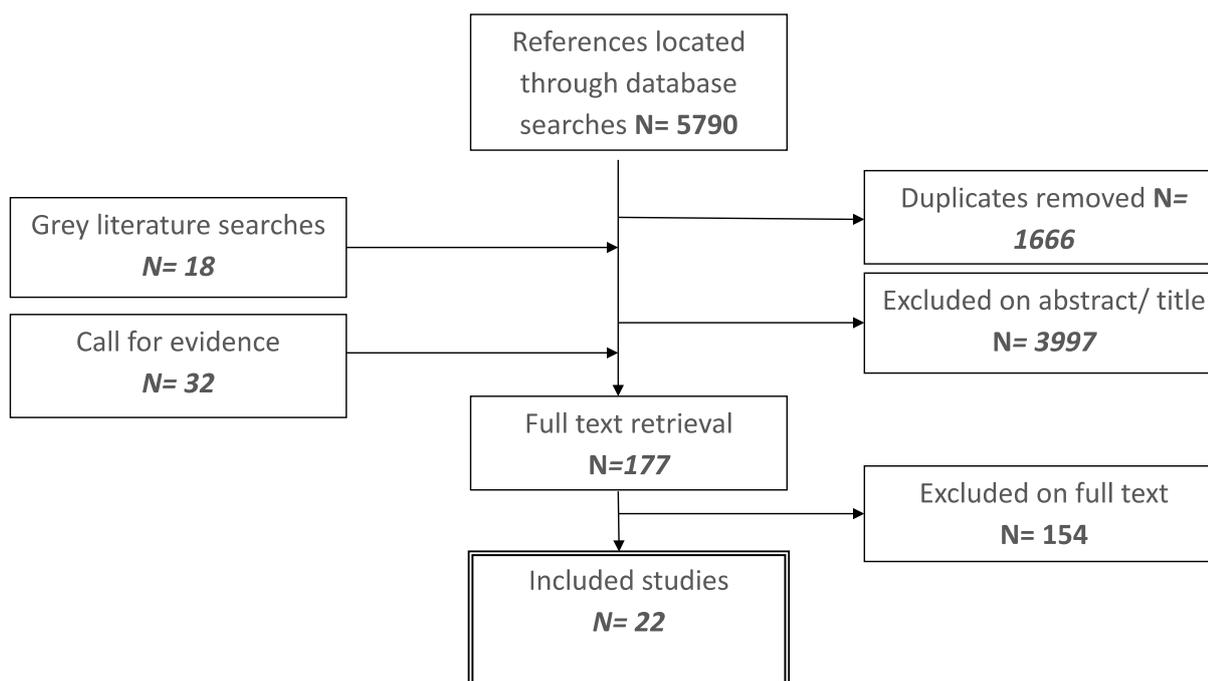
Additional parameters were also sourced around the direct financial benefits of a return to work, and this was done through contact with the DWP.

Finally, population demographics, worklessness figures, and mental health prevalence were also sourced for each local authority, and mapped to other geographical locations such as CCGs and STPs in England.

## 4. Findings from evidence review

### 4.1. Overview of included studies

The flow of studies through the evidence review is shown in the diagram below.



#### 4.1.1. Summary of evidence review findings

Appendix F shows a full data extraction of each shortlisted study. All included systematic review studies were quality assessed using the AMSTAR – Assessing the Methodological Quality of Systematic Reviews<sup>15</sup> – checklist tool. All other quantitative studies were quality assessed using the quality assessment tool for quantitative studies created by the Effective Public Health Practice Project (EPHPP). On the basis of the answers to the questions within these tools, each study was given an overall quality score. The three studies ultimately selected to be included in the tool or sensitivity analysis, Schuring 2010 (tool), Butterworth et. al. 2010 (sensitivity analysis), and Galić & Šverko, 2008 (sensitivity analysis), all used the EPHPP assessment tool as they were all quantitative studies. The completed assessment of their quality can be found in Appendix D with the quality assessment outcomes for shortlisted, but ultimately not included, studies in Appendix F.

<sup>15</sup> AMSTAR checklist can be accessed at [https://amstar.ca/Amstar\\_Checklist.php](https://amstar.ca/Amstar_Checklist.php)

#### 4.1.2. Summary of health scale findings

All shortlisted papers reported health gains in terms of changes in health scores across a wide range of health scales aimed at measuring changes in mental health. Table 11 in Appendix G, provides an overview of different health scales used in the abovementioned studies, providing a reason for inclusion or exclusion.

## 5. Economic analysis

### 5.1. Assessment for modelling and feasibility findings

Once the full data extraction of shortlisted studies was completed, key information was categorised by:

- outcome metric;
- return to work (RTW) definition (eg full-time, part-time)
- study population

This allowed greater understanding of the potential RTW definitions, population types and outcome metrics that the tool could include. From the evidence gathered, it was clear that different permutations of return to work (eg full-time or part-time) and population types (eg long-term unemployed or short-term unemployed), was not possible to include in the analysis.

The tool, therefore, draws from the population in England who are workless but would be expected to return to work, ie Job Seekers Allowance (JSA), whose claimants are capable of work and actively seeking employment<sup>16</sup>, and Employment and Support Allowance (ESA), whose claimants have an illness or disability that affects their ability to work.<sup>17</sup> The tool includes both the ESA Work-Related Activity Group (ESA WRAG), where there is an expectation of progress towards a return to work, and the ESA Support Group (SG), where there is no expectation. However, ESA SG claimants have also been included in the tool despite the fact that their benefits are not contingent on seeking employment or engaging in work-related activities and they are not expected to do anything to improve their chances of work.<sup>18</sup> In the recently published Department of Health and Department for Work and Pensions report, 'Improving lives: The work, health and disability Green Paper (2016)', it was made clear that this group should not be presumed to be indefinitely incapable of employment. For instance, the report showed that 52% of people in the Support Group do want to work. Given this, the tool will also be of use to those who want to see what the benefits are of ESA SG claimants moving to employment.<sup>19</sup> However, this applies only to those who are able to achieve sustainable employment, either at baseline or with support.

---

<sup>16</sup> Eligibility conditions for JSA can be found at <https://www.gov.uk/jobseekers-allowance/eligibility>

<sup>17</sup> Eligibility conditions for ESA can be found at <https://www.gov.uk/employment-support-allowance>

<sup>18</sup> Citizens Advice – About the ESA support and work-related activity group. Accessed: <https://www.citizensadvice.org.uk/benefits/sick-or-disabled-people-and-carers/employment-and-support-allowance/while-youre-getting-esa/about-the-esa-groups/>

<sup>19</sup> As from April 2017, the WRAG component will no longer be available to new claimants. Those on ESA will receive the same rate as those with JSA, alongside additional support.

The tool does not include Universal Credit (UC) benefits, however the numbers provided in the tool for JSA claimants does include those claiming UC and required to seek work. For the purposes of the tool, these claimants are modelled as if they are claiming JSA. The ESA claimant numbers included in the tool do not include any UC claimants.<sup>20</sup>

In terms of health benefits, the evidence showed that the majority of studies focused solely on mental health. In the few instances where physical health was also taken into consideration, the results were generally insignificant and/or the findings were unable to be converted into usable data in the tool (See Appendix F).

Research, therefore, focused on understanding the mental health outcomes of returning to work. To measure this impact, the evidence gathered used a range of mental and combined mental and physical health scales, usually calculated using results from patient questionnaires, to understand the health impact of employment. An assessment of the quality and applicability of the various scales reported was consequently undertaken (see Table 11 in Appendix G for the outcome of this assessment). The scales deemed to be of best quality and applicability reported were:

- 36-item Short Form Survey (SF-36) which includes the Mental Health Inventory (MHI-5)
- 12-item Short Form Survey (SF-12) which is a shorter version of the SF-36
- 12-item General Health Questionnaire (GHQ-12)

The shortlisted reports also included scales such as the Rosenberg Self-esteem Scale and Comprehensive Quality of Life Scale (ComQol), but these scales were excluded for a number of reasons including a lack of information on the distribution of the scale used or if it was included in a study deemed as weak. In addition, although the results from some papers appeared useful, extra information was required to include their findings in the report and tool, but this had not been made available by the author(s).

Once it was clear which health scales could feasibly be included in the analysis, extra research was conducted to understand how the findings from included reports and scales, could be incorporated in the tool. From our research, the SF-36 scale, which encompasses the MHI-5 health scale, was the only scale where QALYs could be robustly estimated and used within the tool. Unfortunately, due to a lack of information regarding the nature of the distribution of commonly used scales, such as the GHQ-12, the health gains were unable to be calculated from results using these scales.

---

[https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/443232/50325\\_Summer\\_Budget\\_15\\_Web\\_Accessible.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/443232/50325_Summer_Budget_15_Web_Accessible.pdf)

<sup>20</sup> JSA numbers come from the ONS CC01 claimant count, available at <https://www.ons.gov.uk/employmentandlabourmarket/peoplenotinwork/unemployment/datasets/claimantcountbyunitaryandlocalauthorityexperimental> . An average of figures for January 2016, April 2016 and July 2016 was used. ESA numbers come from Nomis: <https://www.nomisweb.co.uk/> An average of figures for February 2016, May 2016 and August 2016 was used, based on an analysis run on 7 July 2017.

## 5.2. The SF-36: health outcomes

The Short Form-36 (SF-36) is a validated, 36-item questionnaire, which measures quality of life across eight domains, namely:

- physical functioning (PF)
- role: physical (RP)
- bodily pain (BP)
- general health (GH)
- vitality (VT)
- social functioning (SF)
- role: emotional (RE)
- mental health (MH)

It is a validated, often used scale with scores ranging from 0-100, where higher scores indicate better-perceived health. The mental health component of the SF-36, known as the Mental Health Inventory (MHI-5), is a stand-alone questionnaire consisting of five questions used for evaluating mental health issues such as anxiety, depression and general distress. Three reports included in our search use the SF-36 to measure the change in mental health in a cohort of people returning to work from unemployment.

The economic tool draws its health benefits from the outcomes reported in Schuring et al. (2010) which provides a holistic approach to changes in SF-36 score as their analysis uses all eight domains of the SF-36. Specifically, Schuring et al. (2010) measures the change in SF-36 scores of people who gained employment from unemployment.

**Table 4** provides a brief overview of Schuring et al. (2010), and the raw change in SF-36 score for its study population.

**Table 4: SF-36 scores for Schuring (2010)**

| Report          | Domains of SF-36 used and follow-up period | Results   |
|-----------------|--|---|
| Schuring (2010) | All eight domains: six-month follow-up     | Physical Function (PF) = +9.2<br>Role – Physical (RP) = +20<br>Bodily Pain (BP) = +11.3<br>General Health (GH) = +7<br>Vitality (VT) = +7.8<br>Social Functioning (SF) = +14.2<br>Role – Emotional (RE) = +22.7<br>Mental Health (MH) = +11 |

These raw scores were subsequently used to calculate the potential health benefits, in terms of QALY gains and the reduction in the number of people with common mental health disorders (CMDs) caused by moving from unemployment/economic inactivity<sup>21</sup> to employment.

### 5.2.1. Health outcomes: conversion to QALYs

Ara & Brazier (2008)<sup>22</sup> provide an algorithm that converts the eight SF-36 values into the SF-6D, a preference-based measure of health reporting QALYs per person. The SF-6D calculates utilities on a 0.296-1 scale, where 0.296 represents the maximum impaired level and 1 represents full health, and was formulated to be used to represent health states in economic evaluations.

$$SF - 6D = 34.31814 + 0.0994 * PF + 0.0215 * RP + 0.1083 * BP + 0.0140 * GH + 0.0479 * VT + 0.1001 * SF + 0.0394 * RE + 0.1269 * MH)/100$$

Using the change in SF-36 scores provided in Schuring (2010), the QALY gains per person/per cohort returning to work from unemployment were calculated. This equates to a 0.0675 QALY gain per person moving from unemployment to employment using the results provided in Schuring et al. (2010), when assuming the benefit is sustained for one year. The full methodology for converting the scales can be found in Appendix J and the results from the calculations can be found in Section 5.4.

### 5.2.2. Common mental health disorders: threshold approach

The SF-36 can be converted into two summary scales, the Physical Component Summary (PCS) and the Mental Component Summary (MCS). Using the average of three UK nationwide SF-36 scores provided in Bowling et. al. (1999)<sup>23</sup> as baseline, the methodology provided by Ritvo et. al. (2007)<sup>24</sup> was used to create the MCS for the target population. The methodology transforms the raw baseline SF-36 scores to a 0-100 scale that is normally distributed by standardising the eight SF-36 domains, weighting and aggregating these scales and transforming them into the MCS. As the MCS is normally distributed, the mean MCS score (50.5) can be calculated and a threshold approach used to calculate the number of people who, from moving from unemployment to employment, fall below the clinical threshold for a common mental

---

<sup>21</sup> It is assumed that these results also apply for those who are economically inactive

<sup>22</sup> Ara R, Brazier J. Predicting the short form-6D preference-based index using the eight mean short form-36 health dimension scores: estimating preference-based health-related utilities when patient level data are not available. *Value in Health*. 2009 Mar 1;12(2):346-53.

<sup>23</sup> Bowling A, Bond M, Jenkinson C, Lamping DL. Short Form 36 (SF-36) Health Survey questionnaire: which normative data should be used? Comparisons between the norms provided by the Omnibus Survey in Britain, the Health Survey for England and the Oxford Healthy Life Survey. *Journal of Public Health*. 1999 Sep 1;21(3):255-70.

<sup>24</sup> Ritvo P, Fischer JS, Miller DM, Andrews H, Paty DW, LaRocca NG. MSQLI—Multiple Sclerosis Quality of Life Inventory. A user's manual. New York: National MS Society. 1997

health disorder. The threshold mean value of 44.8 provided in Kelly et. al. (2008)<sup>25</sup> was used; this is based on the misclassification rate, ie it achieves the lowest sum of false positives and false negatives, and is also the rate from the Kelly paper that gives the highest specificity (true negative rate).

By running the figures from Schuring (2010) through this set of calculations, the proportion of people crossing the threshold from 'mental ill health' to 'mental health' can be calculated. This is represented in the calculations as a proportion of those at baseline with a mental health condition.

In the tool, these figures are then scaled according to the baseline proportion of people with a common mental health disorder (CMD), which includes depression, anxiety, phobias, Obsessive Compulsive Disorder (OCD), panic disorder, and 'CMD – not otherwise specified', which includes those with symptoms of both depression and anxiety. These figures are drawn from the 2014 Adult Psychiatric Morbidity Survey.<sup>26</sup>

Default costs of mental health disorders are drawn from McCrone et al.<sup>27</sup> and factor in the proportion of individuals with CMDs who access health services. These include costs due to medication, residential care, GP, psychiatric inpatient, Social Security Disability, other NHS (non-inpatient), and non-psychiatric inpatient. It is possible that this provides an overestimate of relevant costs, on the assumption that those seeking inpatient treatment are less likely to return to work, but in the absence of data relating to that, all cost types have been included.

Applying these costs to the number crossing the threshold provides an estimate of the cost savings from reduced mental health morbidity.<sup>28</sup>

For both the QALY gains and the reduced cost it is assumed that the benefits of a return to work are sustained for the duration of employment. Although there is no data specifically substantiating this, there is evidence that remaining unemployed causes deterioration in mental health, and this is believed to be a reasonable assumption. Additionally, given the threshold approach undertaken in the analysis, only those that are assumed to cross the predefined threshold and apportion the health and financial gains to these individuals are accounted for; those whose mental health improves but not by an amount that pushes them over the threshold are not included.

---

<sup>25</sup> Kelly MJ, Dunstan FD, Lloyd K, Fone DL. Evaluating cutpoints for the MHI-5 and MCS using the GHQ-12: a comparison of five different methods. *BMC psychiatry*. 2008 Feb 19;8(1):10.

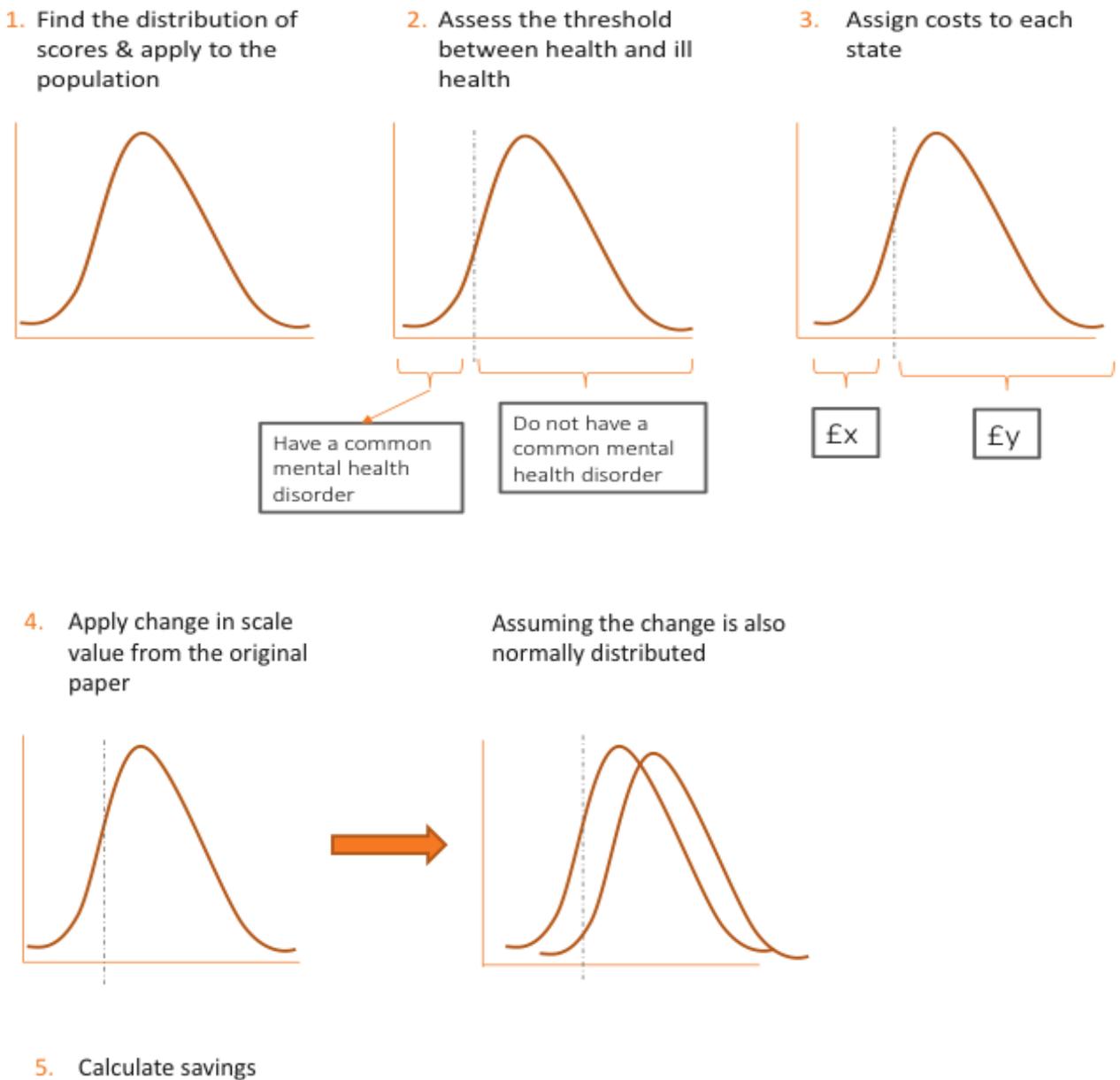
<sup>26</sup> 'Mental health and wellbeing in England: Adult Psychiatric Morbidity Survey 2014. Available at: <http://content.digital.nhs.uk/catalogue/PUB21748/apms-2014-full-rpt.pdf>

<sup>27</sup> McCrone PR, Dhanasiri S, Patel A, Knapp M, Lawton-Smith S. Paying the price: the cost of mental health care in England to 2026. King's Fund; 2008. Available at: [https://www.kingsfund.org.uk/sites/files/kf/Paying-the-Price-the-cost-of-mental-health-care-England-2026-McCrone-Dhanasiri-Patel-Knapp-Lawton-Smith-Kings-Fund-May-2008\\_0.pdf](https://www.kingsfund.org.uk/sites/files/kf/Paying-the-Price-the-cost-of-mental-health-care-England-2026-McCrone-Dhanasiri-Patel-Knapp-Lawton-Smith-Kings-Fund-May-2008_0.pdf)

<sup>28</sup> It is acknowledged that this approach is a simplification, as it assumes all costs disappear for those who cross the threshold (as soon as they return to work), and that there is no change for everyone else. However, in the absence of available data to refine the estimate it is presented as an approximation of cost savings.

Figure 4 provides a graphical illustration of the methodology behind the threshold approach. The full methodology for converting the scales can be found in Appendix K.

**Figure 4: Converting health scores using the threshold approach**



### 5.3. Incorporation of financial data

The model provides a breakdown of the net financial benefits accrued to the exchequer, disaggregated by the body to which they accrue (ie to national government, local authority and national health services), net financial benefits to programme participants and net financial benefits to society, which includes every perspective less transfers between parties. The model mainly builds on information provided by the Department for Work and Pensions (DWP). DWP uses the Policy Simulation Model (PSM) to derive

changes in benefits and indirect tax from moving into work, with the benefit and tax system based on Budget 2016.

On wages and subsequently income tax, national insurance and indirectly childcare costs, the DWP model draws on evidence from the Destinations Survey 2011, a one-off survey of 9,000 customers who have left either Jobseekers Allowance, Income Support or Employment and Support Allowance. The survey includes information on the immediate and substantive destination of customers after leaving benefits, such as employment, other types of benefits, education etc. Data on wages are based on the average earnings of those who ended their claim to enter any type of employment (eg fixed-term contract, permanent or open-ended contract, full-time or part-time employment, uprated with Budget 2016 earnings inflators.<sup>29</sup> A caveat is that the earnings of people with poor mental health might be slightly less than the average of this population.

Responses of the 2012/2013 Family Resources Survey (FRS) participants with low family earnings regarding childcare costs are used to proxy the increased childcare costs resulting from moving into employment.

On welfare, quintile household income estimates produced by the Office for National Statistics using the Living Costs and Food Survey are used to derive the welfare weight, so as to address the issue of diminishing marginal utility of income.

Operational costs refer to costs related to claims maintenance and adviser interventions, estimated by combining benefit caseload data (DABM) with operational cost estimates.

Regarding increased travel costs, the costs and benefits of additional travel are estimated by combining estimates of the amount of commuting using different modes of transport and the cost and benefits of using each mode, obtained from the Department for Transport (DfT) National Travel Survey.

Healthcare cost savings are calculated in the DWP model according to the Fujiwara (2010) methodology<sup>30</sup>, but are replaced with the healthcare estimates in this tool.<sup>31</sup>

---

<sup>29</sup> Adams L, Oldfield K, James AS. Destinations of Jobseeker's Allowance, Income Support and Employment and Support Allowance Leavers 2011. Department for Work and Pensions; 2012.

<sup>30</sup> Fujiwara D. The Department for Work and Pensions social cost-benefit analysis framework. DWP Working Paper 86, London: Department for Work and Pensions; 2010.

<sup>31</sup> Fujiwara's approach for health does not facilitate the inclusion of QALYs and uses a different approach to calculate health cost savings, which is covered in sensitivity analysis.

### 5.4. Overall findings

This model has been developed to act as an aid to decision-makers, for both those who are considering implementing interventions in order to help people move from unemployment/economic inactivity to employment, and those who are attempting to make the case for investment in this area. The model calculates the consequences of a person or person(s) returning to employment, providing the benefit side of the cost-benefit equation, and facilitating comparison against an intervention’s cost.

Users are able to provide the cost of an intervention aimed at returning people to work, as an option in the tool. If a programme cost is included in the user input sheet, the model will calculate Return on Investment (ROI) metrics alongside the consequences of employment, namely: the Incremental Cost-Effectiveness Ratio (ICER), the Net Benefit, and the Benefit-Cost Ratio. Figure 5 provides a high-level snapshot of the tool’s inputs and outputs, with sheet-by-sheet descriptions given in Section 6. (Note that interventions may themselves lead to health benefits, which would not be factored in this tool).

**Figure 5: High-level tool inputs and outputs**

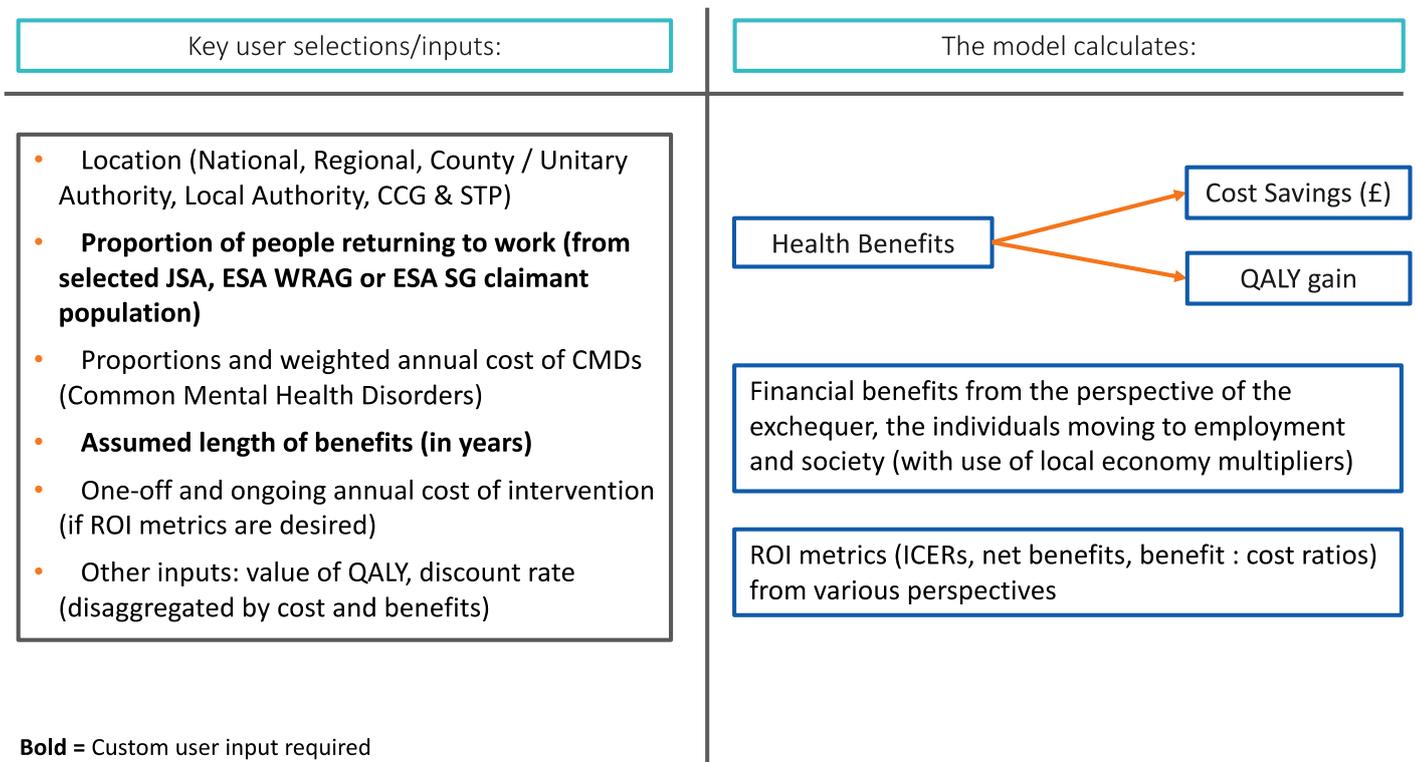


Table 5 provides the key health outcomes of moving from unemployment to employment using the findings from the Schuring et al. (2010) report. These figures relate to a baseline proportion of people with a CMD of 18.9%.

Table 6 provides the key financial outcomes per person returning to work (using figures for England rather than a specific local geography). N.B. All results greater than £200 are rounded to the nearest £100.

**Table 5: Key health outcomes**

|                                       |  |
|---------------------------------------|--|
| <b>Study name</b>                     | Schuring et al. (2010)   |
| <b>Proportion with CMDs improving</b> | 8.8% of the total population returning to work, 46.6% of those with CMD(s) at baseline |
| <b>QALY gain per person</b>           | 0.0675 (per person returning to work)  |

**Table 6: Key financial outcomes**

|   | Per person returning to work |
|---|------------------------------|
| <b>Financial benefits to the individual(s)</b>                  | <b>£3,530</b>                |
| <b>Financial benefit to society</b>                             | <b>£23,070</b>               |
| <b>Financial benefits to the exchequer, of which accrue to:</b> | <b>£12,030</b>               |
| <b>National Government</b>                                      | £11,410                      |
| <b>Local Authority</b>  | £535                         |
| <b>National Health Services</b>                                 | £85                          |

## 5.5. Additional analysis and elements not included in the tool

In order to test the validity of the estimates/compare the findings from the analysis, to other studies that included information that could have been utilised within the tool, an additional economic analysis was conducted. The results are set out below.

As aforementioned, healthcare cost savings are calculated in the DWP model according to the Fujiwara (2010) methodology,<sup>32</sup> but are replaced with Optimity Advisor’s healthcare estimates in this tool.<sup>33</sup> The rest of this chapter outlines what the results would have been if the findings from three reports, namely Butterworth (2011), Galić (2008) and Fujiwara (2010), that investigated the potential benefits of moving from unemployment to employment, instead of Schuring et al. (2010), had been incorporated. Additionally, reasoning for not including these findings in the tool has been provided.

### 5.5.1. Health outcomes

<sup>32</sup> Fujiwara D. The Department for Work and Pensions social cost-benefit analysis framework. DWP Working Paper 86, London: Department for Work and Pensions; 2010.

<sup>33</sup> Fujiwara’s approach for health does not facilitate the inclusion of QALYs and uses a different approach to calculate health cost savings, which is covered in sensitivity analysis.

The results from two initially shortlisted reports: Butterworth et al. (2011) and Galić et al. (2008), both of which included a domain or domains of the SF-36, in their analysis, are shown in Table 7, with Table 8 showing the raw SF-36 scores for each report. These figures show the proportion of people with a common mental health disorder(s) (CMD(s)), who see an improvement in their mental health status.

**Table 7: Sensitivity analysis - health outcomes**

| Study name                                    | Participants  | Proportion of people with CMD(s) improving | QALY gain per person |
|---|---|--|----------------------|
| Schuring et al. (2010)                        | 47 re-employed individuals of 965 unemployed individuals    | 46.6%                                      | 0.0675               |
| Butterworth et al. (2011) – Good quality job  | 693 relevant survey respondents (re-employed or unemployed) | 10.3%                                      | 0.0042               |
| Butterworth et al. (2011) – Poor quality job* |   | -18.7%                                     | -0.0071              |
| Galić et al. (2008)**                         | 166 re-employed individuals (of 372 unemployed at baseline) | 17.5%                                      | 0.0095               |

\*Negative signs indicate proportion of mental health state deteriorating and QALY loss, respectively.<sup>34</sup>

\*\* Results for Galić et al. (2008) are provided from the averaged results over the four time periods and the three SF-36 domains, included in their analysis and shown in Table 4.

**Table 8: SF-36 scores for shortlisted studies**

| Report                                | Domains of SF-36 used and follow-up period | Results  |
|---------------------------------------|--|--|
| Schuring (2010)                       | All eight domains: six-month follow-up     | Physical Function (PF) = +9.2<br>Role – Physical (RP) = +20<br>Bodily Pain (BP) = +11.3<br>General Health (GH) = +7<br>Vitality (VT) = +7.8<br>Social Functioning (SF)= +14.2<br>Role – Emotional (RE) = +22.7<br>Mental Health (MH) = +11 |
| Butterworth (2011) – good quality job | MHI-5: 12-month follow-up                  | Mental Health (MH) = +3.3  |

<sup>34</sup> More information on how the authors derived their definitions of a 'good quality' and 'bad quality' job can be found in the following reports:

1. Butterworth P, Leach LS, Strazdins L, Olesen SC, Rodgers B, Broom DH. The psychosocial quality of work determines whether employment has benefits for mental health: results from a longitudinal national household panel survey. *Occupational and environmental medicine*. 2011 Jan 1; oem-2010.
2. Leach LS, Butterworth P, Rodgers B, et al. Deriving an evidence-based measure of job quality from the HILDA Survey. *Aust Soc Pol J* 2010;9:67e86.
3. Butterworth P, Rodgers B, Windsor TD. Financial hardship, socio-economic position and depression: results from the PATH Through Life Survey. *Soc Sci Med* 2009;69:229e37.

| Report                                | Domains of SF-36 used and follow-up period   | Results  |
|---------------------------------------|--|--|
| Butterworth (2011) – poor quality job | MHI-5: 12-month follow-up  | Mental Health (MH) = -5.6  |
| Galić (2008)                          | <p>Averaged results for the three domains concerning general mental health at four different time periods (T1: baseline):</p> <p>T2: 1.5-year follow-up</p> <p>T2: 12-month follow-up</p> <p>(A) UU → E (T3 – T2)</p> <p>(B) UU → E (T3 – T1)</p> <p>(C) U → EE (T2 – T1)</p> <p>(D) U → EE (T3 – T1)</p> <p>UU → E n=53</p> <p>U → EE n=113</p> | <p>Avg. Social Functioning (SF), Role – Emotional (RE) and Mental Health (MH) =</p> <p>(A) = +7.1</p> <p>(B) = +2.90</p> <p>(C) = +2.60</p> <p>(D) = +1.60</p> |

\*Note U = unemployed, E = employed. Eg UU→E represents the cohort of people who were unemployed at T1 and T2 and re-employed in T3; U→EE represents the cohort of people who were unemployed in T1, re-employed in T2 and remained employed in T3.

The summaries provided below outline the distinction between the Schuring et al. (2010) findings and the findings of Butterworth et al. (2011) and Galić et al. (2008).

Schuring et al. (2010): measures the change in SF-36 scores of people who were unemployed and gained employment across all eight domains of the SF-36;

Butterworth et al. (2011): measures only the impact on mental health using the MHI-5, Butterworth et al. (2011) compares the impact of gaining a good-quality job with obtaining a poor-quality job. The study uses several different statistical methods to develop measures of the psychosocial characteristics of jobs using 12 items included in the Household, Income and Labour Dynamics in Australia (HILDA) survey.<sup>35</sup>

Galić et al. (2008): measures the wider impact on mental health by averaging the three domains specifically aimed at measuring general mental health; Social Functioning (SF), Role – Emotional (RE) and Mental Health MH.

Due to its study design, specifically the inclusion of all eight domains of the SF-36, findings using the Schuring et al. (2010) figures find a higher percentage of people crossing the threshold explained in Section 5.2.2, than would be the case for studies

<sup>35</sup> See Butterworth et al. (2010) report; 'Leach LS, Butterworth P, Rodgers B, et al. Deriving an evidence-based measure of job quality from the HILDA Survey'; and 'Butterworth P, Rodgers B, Windsor TD. Financial hardship, socio-economic position and depression: results from the PATH Through Life Survey,' for details regarding how the authors measure a good-quality/bad-quality job.

using only one or a combination of some of the eight domains. Galić uses the average of three domains of the SF-36, while Butterworth uses only the MHI-5 component of the SF-36, therefore, a smaller number of people are estimated to cross the threshold as less factors change within the calculations when using their findings. The size of the evidence base on the effect of job quality at present is unsure, although this would be an interesting area to explore in future research.

In the tool, only findings using Schuring et al. (2010) are presented due to its breadth in terms of capturing changes in all eight domains of the SF-36. With regards to the Galić et. al (2008) paper, concerns were raised with regards to its applicability to a UK context as the study is conducted in Croatia. As Butterworth et al. (2011) only used one domain of the SF-36, the MHI-5, it was felt that Schuring et al. (2010) captured the impact reflected in its results.

Fujiwara (2010) estimates that when an unemployed person, without a disability (somewhat of a proxy for those on JSA), moves into work, they incur £575 less in NHS costs per annum and with a disability (somewhat of a proxy for those on ESA), they incur £1,150 less in NHS costs per annum.<sup>36</sup> This is calculated by using an estimate of the average NHS cost per person of working age per annum, which was £1,220 in 2008 and applying this impact of employment on medical costs, while acknowledging that those who begin with a poorer health status are likely to incur higher initial costs to the NHS. Fujiwara (2010) measures health benefits but based on a study looking at the impact of redundancy on health, based on factory closure, which raises some questions about causality: ie presumably the health loss from redundancy is not equivalent to the health gain from employment. The findings, therefore, should provide a more accurate estimation of the benefits of moving from unemployment to employment.

In the analysis, the benefit to the health service per person whose mental health improves is equal to £1,246, without taking into consideration the potential differences of initial health status.

### 5.5.2. Financial outcomes

Alongside health outcomes and direct financial benefits, covered in the analysis through the literature review and data provided from DWP, a number of other impacts from a return to work could potentially have been included in the tool. These, and the reasons for not including them, are discussed below.

---

<sup>36</sup> Please note costs have been uplifted to 2016/17 prices (£508 and £1,016, respectively, at 2008 prices) using GDP deflators provided by the ONS.

## Welfare weight on income

Both HM Treasury's Green Book on appraisal and evaluation in central government<sup>37</sup> and Fujiwara (2010),<sup>38</sup> a Department for Work and Pensions (DWP) working paper on integrating social impacts into cost-benefit analysis for employment programmes, recommend including a welfare weight on the value of income to the individual. This refers to weighting the value of income received to an individual according to their income, reflecting the theory that any money received by someone with a lower income is worth more to them than the same amount provided to someone with a higher income.

Based on the evidence, the Green Book suggests "an extra £1 of consumption received by someone earning £10,000 a year will be worth twice as much as when it is paid to a person earning £20,000 per annum."<sup>39</sup> For those in the bottom income quintile, the Green Book suggests a weight of 1.9-2.0 times actual net income, while Fujiwara suggests a weighting of 2.5.

It was decided not to include this in the tool to avoid confusion with the actual financial figure received for individuals. However, the figures in the indicative case study below have been included.

## Reduction in crime resulting from reduced unemployment

Fujiwara (2010) recommends including the benefits of a reduction in crime due to reduced unemployment. It was noted that while the literature contains no conclusive evidence of a link between employment status and crime, there is an established link between income and crime (including literature from the UK).

Due to factors around the way data were presented, Fujiwara uses a US study to calculate the potential reduction in crime from employment for different age and gender groups, and values these using costs of crime published by the Home Office.

However, given the potential issues around generalisability of US findings, a desire to be conservative in estimates and in order to not distract from the impact on health, the

---

<sup>37</sup> Treasury HM. The green book: appraisal and evaluation in central government.

<https://www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-government>

<sup>38</sup> Fujiwara D. The Department for Work and Pensions social cost-benefit analysis framework. DWP Working Paper 86, London: Department for Work and Pensions; 2010.

[https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/214384/WP86.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/214384/WP86.pdf)

<sup>39</sup> Treasury HM. The Green Book: appraisal and evaluation in central government.

<https://www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-government>

primary focus of the development of the tool, crime impacts have not been included in the tool. They have, however, been considered in the indicative case study below.

### Substitution effects and social costs associated with funding programmes through taxation

Fujiwara (2010) suggests that substitution effects and social costs should theoretically be included in analysis, but that they are sensitive to a number of assumptions and estimates are varied and recommends inclusion only in sensitivity analysis for economic assessment of employment programmes. Specifically, substitution effects refer to the notion that employment programmes to get people into work may displace other workers; while social costs associated with funding programmes through taxation refers to the notion that increased taxation affects people's behaviour, and that funding employment programmes through taxation may have welfare consequences.

Neither of these elements is included in the tool due to this recommendation and the fact that the tool itself does not examine specific employment programmes, although the user is encouraged to consider the potential for substitution effects when entering the incremental proportion of people who return to work.

## 6. Economic tool

The tool begins with an introductory set of worksheets, providing background information on the project, its aims, a guide on how the tool hopes to facilitate decision making around worklessness and health, as well as a guide on how to navigate and use it. Note, although the tool is already populated with population data, prevalence data, the data extracted from Schuring (2010), and data provided by DWP, the user can make changes to the population data and prevalence of mental health conditions for their selected location. In addition, the user can include intervention information, specifically one-off and ongoing costs of an intervention, which allows the tool to calculate Return on Investment (ROI) metrics such as the incremental cost-effectiveness ratio (ICER) and Benefit-Cost Ratio from relevant different perspectives. Below is a more detailed description of the tools inputs and outputs.

### 6.1. Tool inputs

#### Population sheet

To begin, a geographical area needs to be selected in the Population sheet. This can be a CCG, a STP, a combined authority, a district/unitary authority, a county/unitary authority, a region, or England as a whole. The working age population is then provided for the selected area, alongside the population of interest, namely the proportion (and the number of people) claiming Jobseeker's Allowance, Employment and Support Allowance Work Related Activity Group (ESA WRAG) or the Employment and Support Allowance Support Group (ESA SG). Once population information is populated, which the user can override with updated figures, the user must then specify the proportion of the population of interest that is assumed to return to work.

The decision to use JSA, ESA WRAG & ESA SG claimants as the population of interest, as opposed to just the unemployed population as a whole, was due to the fact that those claiming ESA WRAG are deemed to be in a position where they can work towards returning to work. Although this is not the case for the ESA SG, there is recognition that many ESA SG claimants would like gainful employment, and thus this analysis may be relevant depending on the extent of support provided by an intervention aimed at helping people get back to work, and the timeline over which this can occur.

#### Mental Health Conditions sheet

In the Mental Health Conditions sheet, the percentage of the working age population with Common Mental Health Disorders (CMDs) is provided. This value can be overridden. The breakdown of CMDs into six categories with the respective number of

people experiencing each condition is provided for information purposes. A person may experience more than one condition simultaneously; as such, their sum exceeds the overall number of people with CMDs. A weighted annual unit cost of CMDs is provided, based on the annual unit costs of depression and anxiety, adjusted for the proportion of people who access services, as estimated by McCrone et al. (2008)<sup>40</sup> and adjusted for 2016/2017 prices.

### Return to work sheet

The number of people returning to work who have CMDs is provided, based on the user's inputs from previous sheets. This number has been calculated by using the prevalence ratio of CMDs of the working age population. Since the prevalence ratio of CMDs is likely to be higher among JSA and ESA-WRAG claimants, this number is likely to be an underestimate. The user is requested to enter the duration of the expected benefit, namely the amount of time the average individual spends in work when they would otherwise have been unemployed/economically inactive. Finally, this sheet provides users with the option to enter the total cost, ie the one-off fixed cost and the ongoing annual cost, that is required to achieve the return to work and which allows ROI metrics to be calculated.

### Other Inputs sheet

In the Other Inputs sheet, the user can input a customised monetary value for a QALY gain or keep the default of £60,000 (as used by PHE), which represents the willingness to pay per QALY value gained. Finally, differential discounting is proposed for the health effects and for all the other financial costs and benefits. The default values are set at 1.5% and 3.5% respectively, as recommended by PHE.

## 6.2. Tool Outputs

### Direct health benefits sheet

The tool provides per person and total QALY gains and their respective value and calculates the proportion of people whose mental health status improves due to gaining employment. In addition, the cost savings from a reduction in the number of people facing CMDs is provided, as well as the total health value, ie the value of the QALY gains and the cost savings to the healthcare system.

---

<sup>40</sup> McCrone PR, Dhanasiri S, Patel A, Knapp M, Lawton-Smith S. Paying the price: the cost of mental health care in England to 2026. King's Fund; 2008.

## Financial benefits sheet

The tool provides the benefits to the exchequer, the benefits to the programme participant(s), and the benefits to society, broken down by category (eg increase in income tax, increase in wages, reduction of housing benefit payments etc.). The benefits to the exchequer consist of benefits to the national government, benefits to the local authority (excluding health and social care cost savings) and benefits to the health and care services.

Health and social care benefits are broken down by the agency they accrue to using data provided by New Economy, (a trading arm of the Greater Manchester Combined Authority which provides policy, strategy and research advice). The other reported financial figures are based on DWP estimations and are rounded to the nearest 10 where applicable. The values of Council Tax Benefit and Free School Meals payments can be overridden to reflect area- and person-specific values respectively. Housing benefits presented in the tool have been localised by local authority, using a tool provided by DWP, and mapped for each population area provided (eg CCG, STP, etc.). It is an average weighted by the proportion of people returning to work from each claimant group – JSA, ESA WRAG and ESA SG, and provides more accurate figures for change in housing benefits due to gainful employment. Note that while Universal Credit is not included in the tool, UC claimants who are required to seek work are included in the model as JSA claimants.

Total healthcare cost savings are calculated by multiplying the cost of CMDs with the number of people who are no longer experiencing CMDs, as estimated by Schuring et al. (2010). The benefits to society are provided using local economy average and health specific multipliers, as estimated by the Scottish Government (2016).<sup>41</sup>

## Key findings sheet

This sheet provides a summary of the reported results, and if the user has provided an intervention cost in the Other Inputs sheet, the model will report Return on Investment (ROI) metrics, namely: the Incremental Cost-Effectiveness Ratio (ICER), the Net Benefit, and the Benefit-Cost Ratio.

## Charts sheet

Finally, charts, providing graphical representations of key outputs, include the breakdown of total benefits to the exchequer and per participant due to the movement

---

<sup>41</sup> Scottish Government. Input-Output Classification Table SIC(2007). Available at <http://www.gov.scot/Topics/Statistics/Browse/Economy/Input-Output/Downloads/IOC098-SIC07>.

to employment, the breakdown of total economic value for society and the breakdown of healthcare cost savings by agency from the inputs selected in the previous relevant sheets in the tool.

### 6.3. Indicative case study

As an example of a scenario run through the model, a decision-maker in Bedford may choose to implement an intervention targeted at the population of interest; the JSA, ESA WRAG and ESA SG population within Bedford which amounts to 6,287 people. This decision-maker must then select the proportion of this population of interest that successfully returns to work. As an example, it is assumed that the user selects 5% of their population of interest to return to work, ie 314 people. With all the default figures unchanged, the prevalence ratio of people with common mental health disorders (CMDs) is 18.9%.

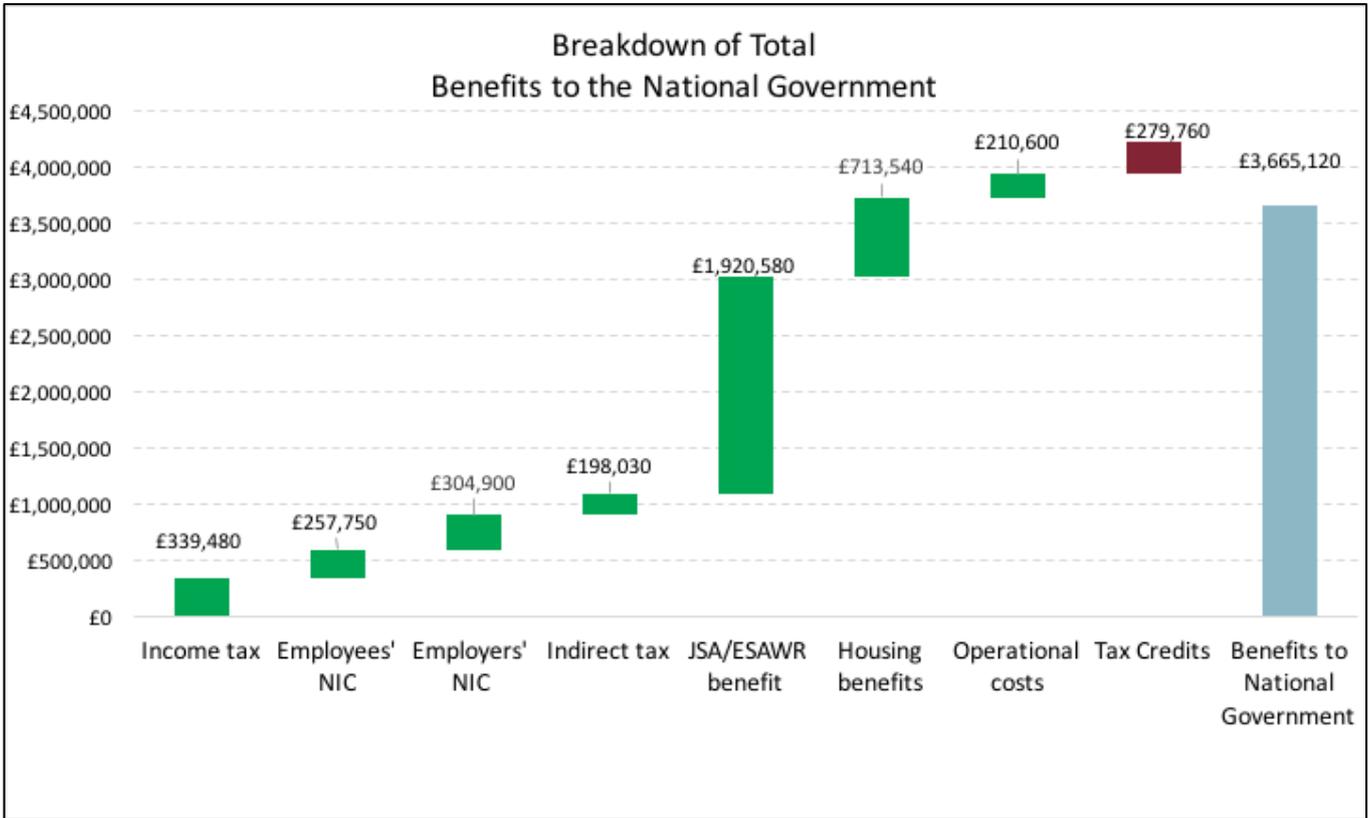
In this example, the user selects that the amount of time the average individual spends in work when they would otherwise have been unemployed/economically inactive is one year. To achieve re-employment the user faces a one-off fixed cost of £2.5m (£7953 per person) and the default figures for QALY value and discount rate are kept unchanged.

Using results derived from Schuring (2010), the per person QALY gain is 0.0675 QALYs (equivalent to £4,100 per person in value if the value of a QALY is £60,000), and 28 people (8.8% of those returning to work, 46.6% of those returning to work who had a CMD(s) at baseline) in the selected return to work cohort who had a CMD(s) and improved due to employment, leading to an £85 financial gain to the health services per person returning to work.

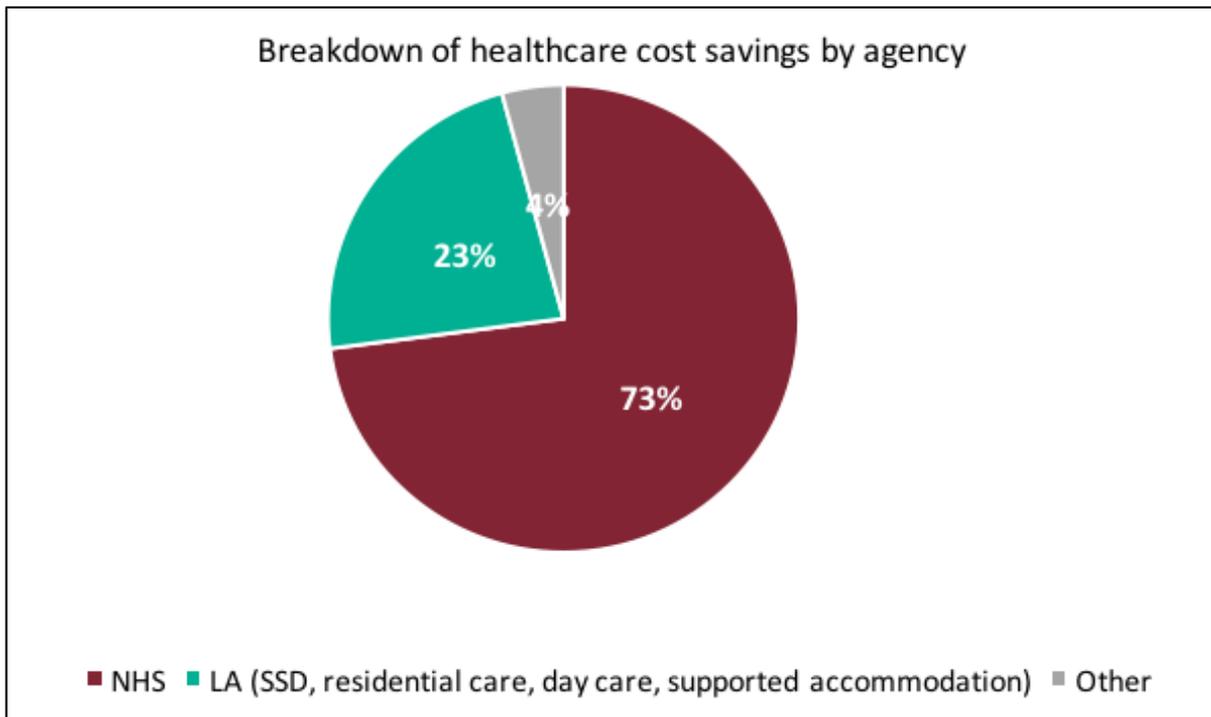
Based on the person returning to employment, there is, per person returning to work, a £3,500 financial gain to the individual, a £500 financial gain to the local authority and £11,700 savings to national government. The overall gain to society (excluding transfers between individual and government) and based on a local economy multiplier of 1.6 and 1.66 health-specific multiplier, is £23,400 per person returning to work.

If the user inputs a one-off fixed cost of intervention £2.5m (£7953 per person) and ongoing annual cost of £100,000 (£318 per person) to achieve this change, the ROI metrics would show from the public sector perspective a 'dominant' ICER (ie the intervention saves money and delivers a positive QALY gain), a total Net Benefit of £1.4m and a Benefit-Cost Ratio of £1.54 per £1 spent.

**Figure 6: Indicative case study: breakdown of total benefits to exchequer for 2016/2017**



**Figure 7 – Indicative example: breakdown of healthcare cost savings by agency**



## Other analysis elements not included in the tool

As detailed in Section 5.5, the tool does not include figures on the welfare weight of an increase in income, or the benefits of crime reduction due to a return to work. Were these to be included, this is how they would add to the benefits included for the Bedford example above.

### Welfare weight

The net financial gain to the average individual equates to £3,500. With the assumption that the average unemployed individual is in the bottom quintile of the population with regards to income, the HM Treasury Green Book suggests a weighting of 1.9-2.0 times the financial value, and Fujiwara (2010) suggests a weighting of 2.5. This would lead to the value to the individual being equivalent to £6,700-£7,100 or £8,900 respectively.

### Crime reduction

Fujiwara (2010) recommends calculating the percentage reduction in crime by multiplying the percentage increase in income by 0.6. Given the average individual has a net income increase of £3,500, and a net income of £14,500 while in work, this means net income prior to returning to work was £11,000. Income has, therefore, increased by 32%, and crime has reduced by 19%.

Fujiwara then provides the cost of crime per relevant age and gender group, which can be multiplied by the reduction in crime to calculate the total savings through crime reduction. The cost represents the cost of crime in a value sense, and includes such factors as the value of pain and suffering caused by crime. From a financial perspective, only approximately 19% of these figures relate to fiscal costs.

Fujiwara suggests that for male employment programme participants, aged 17-24, the average cost of crime is £5,170. For males 25+, it is £2,610. For females, aged 17-24, it is £1,250, and for females aged 25+, it is £444.

Based on figures for Bedford for overall population, from Office of National Statistics (ONS)<sup>42</sup>, 9% of the working age population are males 16-24<sup>43</sup>; 41% are males 25+; 8% are females 16-24, and 42% are females 25+. This gives a weighted cost of crime at baseline of £1,800 per person, and an average reduction in crime costs of £400 per person due to a return to work.

---

<sup>42</sup> Office of National Statistics (ONS), Population estimates for high level areas, 2015

<sup>43</sup> Figures for 16-24 year olds were available rather than 17-24 year olds. It has been assumed that the effect is the same.

## 7. Conclusion

This review and analysis has shown that there are significant mental health benefits and consequent financial savings from returning to work, as well as direct financial benefits, and has provided a tool to facilitate the inclusion of these findings in decision making at the local level.

The tool allows users to understand, for a given number of people who return to work, the health benefits and financial benefits to a variety of stakeholders. This can also be compared against an intervention cost to understand if the intervention is cost saving (financial benefits outweigh the cost) and/or cost-effective (the cost per QALY gained is below a given threshold).

The evidence incorporated within the tool was based on a rapid evidence assessment of available research. This picked up an array of recent research on the topic of the health benefits of returning to work, but was not as comprehensive as a full systematic review. For instance, papers before 2006, or studies conducted outside of the EU countries, Switzerland, Canada, New Zealand, Australia, Norway or the USA, were not included. Citations from before 2006, such as those mentioned in Waddell & Burton's overview of the evidence from 2006, were also not included.

Much of the current evidence on the health benefits of employment is found in evaluations of particular interventions. While this is valuable research and should be encouraged, non-intervention specific research can also help separate out the impact of an intervention versus the impact of employment more generally. More research, examining both the quality of life impact to individuals and the actual cost savings resulting from a return to work in the form of longitudinal studies with matched control groups, is recommended.

The evidence search also clearly highlighted the scarcity of evidence around the potential benefits of employment on physical health. Of the few reports that included the impact of transitioning from unemployment to employment on physical health, a number showed some level of (non-significant) improvement. However, none of the studies focused primarily on physical health. Consequently, due to the low level of robust evidence around physical health impacts, these potential benefits were omitted from the tool. More research, examining the effects of employment, including the transition from unemployment to employment, on physical health, is advocated.

In addition, it would be interesting to explore how various variables impact on the health benefits of returning to work, such as those highlighted in the original research questions that were not possible to answer on current evidence, including duration of unemployment, health conditions suffered by the individual, and reason for unemployment (eg whether or not it is health-related).

Further research to build on the Butterworth study, with regard to investigating whether health benefits are dependent on the type or quality of job, is also encouraged. A more comprehensive evidence base could help determine whether or not, and the extent, to which type of quality of job affects wellbeing. Additionally, if job quality was better defined in the literature, this would help policy makers focus efforts on creating more of those types of jobs and/or steering those with health conditions towards them.

There are other benefits to employment, such as a reduction in crime, which was not included due to uncertainty around the estimates but which further research could develop.

Ultimately, the tool, combined with other available evidence around the benefits of employment, equips decision-makers with the evidence needed to support the case for better investment in areas aimed at alleviating worklessness. Better investment will likely lead to better outcomes for the individuals transitioning into employment, their families and their wider local communities, leading to an overall more healthy and inclusive society.

## 8. Bibliography

Andersen H H, Mühlbacher A, Nübling M, Schupp J, Wagner G G. Computation of standard values for physical and mental health scale scores using the SOEP version of SF-12v2. *Schmollers Jahrbuch*. 2007;127(1):171-82.

Ara R, Brazier J. Predicting the short form-6D preference-based index using the eight mean short form-36 health dimension scores: estimating preference-based health-related utilities when patient level data are not available. *Value in Health*. 2009 Mar 1;12(2):346-53.

Audhoe S S, Hoving J L, Sluiter J K, Frings-Dresen M H. Vocational interventions for unemployed: effects on work participation and mental distress. A systematic review. *Journal of occupational rehabilitation*. 2010 Mar 1;20(1):1-3.

Ayala L, Rodríguez M. Health-related effects of welfare-to-work policies. *Social Science & Medicine*. 2013 Sep 30;93:103-12.

Beyer S, Brown T, Akandi R, Rapley M. A comparison of quality of life outcomes for people with intellectual disabilities in supported employment, day services and employment enterprises. *Journal of Applied Research in Intellectual Disabilities*. 2010 May 1;23(3):290-5.

Böckerman P, Ilmakunnas P. Unemployment and self-assessed health: evidence from panel data. *Health economics*. 2009 Feb 1;18(2):161-79.

Bowling A, Bond M, Jenkinson C, Lamping D L. Short Form 36 (SF-36) Health Survey questionnaire: which normative data should be used? Comparisons between the norms provided by the Omnibus Survey in Britain, the Health Survey for England and the Oxford Healthy Life Survey. *Journal of Public Health*. 1999 Sep 1;21(3):255-70.

Bush P W, Drake R E, Xie H, McHugo G J, Haslett W R. The long-term impact of employment on mental health service use and costs for persons with severe mental illness. *Psychiatric Services*. 2009 Aug;60(8):1024-31.

Butterworth P, Leach L S, Strazdins L, Olesen S C, Rodgers B, Broom D H. The psychosocial quality of work determines whether employment has benefits for mental health: results from a longitudinal national household panel survey. *Occupational and environmental medicine*. 2011 Jan 1:oem-2010.

Carlier B E, Schuring M, Lötters F J, Bakker B, Borgers N, Burdorf A. The influence of re-employment on quality of life and self-rated health, a longitudinal study among unemployed persons in the Netherlands. *BMC Public Health*. 2013 May 24;13(1):503.

Crowther R, Marshall M, Bond G, Huxley P. Vocational rehabilitation for people with severe mental illness. *Cochrane Database Syst Rev*. 2001;2(2).

Curnock E, Leyland A H, Popham F. The impact on health of employment and welfare transitions for those receiving out-of-work disability benefits in the UK. *Social Science & Medicine*. 2016 Aug 31;162:1-0.

- Ferreira J A, Reitzle M, Lee B, Freitas R A, Santos E R, Alcoforado L, Vondracek F W. Configurations of unemployment, re-employment, and psychological wellbeing: A longitudinal study of unemployed individuals in Portugal. *Journal of Vocational Behavior*. 2015 Dec 31;91:54-64.
- Galić Z, Šverko B. Effects of prolonged unemployment and re-employment on psychological and physical health. *Review of psychology*. 2008 Dec 16;15(1-2):3-10.
- Gebel M, Voßemer J. The impact of employment transitions on health in Germany. A difference-in-differences propensity score matching approach. *Social Science & Medicine*. 2014 May 31;108:128-36.
- Grassi M, Nucera A. Dimensionality and summary measures of the SF-36 v1. 6: comparison of scale-and item-based approach across ECRHS II adults population. *Value in health*. 2010 Jun 1;13(4):469-78.
- Hao Y. Productive activities and psychological wellbeing among older adults. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*. 2008 Mar 1;63(2):S64-72.
- Hoare P N, Machin M A. The impact of re-employment on access to the latent and manifest benefits of employment and mental health. *Journal of Occupational and Organizational Psychology*. 2010 Sep 1;83(3):759-70.
- Huber M, Lechner M, Wunsch C. Does leaving welfare improve health? Evidence for Germany. *Health Economics*. 2011 Apr 1;20(4):484-504.
- Jacobs A W, Hill T D, Tope D, O'Brien L K. Employment Transitions, Child Care Conflict, and the Mental Health of Low-Income Urban Women With Children. *Women's Health Issues*. 2016 Aug 31;26(4):366-76.
- Kelly M J, Dunstan F D, Lloyd K, Fone D L. Evaluating cutpoints for the MHI-5 and MCS using the GHQ-12: a comparison of five different methods. *BMC Psychiatry*. 2008 Feb 19;8(1):10.
- McCrone P R, Dhanasiri S, Patel A, Knapp M, Lawton-Smith S. Paying the price: the cost of mental health care in England to 2026. King's Fund; 2008.
- McManus S, Bebbington P, Jenkins R, Brugha T. (eds.) (2016) *Mental health and wellbeing in England: Adult Psychiatric Morbidity Survey 2014*. Leeds: NHS Digital.
- NOMIS, Official Labour Market Statistics. (2017). Benefit claimants - Employment and Support Allowance  
<https://www.nomisweb.co.uk/query/construct/summary.asp?mode=construct&version=0&datasheet=134>
- Office for National Statistics. (2017). CC01: Regional labour market: Claimant Count by unitary and local authority (experimental).  
<https://www.ons.gov.uk/employmentandlabourmarket/peoplenotinwork/unemployment/datasets/claimantcountbyunitaryandlocalauthorityexperimental/current>

Office for National Statistics. (2017). National and Local Authority Level Population Estimates

Reine I, Novo M, Hammarström A. Does transition from an unstable labour market position to permanent employment protect mental health? Results from a 14-year follow-up of school-leavers. *BMC Public Health*. 2008 May 13;8(1):159.

Ritvo P, Fischer J S, Miller D M, Andrews H, Paty D W, LaRocca N G. MSQLI – Multiple Sclerosis Quality of Life Inventory. A user's manual. New York: National MS Society. 1997.

Schneider J, Boyce M, Johnson R, Secker J, Slade J, Grove B, Floyd M. Impact of supported employment on service costs and income of people with mental health needs. *Journal of Mental Health*. 2009 Dec 1;18(6):533-42.

Schuring M, Mackenbach J, Voorham T, & Burdorf A (2010). The effect of re-employment on perceived health. *Journal of Epidemiology & Community Health*, jech-2009.

Steele F, French R, Bartley M. Adjusting for selection bias in longitudinal analyses using simultaneous equations modeling: the relationship between employment transitions and mental health. *Epidemiology*. 2013 Sep 1;24(5):703-11.

Thomas C, Benzeval M, Stansfeld S. Psychological distress after employment transitions: the role of subjective financial position as a mediator. *Journal of Epidemiology & Community Health*. 2007 Jan 1;61(1):48-52.

van der Noordt M, IJzelenberg H, Droomers M, Proper K I. Health effects of employment: a systematic review of prospective studies. *Occup Environ Med*. 2014 Feb 20:oemed-2013.

van Rijn R M, Carlier B E, Schuring M, Burdorf A. Work as treatment? The effectiveness of re-employment programmes for unemployed persons with severe mental health problems on health and quality of life: a systematic review and meta-analysis. *Occup Environ Med*. 2016 Apr 1;73(4):27

## 9. Appendix

### A. Search strategy & database search results

The search strategy below was used for the Embase database and consequently adapted for the other databases searched. The search was carried out by Nicola Pearce-Smith from the PHE Knowledge and Library Services.

**Database:** Embase <1974 to 2016 December 06>

**Search Strategy:**

- 
1. (work or employment or re-employment or reemployment or employed or re-employed or reemployed or RTW).tw.
  2. employment/ or employment, supported/ or return to work/
  3. 1 or 2
  4. (unemployment or unemployed or out-of-work or worklessness or economic\* inactive).tw.
  5. unemployment/
  6. 4 or 5
  7. 3 and 6
  8. (employment transition\* or unemployment transition\* or labor market status transition\* or welfare transition\*).tw.
  9. 7 or 8
  10. (health or ill-health or wellbeing or disease\* or disorders or complaints or psychosocial or psychological or sleep or mental or depression or anxiety or stress or distress or ageing or aging or quality of life or mortality or morbidity or sickness).tw.
  11. mental health/ or mental disorders/ or depression/ or depressive disorder/ or anxiety/
  12. (government\* adj3 gain\*).tw.
  13. (benefits or savings or taxation).tw.
  14. (health\* adj3 cost\*).tw.
  15. (income or incomes).tw.
  16. INSURANCE BENEFITS/ or Income/ or Taxes/ or Health Care Costs/
  17. Socioeconomic Factors/
  18. financial outcome\*.tw.
  19. social outcome\*.tw.
  20. (crime or offending or reoffend\* or re-offend\* or recidivis\*).tw.
  21. crime/ or fraud/ or homicide/ or theft/ or violence/
  22. housing/ or public housing/
  23. (house\* or housing).tw.
  24. exp Life Style/
  25. (smoking or smoke\*).tw.
  26. Smoking/

- 27. Smoking Cessation/
- 28. (exercise or physical activit\*).tw.
- 29. exercise/ or running/ or swimming/ or walking/
- 30. Caregivers/
- 31. caregiv\*.tw.
- 32. lifestyle factor\*.tw.
- 33. 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32
- 34. 9 and 33
- 35. ((work or employment or re-employment or reemployment or employed or re-employed or reemployed or RTW or employment transition\* or unemployment transition\* or labo?r market status transition\* or welfare transition\* or unemployment or unemployed or out-of-work or worklessness or economic\* inactive) and health\* and (impact\* or gain\* or benefit\* or effect\*)).ti.
- 36. limit 34 to "reviews (best balance of sensitivity and specificity)"
- 37. longitudinal studies/ or prospective studies/
- 38. (longitudinal or prospective).tw.
- 39. 37 or 38
- 40. 34 and 39
- 41. 35 or 36 or 40

-----

**Results from the database searches:**

| <b>Source</b>                     | <b>Results</b>                  |
|-----------------------------------|---------------------------------|
| Medline                           | 1296 (1102 after deduplication) |
| Embase                            | 2991 (1985 after deduplication) |
| PsycINFO                          | 1503 (1045 after deduplication) |
| <b>Total number of references</b> | <b>4132 unique references</b>   |

## B. Call for evidence

### Call for evidence: Estimation of benefits from moving an individual from unemployment into sustainable employment

Dear colleagues,

**We are issuing a call for evidence on the benefits that come from moving from unemployment (or economic inactivity) to sustainable employment, specifically in terms of:**

- Health benefits to the individual
- Health and social care utilisation
- Cost savings to government
- Financial benefits to the individual and the local economy
- Any intermediate factor that may lead to one or more of the above

This is to support work commissioned by Public Health England (PHE) to undertake a review of the evidence of the health, financial and other social impacts of a move from worklessness to work. The aim of this project is to assess **the cost savings that result to the public purse from employment itself, and improved health outcomes**. This work will not look at the effectiveness of interventions to get people into work. It will attempt to show the benefits that can result beyond individual remuneration and productivity.

We are interested in receiving references to evidence in any format which is permitted within the UK copyright law including: links to websites, abstracts, or references for published articles, as well as excel files or links to any tool available online. If you wish to submit commercially or academically sensitive information, unpublished information or research, please highlight any sections that are confidential or sensitive.

Please send any relevant details by 13<sup>th</sup> January 2017 to:

Obinna Onwude

[obinna.onwude@optimityadvisors.com](mailto:obinna.onwude@optimityadvisors.com)

**or**

Rory Tierney

[rory.tierney@optimityadvisors.com](mailto:rory.tierney@optimityadvisors.com)

Optimity Advisors  
1<sup>st</sup> Floor Kemp House  
152-160 City Road  
London EC1V 2NP  
t: +44 (0) 20 7553 4800

We look forward to receiving the information and thank you in advance for your help.

Yours sincerely,  
Obinna Onwude  
Senior Consultant

## C. Inclusion / exclusion criteria

**Table 9: Full inclusion / exclusion criteria**

|           | <b>Criteria</b>   | <b>Inclusion</b>                           | <b>Exclusion Code</b>         | <b>Notes</b>   |
|-----------|---|--|-------------------------------|--|
| <b>C1</b> | <b>Date</b><br>Studies published in 2006 or later.  | If YES or UNCLEAR, move to next criterion. | If not<br><b>1_EX.DATE</b>    | Exclude studies published before 2006.   |
| <b>C2</b> | <b>Country</b><br>Studies conducted in EU countries, and any of Switzerland, Canada, New Zealand, Australia, Norway, USA. | If YES or UNCLEAR, move to next criterion. | If not<br><b>2_EX.COUNTRY</b> | Exclude reports not related to one or more of the following countries: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, United Kingdom, Switzerland, Canada, New Zealand, Australia, Norway and USA. |
| <b>C3</b> | <b>Language</b><br>Only papers published in English.  | If YES or UNCLEAR, move to next criterion. | If not,<br><b>3_EX.LANG</b>   | Exclude all non-English papers.  |
| <b>C4</b> | <b>Topic</b><br>Studies focused on the health effects of employment.  | If YES or UNCLEAR, move to next criterion. | If not <b>4_EX.TOPIC</b>      | Exclude papers that do not focus on the health effects of employment.  |

Estimation of benefits from moving an individual from unemployment into sustainable employment

|    |  |  |   |  |
|----|--|--|---|--|
| C5 | <p><b>Population</b><br/>Include papers that report on adult populations or sub-populations moving from worklessness to work (16-75).</p>  | If YES or UNCLEAR, move to next criterion. | If not <b>5_EX.POP</b>  | Exclude papers that do not focus on adult populations or sub-populations moving from worklessness to work.   |
| C6 | <p><b>Interventions</b><br/>General population interventions aimed at moving people from worklessness to work.</p>   | If YES or UNCLEAR, move to next criterion. | If not <b>6_EX.INT</b>  | Exclude papers on specific interventions targeted at improving health or targeted at health conditions to enable people to return to work.   |
| C7 | <p><b>Outcomes</b><br/>Only studies that report relevant health (mental or physical health) or financial benefits.</p>   | If YES or UNCLEAR, move to next criterion. | If not <b>7_EX.OUT</b>  | Exclude studies that do not report outcomes related to the relevant outcomes.  |
| C8 | <p><b>Data</b><br/>Only studies that report empirical/quantitative data.</p>   | If YES or UNCLEAR, move to next criterion. | If not <b>8_EX.DATA</b>   | Exclude studies that do not report empirical/quantitative data. These might include commentaries, editorials, position papers, policy papers etc.  |
| C9 | <p><b>Inclusion codes</b><br/><b>9_IN.HEALTH – health impact/outcome studies</b><br/><b>10_IN.FIN – financial impact studies</b><br/><b>11_IN.OTHER – impact on other sectors eg justice, housing etc.</b></p> |  | <p>Include studies that meet the inclusion criteria, please distinguish according to their main focus:<br/><b>(9_IN.HEALTH)</b><br/><b>(10_IN.FIN)</b><br/><b>(11_IN.OTHER)</b></p> | <p>Reports should be about impact on health, financial benefits and impact on other sectors</p> <p>Systematic reviews that include any relevant study listed above will be identified and reported separately.</p> |

## D. Quality assessment tool



# Quality Assessment Tool for Quantitative Studies Dictionary<sup>44</sup>

The purpose of this dictionary is to describe items in the tool thereby assisting raters to score study quality. Due to under-reporting or lack of clarity in the primary study, raters will need to make judgements about the extent that bias may be present. When making judgements about each component, raters should form their opinion based upon information contained in the study rather than making inferences about what the authors intended.

### SELECTION BIAS

(Q1) Participants are more likely to be representative of the target population if they are randomly selected from a comprehensive list of individuals in the target population (score very likely). They may not be representative if they are referred from a source (eg clinic) in a systematic manner (score somewhat likely) or self-referred (score not likely).

(Q2) Refers to the % of subjects in the control and intervention groups that agreed to participate in the study before they were assigned to intervention or control groups.

### STUDY DESIGN

In this section, raters assess the likelihood of bias due to the allocation process in an experimental study. For observational studies, raters assess the extent that assessments of exposure and outcome are likely to be independent. Generally, the type of design is a good indicator of the extent of bias. In stronger designs, an equivalent control group is present and the allocation process is such that the investigators are unable to predict the sequence.

#### Randomized Controlled Trial (RCT)

An experimental design where investigators randomly allocate eligible people to an intervention or control group. A rater should describe a study as an RCT if the randomization sequence allows each study participant to have the same chance of receiving each intervention and the investigators could not predict which intervention was next. If the investigators do not describe the allocation process and only use the words 'random' or 'randomly', the study is described as a controlled clinical trial.

See below for more details.

*Was the study described as randomized?*

Score YES, if the authors used words such as random allocation, randomly assigned, and random assignment. Score NO, if no mention of randomization is made.

*Was the method of randomization described?*

Score YES, if the authors describe any method used to generate a random allocation sequence.

Score NO, if the authors do not describe the allocation method or describe methods of allocation such as alternation, case record numbers, dates of birth, day of the week, and any allocation procedure that is entirely transparent before assignment, such as an open list of random numbers of assignments.

If NO is scored, then the study is a controlled clinical trial.

---

<sup>44</sup> Effective Public Health Practice Project. [Cited 28 July 2017.] Quality Assessment Tool for Quantitative Studies. Available from <http://www.ehpp.ca/tools.html>

*Was the method appropriate?*

Score YES, if the randomization sequence allowed each study participant to have the same chance of receiving each intervention and the investigators could not predict which intervention was next. Examples of appropriate approaches include assignment of subjects by a central office unaware of subject characteristics, or sequentially numbered, sealed, opaque envelopes.

Score NO, if the randomization sequence is open to the individuals responsible for recruiting and allocating participants or providing the intervention, since those individuals can influence the allocation process, either knowingly or unknowingly.

If NO is scored, then the study is a controlled clinical trial.

**Controlled Clinical Trial (CCT)**

An experimental study design where the method of allocating study subjects to intervention or control groups is open to individuals responsible for recruiting subjects or providing the intervention. The method of allocation is transparent before assignment, eg an open list of random numbers or allocation by date of birth, etc.

**Cohort analytic (two group pre and post)**

An observational study design where groups are assembled according to whether or not exposure to the intervention has occurred. Exposure to the intervention is not under the control of the investigators. Study groups might be non-equivalent or not comparable on some feature that affects outcome.

**Case control study**

A retrospective study design where the investigators gather 'cases' of people who already have the outcome of interest and 'controls' who do not. Both groups are then questioned or their records examined about whether they received the intervention exposure of interest.

**Cohort (one group pre + post (before and after))**

The same group is pretested, given an intervention, and tested immediately after the intervention. The intervention group, by means of the pretest, act as their own control group.

**Interrupted time series**

A time series consists of multiple observations over time. Observations can be on the same units (eg individuals over time) or on different but similar units (eg student achievement scores for particular grade and school). Interrupted time series analysis requires knowing the specific point in the series when an intervention occurred.

**CONFOUNDERS**

By definition, a confounder is a variable that is associated with the intervention or exposure and causally related to the outcome of interest. Even in a robust study design, groups may not be balanced with respect to important variables prior to the intervention. The authors should indicate if confounders were controlled in the design (by stratification or matching) or in the analysis. If the allocation to intervention and control groups is randomized, the authors must report that the groups were balanced at baseline with respect to confounders (either in the text or a table).

**BLINDING**

(Q1) Assessors should be described as blinded to which participants were in the control and intervention groups. The purpose of blinding the outcome assessors (who might also be the care providers) is to protect against detection bias.

(Q2) Study participants should not be aware of (ie blinded to) the research question. The purpose of blinding the participants is to protect against reporting bias.

**DATA COLLECTION METHODS**

Tools for primary outcome measures must be described as reliable and valid. If 'face' validity or 'content' validity has been demonstrated, this is acceptable. Some sources from which data may be collected are described below:

Self reported data includes data that is collected from participants in the study (eg completing a questionnaire, survey, answering questions during an interview, etc.).

Assessment/Screening includes objective data that is retrieved by the researchers. (eg observations by investigators).

Medical Records/Vital Statistics refers to the types of formal records used for the extraction of the data.

Reliability and validity can be reported in the study or in a separate study. For example, some standard assessment tools have known reliability and validity.

#### WITHDRAWALS AND DROP-OUTS

Score YES if the authors describe BOTH the numbers and reasons for withdrawals and drop-outs. Score NO if either the numbers or reasons for withdrawals and drop-outs are not reported.

The percentage of participants completing the study refers to the % of subjects remaining in the study at the final data collection period in all groups (ie control and intervention groups).

#### INTERVENTION INTEGRITY

The number of participants receiving the intended intervention should be noted (consider both frequency and intensity). For example, the authors may have reported that at least 80% of the participants received the complete intervention. The authors should describe a method of measuring if the intervention was provided to all participants the same way. As well, the authors should indicate if subjects received an unintended intervention that may have influenced the outcomes. For example, co-intervention occurs when the study group receives an additional intervention (other than that intended). In this case, it is possible that the effect of the intervention may be over-estimated.

Contamination refers to situations where the control group accidentally receives the study intervention. This could result in an under-estimation of the impact of the intervention.

#### ANALYSIS APPROPRIATE TO QUESTION

Was the quantitative analysis appropriate to the research question being asked?

An intention-to-treat analysis is one in which all the participants in a trial are analyzed according to the intervention to which they were allocated, whether they received it or not. Intention-to-treat analyses are favoured in assessments of effectiveness as they mirror the noncompliance and treatment changes that are likely to occur when the intervention is used in practice, and because of the risk of attrition bias when participants are excluded from the analysis.

#### **Component Ratings of Study:**

For each of the six components A – F, use the following descriptions as a roadmap.

##### **SELECTION BIAS**

Strong: The selected individuals are very likely to be representative of the target population (Q1 is 1) and there is greater than 80% participation (Q2 is 1).

Moderate: The selected individuals are at least somewhat likely to be representative of the target population (Q1 is 1 or 2); and there is 60-79% participation (Q2 is 2). 'Moderate' may also be assigned if Q1 is 1 or 2 and Q2 is 5 (can't tell).

Weak: The selected individuals are not likely to be representative of the target population (Q1 is 3); or there is less than 60% participation (Q2 is 3) or selection is not described (Q1 is 4); and the level of participation is not described (Q2 is 5).

##### **DESIGN**

Strong: will be assigned to those articles that described RCTs and CCTs.

Moderate: will be assigned to those that described a cohort analytic study, a case control study, a cohort design, or an interrupted time series.

Weak: will be assigned to those that used any other method or did not state the method used.

### **CONFOUNDERS**

Strong: will be assigned to those articles that controlled for at least 80% of relevant confounders (Q1 is 2); or (Q2 is 1). Moderate: will be given to those studies that controlled for 60-79% of relevant confounders (Q1 is 1) and (Q2 is 2). Weak: will be assigned when less than 60% of relevant confounders were controlled (Q1 is 1) and (Q2 is 3) or control of confounders was not described (Q1 is 3) and (Q2 is 4).

### **BLINDING**

Strong: The outcome assessor is not aware of the intervention status of participants (Q1 is 2); and the study participants are not aware of the research question (Q2 is 2).

Moderate: The outcome assessor is not aware of the intervention status of participants (Q1 is 2); or the study participants are not aware of the research question (Q2 is 2); or blinding is not described (Q1 is 3 and Q2 is 3).

Weak: The outcome assessor is aware of the intervention status of participants (Q1 is 1); and the study participants are aware of the research question (Q2 is 1).

### **DATA COLLECTION METHODS**

Strong: The data collection tools have been shown to be valid (Q1 is 1); and the data collection tools have been shown to be reliable (Q2 is 1).

Moderate: The data collection tools have been shown to be valid (Q1 is 1); and the data collection tools have not been shown to be reliable (Q2 is 2) or reliability is not described (Q2 is 3).

Weak: The data collection tools have not been shown to be valid (Q1 is 2) or both reliability and validity are not described (Q1 is 3 and Q2 is 3).

### **WITHDRAWALS AND DROP-OUTS – a rating of:**

Strong: will be assigned when the follow-up rate is 80% or greater (Q2 is 1).

Moderate: will be assigned when the follow-up rate is 60-79% (Q2 is 2) OR Q2 is 5 (N/A).

Weak: will be assigned when a follow-up rate is less than 60% (Q2 is 3) or if the withdrawals and drop-outs were not described (Q2 is 4).

### E. Quality assessment results for included studies

|                                      |  |                      |
|--------------------------------------|--|----------------------|
| <b>Study ID:</b>                     | <b>Schuring, 2010</b>  |                      |
| <b>Overall Rating for the paper:</b> | STRONG (++)  |                      |
| <b>A)</b>                            | <b>Selection bias</b>  |                      |
| Q1                                   | Are the individuals selected to participate in the study likely to be representative of the target population                              | 1. Very Likely       |
| Q2                                   | What percentage of selected individuals agreed to participate?   | 1. 80-100% agreement |
| <b>Rate this section:</b>            |  | STRONG (++)          |
| <b>B)</b>                            | <b>Study design</b>  |                      |
|                                      | Indicate the study design  | Longitudinal study   |
|                                      | Was the study described as randomised? (If no, go to C))   | 2. No                |
|                                      | If yes, was the method of randomisation described? (See dictionary)  |                      |
|                                      | If yes, was the method appropriate? (See dictionary)   |                      |
| <b>Rate this section:</b>            |  | STRONG (++)          |
| <b>C)</b>                            | <b>Confounders</b>   |                      |
| Q1                                   | Were there important differences between groups prior to intervention?   | 2. No                |
| Q2                                   | If yes, indicate the percentage of relevant confounders that were controlled (either in design (eg stratification, matching) or analysis). |                      |
| <b>Rate this section:</b>            |  | STRONG (++)          |
| <b>D)</b>                            | <b>Blinding</b>  |                      |
| Q1                                   | Was (were) the outcome assessor(s) aware of the intervention or exposure status of participants?   | 2. No                |
| Q2                                   | Were the study participants aware of the research question?  | 1. Yes               |
| <b>Rate this section:</b>            |  | STRONG (++)          |

|                           |   |                     |
|---------------------------|---|---------------------|
| <b>Study ID:</b>          | <b>Schuring, 2010</b>   |                     |
| <b>E)</b>                 | <b>Data collection methods</b>  |                     |
| Q1                        | Were data collection tools shown to be valid?   | 1. Yes              |
| Q2                        | Were data collection tools shown to be reliable?  | 1. Yes              |
| <b>Rate this section:</b> |   | <b>STRONG (++)</b>  |
| <b>F)</b>                 | <b>Withdrawals and drop-outs</b>  |                     |
| Q1                        | Were withdrawals and drop-outs reported in terms of numbers and/or reasons per group?   | 1. Yes              |
| Q2                        | Indicate the percentage of participants completing the study. (If the percentage differs by group, record the lowest).            | 3. Less than 60%    |
| <b>Rate this section:</b> |   | <b>MODERATE (+)</b> |
| <b>G)</b>                 | <b>Intervention integrity</b>   |                     |
| Q1                        | What percentage of participants received the allocated intervention or exposure of interest?                                      | 1. 80-100%          |
| Q2                        | Was the consistency of the intervention measured?   | 1. Yes              |
| Q3                        | Is it likely that subjects received an unintended intervention (contamination or co-intervention) that may influence the results? | 2. No               |
| <b>H)</b>                 | <b>Analyses</b>   |                     |
| Q1                        | Indicate the unit of allocation   | Individual          |
| Q2                        | Indicate the unit of analysis   | Individual          |
| Q3                        | Are the statistical methods appropriate for the study design?   | 1. Yes              |
| Q4                        | Is the analysis performed by intervention allocation status (ie intention to treat) rather than the actual intervention required? | 1. Yes              |

|                                      |  |  |
|--------------------------------------|--|--|
| <b>Study ID:</b>                     | <b>Butterworth, 2011</b>   |  |
| <b>Overall Rating for the paper:</b> | STRONG (++)  |  |
| <b>A)</b>                            | <b>Selection bias</b>  |  |
| Q1                                   | Are the individuals selected to participate in the study likely to be representative of the target population                              | 1. Very Likely   |
| Q2                                   | What percentage of selected individuals agreed to participate?   | 2. 60-79% agreement  |
| <b>Rate this section:</b>            |  | MODERATE (+)   |
| <b>B)</b>                            | <b>Study design</b>  |  |
|                                      | Indicate the study design  | Longitudinal   |
|                                      | Was the study described as randomised? (If no, go to C))   | 2. No  |
|                                      | If yes, was the method of randomisation described? (See dictionary)  |  |
|                                      | If yes, was the method appropriate? (See dictionary)   |  |
| <b>Rate this section:</b>            |  | MODERATE (+)   |
| <b>C)</b>                            | <b>Confounders</b>   |  |
| Q1                                   | Were there important differences between groups prior to intervention?   | 2. No  |
| Q2                                   | If yes, indicate the percentage of relevant confounders that were controlled (either in design (eg stratification, matching) or analysis). |  |
| <b>Rate this section:</b>            |  |  |
| <b>D)</b>                            | <b>Blinding</b>  |  |
| Q1                                   | Was (were) the outcome assessor(s) aware of the intervention or exposure status of participants?   | Not applicable (data drawn from seven waves of national household surveys) |
| Q2                                   | Were the study participants aware of the research question?  |  |
| <b>Rate this section:</b>            |  |  |

|                           |   |  |
|---------------------------|---|--|
| <b>Study ID:</b>          | <b>Butterworth, 2011</b>  |  |
| <b>E)</b>                 | <b>Data collection methods</b>  |  |
| Q1                        | Were data collection tools shown to be valid?   | 1. Yes   |
| Q2                        | Were data collection tools shown to be reliable?  | 1. Yes   |
| <b>Rate this section:</b> |   | <b>STRONG (++)</b>   |
| <b>F)</b>                 | <b>Withdrawals and drop-outs</b>  |  |
| Q1                        | Were withdrawals and drop-outs reported in terms of numbers and/or reasons per group?   | Not applicable (data drawn from seven waves of national household surveys) |
| Q2                        | Indicate the percentage of participants completing the study. (If the percentage differs by group, record the lowest).            |  |
| <b>Rate this section:</b> |   |  |
| <b>G)</b>                 | <b>Intervention integrity</b>   |  |
| Q1                        | What percentage of participants received the allocated intervention or exposure of interest?                                      | Not applicable (data drawn from seven waves of national household surveys) |
| Q2                        | Was the consistency of the intervention measured?   |  |
| Q3                        | Is it likely that subjects received an unintended intervention (contamination or co-intervention) that may influence the results? |  |
| <b>H)</b>                 | <b>Analyses</b>   |  |
| Q1                        | Indicate the unit of allocation   | Community  |
| Q2                        | Indicate the unit of analysis   | Individual   |
| Q3                        | Are the statistical methods appropriate for the study design?   | 2. No  |
| Q4                        | Is the analysis performed by intervention allocation status (ie intention to treat) rather than the actual intervention required? | Not applicable (data drawn from seven waves of national household surveys) |

|                                      |   |                                  |
|--------------------------------------|---|----------------------------------|
| <b>Study ID:</b>                     | <b>Galić &amp; Šverko, 2008</b>   |                                  |
| <b>Overall Rating for the paper:</b> | MODERATE (+)  |                                  |
| <b>A)</b>                            | <b>Selection bias</b>   |                                  |
| Q1                                   | Are the individuals selected to participate in the study likely to be representative of the target population                             | 1. Very Likely                   |
| Q2                                   | What percentage of selected individuals agreed to participate?  | 2. 60-79% agreement              |
| <b>Rate this section:</b>            |   | MODERATE (+)                     |
| <b>B)</b>                            | <b>Study design</b>   |                                  |
|                                      | Indicate the study design   | Longitudinal                     |
|                                      | Was the study described as randomised? (If no, go to C))  | 1. Yes                           |
|                                      | If yes, was the method of randomisation described? (See dictionary)   | 1. Yes                           |
|                                      | If yes, was the method appropriate? (See dictionary)  | 1. Yes                           |
| <b>Rate this section:</b>            |   | MODERATE (+)                     |
| <b>C)</b>                            | <b>Confounders</b>  |                                  |
|                                      |   | Age, education                   |
| Q1                                   | Were there important differences between groups prior to intervention?  | 1. Yes                           |
| Q2                                   | If yes, indicate the percentage of relevant cofounders that were controlled (either in design (eg stratification, matching) or analysis). |                                  |
| <b>Rate this section:</b>            |   |                                  |
| <b>D)</b>                            | <b>Blinding</b>   |                                  |
| Q1                                   | Was (were) the outcome assessor(s) aware of the intervention or exposure status of participants?  | Not applicable (no intervention) |
| Q2                                   | Were the study participants aware of the research question?   | 1. Yes                           |
| <b>Rate this section:</b>            |   |                                  |
| <b>E)</b>                            | <b>Data collection methods</b>  |                                  |

|                           |   |                                  |
|---------------------------|---|----------------------------------|
| <b>Study ID:</b>          | <b>Galić &amp; Šverko, 2008</b>   |                                  |
| Q1                        | Were data collection tools shown to be valid?   | 1. Yes                           |
| Q2                        | Were data collection tools shown to be reliable?  | 3. Can't tell                    |
| <b>Rate this section:</b> |   | <b>MODERATE (+)</b>              |
| <b>F)</b>                 | <b>Withdrawals and drop-outs</b>  |                                  |
| Q1                        | Were withdrawals and drop-outs reported in terms of numbers and/or reasons per group?   | 1. Yes                           |
| Q2                        | Indicate the percentage of participants completing the study. (If the percentage differs by group, record the lowest).            | 3. Less than 60%                 |
| <b>Rate this section:</b> |   | <b>WEAK (-)</b>                  |
| <b>G)</b>                 | <b>Intervention integrity</b>   |                                  |
| Q1                        | What percentage of participants received the allocated intervention or exposure of interest?                                      | Not applicable (no intervention) |
| Q2                        | Was the consistency of the intervention measured?   |                                  |
| Q3                        | Is it likely that subjects received an unintended intervention (contamination or co-intervention) that may influence the results? |                                  |
| <b>H)</b>                 | <b>Analyses</b>   |                                  |
| Q1                        | Indicate the unit of allocation   | Community                        |
| Q2                        | Indicate the unit of analysis   | Individual                       |
| Q3                        | Are the statistical methods appropriate for the study design?   | 1. Yes                           |
| Q4                        | Is the analysis performed by intervention allocation status (ie intention to treat) rather than the actual intervention required? | Not applicable (no intervention) |

## F. High-level overview and quality assessment results for shortlisted studies

**Table 10: Overview of included studies**

| Study ID                     | Study type         | Study population   | RTW metric   | Outcome measures  | Tool used to measure outcomes      | Study quality |
|------------------------------|--------------------|--|--|---|------------------------------------|---------------|
| Audhoe et. al, 2010          | Systematic Review  | Unemployed general population                                    | Any employment   | Psychological distress                                  | Review of studies                  | Moderate (+)  |
| Ayala et. al, 2013           | Quasi-experimental | Long-term unemployed general population                          | Participating in work-related activities                         | Physical and mental health                              | Propensity score matching analysis | Moderate (+)  |
| Beyer et. al, 2010           | Quasi-experimental | Population with intellectual disabilities                        | Any supported employment, employment enterprises or day services | Quality of life   | ComQoI-A/ComQoI-I                  | Weak (-)      |
| Bockerman & Ilmakunnas, 2009 | Quasi-experimental | General population   | Any employment   | Self-assessed general health                            | Non-validated questionnaire        | Moderate (+)  |
| Bush et. al, 2009            | Observational      | Unemployed population with severe mental health problems         | Any employment   | Financial – utilisation costs of outpatient service use | Latent-class growth analysis       | Strong (++)   |
| Butterworth et. al, 2010     | Longitudinal       | Unemployed general population (assesses quality of job)          | Employed in a good/bad quality job                               | Mental health   | MHI-5                              | Strong (++)   |
| Carlier et. al, 2013         | Prospective study  | Unemployed general population receiving social security benefits | Any employment   | Quality of life   | Non-validated questionnaire        | Strong (++)   |
| Crowther et. al, 2010        | Systematic review  | Unemployed population with severe mental health disorders        | Any employment   | Quality of life / mental health                         | Global outcomes scale' Rosenberg   | Strong (++)   |

Estimation of benefits from moving an individual from unemployment into sustainable employment

| Study ID               | Study type         | Study population  | RTW metric                                 | Outcome measures                   | Tool used to measure outcomes                                 | Study quality |
|------------------------|--------------------|---|--|------------------------------------|---|---------------|
|                        |                    |   |  |                                    | scale, Lehman's scale, Brief Psychiatric Rating Scale (BPRS)  |               |
| Curnock et. al, 2016   | Longitudinal       | Unemployed general population receiving disability benefits                   | Any employment                             | Mental and physical health         | SF-12   | Strong (++)   |
| Ferreira et. al, 2015  | Longitudinal       | Unemployed general population   | Any employment                             | Mental health                      | Satisfaction with life scale                                  | Moderate (+)  |
| Galić & Šverko, 2008   | Longitudinal       | Unemployed general population receiving social security benefits              | Any employment                             | Mental health                      | SF-36   | Moderate (+)  |
| Gebel & Voßemer, 2014  | Quasi-experimental | General continuously employed, unemployed and long-term unemployed population | Any paid temporary or permanent employment | Psychological and mental health    | Life satisfaction scale                                       | Moderate (+)  |
| Hao, 2008              | Longitudinal       | General population, aged 55-66 (no activity, paid work or voluntary work)     | Any paid or formal volunteering            | Mental health                      | Centre for Epidemiological Studies – Depression Scale (CES-D) | Moderate (+)  |
| Hoare & Machin, 2010   | Longitudinal       | Unemployed general population   | Any employment                             | Financial hardship / mental health | Latent and Manifest Benefits Scale (LAMB); GHQ-12             | Weak (-)      |
| Huber et. al, 2009     | Prospective study  | Unemployed general population   | Any employment, min. one month             | General health                     | Non-validated questionnaire                                   | Moderate (+)  |
| Jacobs et. al, 2016    | Longitudinal       | Low-income urban women with at least one child                                | Any employment                             | Mental health                      | Brief Symptom Inventory (BSI-18)                              | Moderate (+)  |
| Schneider et. al, 2009 | Pre-post study     | Unemployed population with severe mental health problems                      | Any employment                             | Service use and frequency          | Analysis of service use                                       | Strong (++)   |

Estimation of benefits from moving an individual from unemployment into sustainable employment

| Study ID                   | Study type        | Study population   | RTW metric                     | Outcome measures                              | Tool used to measure outcomes   | Study quality |
|----------------------------|-------------------|--|--------------------------------|---|---|---------------|
| Schuring et. al, 2010      | Longitudinal      | Unemployed general population receiving social security benefits | Any employment                 | Quality of life                               | SF-36   | Strong (++)   |
| Steele et. al, 2013        | Observational     | Unemployed male general population                               | Any employment                 | Mental health                                 | GHQ   | Strong (++)   |
| Thomas et. al, 2007        | Longitudinal      | Unemployed general population                                    | Any employment                 | Mental health                                 | GHQ-12  | Strong (++)   |
| van de Noordt et. al, 2014 | Systematic review | Unemployed general population                                    | Any employment                 | Mental health                                 | Various including: Hopkins Symptom checklist, Beck Depression inventory, CES-D, Rosenberg's scale, Edinburgh Postnatal Depression Scale (EPDS), GHQ | Strong (++)   |
| van Rijn et. al, 2015      | Systematic review | Unemployed population with severe mental health problems         | Any employment, min. one month | Functioning / mental health / quality of life | Various including: Global Assessment of Functioning (GAF); SF-36; BPRS.   | Moderate (+)  |

## G. Overview of health scales

**Table 11: Health scales: inclusion/exclusion summaries**

| Outcome Scales   | Scale Included / Excluded | Reason for Inclusion / Exclusion                                |
|--|---------------------------|---|
| 1-5 general health scale (non-validated)               | Exclude                   | Non-validated scale used; not able to be robustly generalised   |
| Beck Depression Inventory                              | Exclude                   | Unable to convert to QALYs                                      |
| Brief Psychiatric Rating Scale (BPRS)                  | Exclude                   | Unable to convert to QALYs                                      |
| Brief Syndrome Inventory                               | Exclude                   | Too specific population / unable to convert to QALYs            |
| Capacity to work hours per day (non-validated)         | Exclude                   | Non-validated scale used; not able to be robustly generalised   |
| CES Depression Scale                                   | Exclude                   | Too specific population /unable to convert to QALYs             |
| Comprehensive Quality of Life Scale                    | Exclude                   | Weak study quality  |
| Edinburgh Postnatal Depression Scale                   | Exclude                   | Too specific population   |
| Financial hardship 1-6 self-reported                   | Exclude                   | Weak study quality  |
| GHQ-12   | Exclude                   | Weak study quality / unable to convert to QALYs                 |
| Global Assessment of Functioning                       | Exclude                   | Unable to convert to QALYs                                      |
| Hopkins Symptom Checklist (MH)                         | Exclude                   | Unable to convert to QALYs                                      |
| Hospital, Anxiety and Depression Scale (HADS)          | Exclude                   | Unable to convert to QALYs                                      |
| Lancashire Quality of Life Profile                     | Exclude                   | Unable to convert to QALYs                                      |
| Latent and Manifest Benefits scale                     | Exclude                   | Weak study quality / unable to convert to QALYs                 |
| Lehman's Scale (QoL)                                   | Exclude                   | Unable to convert to QALYs                                      |
| Life satisfaction scale                                | Exclude                   | Unable to convert to QALYs                                      |
| Manchester Short Assessment of Quality of Life (MANSA) | Exclude                   | Unable to convert to QALYs                                      |
| MHI-5  | Include                   | Validated questionnaire and able to convert findings into QALYs |

Estimation of benefits from moving an individual from unemployment into sustainable employment

| Outcome Scales                              | Scale Included / Excluded | Reason for Inclusion / Exclusion                                      |
|---|---------------------------|---|
| Positive and Negative Affect Scales (PANAS) | Exclude                   | Unable to convert to QALYs  |
| Rosenberg's Self-Esteem Scale               | Exclude                   | Measurement (self-esteem) too subjective / unable to convert to QALYs |
| Satisfaction with Life Scale                | Exclude                   | Measurement (self-esteem) too subjective / unable to convert to QALYs |
| Self-reported survey data                   | Exclude                   | Non-validated survey  |
| SF-12                                       | Exclude                   | Unable to convert to QALYs  |
| SF-36                                       | Include                   | Validated questionnaire and able to convert findings into QALYs       |
| VAS (1-10)                                  | Exclude                   | Non-validated scale used; not able to be robustly generalised         |
| WHO Quality of Life Measure                 | Exclude                   | Unable to convert to QALYs  |
| Wisconsin Quality of Life Index             | Exclude                   | Unable to convert to QALYs  |

## H. Full data extraction for shortlisted and included studies

| <b>Study ID</b>                | <b>Ayala, 2013</b>   |
|--------------------------------|--|
| Country                        | Spain  |
| Study Population (N)           | N = 2,300 households<br><br>Authors used the socioeconomic information from administrative records of households in Madrid's Welfare Program (IMI) to estimate the probability of taking part in a given activity. Authors used this data to identify the 'treatments'. The 2001 survey of 2,300 sample households uses, as initial universe, those in the program at a given moment in the previous decade. |
| Study Population               | Unemployed general population receiving specific benefits via Madrid Welfare Program (IMI)   |
| Study type                     | Quasi-experimental   |
| Definition of employment       | Participation in work-related activities including both general labour-orientated activities (access to specific employment offers, general job search assistance, and training) and labour-intensive sub-programs such as subsidised employment and social enterprises.   |
| Length of follow-up            | 10 years   |
| Duration of employment         | Continuously employed  |
| Duration of unemployment       | Not applicable – propensity score matching process   |
| Outcome type: health           | Yes  |
| Outcome type: financial        | No   |
| Outcome type: Other            | No   |
| Outcome metric                 | Changes in physical and mental health  |
| Statistical analysis           | Propensity score matching analysis   |
| Employment status at baseline  | Unemployed   |
| Employment status at follow-up | Not applicable – propensity score matching process   |

| Study ID      | Ayala, 2013   |
|---------------|---|
| Results       | <p>Statistics confidently show that engaging welfare recipients in work-related activities yields a reduction of 5.5% in physical health problems.</p> <p>However, with regards to mental health, results show that participation in work-related activities produces only modest positive effects; a reduction of 1.4% in mental health problems. This is not statistically significant.</p> <p>Participation in work-oriented measures would result in a 5.9% reduction in overall health problems – including health outcomes and behaviours</p> <p>As regards health behaviours, the effects on alcoholism and drug addiction were positive but not significant. Gambling addiction had a 1.2% reduction and was significant.</p> |
| Study quality | Moderate (+)  |

| Study ID                      | Beyer 2010   |
|-------------------------------|--|
| Country                       | United Kingdom   |
| Study Population (N)          | A purposive sample of people with intellectual disabilities was drawn from two supported employment agencies in South Wales (n = 17), two employment enterprises developed as part of day service modernisation (n = 10), and from two day services, five from a traditional day centre and five from an outreach centre, together representing a range of activities widely available today (n = 10). A further group (n = 17) were non-disabled co-workers of the supported employment participants. |
| Study Population              | Population with intellectual disabilities  |
| Study type                    | Quasi-experimental   |
| Definition of employment      | Progress towards employment achieved in partnership with specialist supported employment agencies  |
| Length of follow-up           | N/A  |
| Duration of employment        | Not reported   |
| Duration of unemployment      | Not reported   |
| Outcome type: health          | Yes  |
| Outcome type: financial       | No   |
| Outcome type: Other           | No   |
| Outcome metric                | Quality of Life was measured using the adult (ComQol-A / ComQol-I)   |
| Statistical analysis          | Non-parametric statistics were used in analysis.   |
| Employment status at baseline | The study compared a group of participants who gained supported employment, against a group of similar participants attending adult day centres.   |

| Study ID                       | Beyer 2010  |
|--------------------------------|---|
| Employment status at follow-up | Employed  |
| Results                        | Objective measure mean (SD):<br>Supported employees – 71.5 (7.7)<br>Day service attendees – 58.1 (11.0)<br><br>Subjective measure mean (SD):<br>Supported employees – 100.9 (26.1)<br>Day service attendees – 68.9 (50.3) |
| Study quality                  | Weak (-)  |

| Study ID                      | Bockerman, 2009  |
|-------------------------------|--|
| Country                       | Finland  |
| Study Population (N)          | Persons unemployed at least once over the period 1996-2001. The reference (employed) group consists of those continually in work.  |
| Study Population              | General continuously employed, unemployed and long-term unemployed population  |
| Study type                    | Quasi-experimental   |
| Definition of employment      | Employed = continuously employed<br>Unemployed = unemployed for < six months<br>Long-term unemployed = unemployed continuously >= six months.  |
| Length of follow-up           | Six years  |
| Duration of employment        | For employed group: continuously employed.   |
| Duration of unemployment      | Information about the unemployment duration for the persons interviewed. However, the data record monthly activity statuses (unemployed being one possible alternative) for each person for the whole year before the interview.                                       |
| Outcome type: health          | Yes  |
| Outcome type: financial       | No   |
| Outcome type: Other           | No   |
| Outcome metric                | Self-assessed health --> measured on an ordinal five-point Likert scale with alternatives five ('very good'), four ('good'), three ('fair'), two ('bad') or one ('very bad'). Hence, a higher value on this scale means that a person feels currently healthier.       |
| Statistical analysis          | Panel data analysis. Apply difference-in-differences models and matching methods. In particular, the use of matching methods allows the selection for unemployment and the possibility of reverse causality from poor health to unemployment to be taken into account. |
| Employment status at baseline | Varies for people who experience unemployment at some point during the study.<br><br>Those in the 'control' group are continuously employed.   |

| Study ID                       | Bockerman, 2009   |
|--------------------------------|---|
| Employment status at follow-up | Varies for people who experience unemployment at some point during the study.<br><br>Those in the 'control' group are continuously employed.  |
| Results                        | Reference group for all models is 'continuously employed'. Results for employed after unemployment in models assessing the effect of labour market status on self-assessed health:<br>OLS: -0.30 (0.031), OLS: -0.018 (0.029)<br>Fixed effects: -0.009 (0.024), -0.008 (0.024)<br>Results are not significant.<br><br>Those who are unemployed but become employed again at some stage have a somewhat higher health level than the reference group. On the other hand, when those working become unemployed, their self-assessed health status does not deteriorate and when those who are unemployed find a job, their health status does not improve. Taken together, the results show that unemployment as such does not seem to worsen the level of self-assessed health. It is more the case that the persons who experience poor health are being selected for the pool of unemployed persons and those who manage to escape unemployment tend to have better health in the first place. |
| Study quality                  | Moderate (+)  |

| Study ID             | Bush, 2009   |
|----------------------|--|
| Country              | USA  |
| Study Population (N) | N = 187 (originally 223 – included only those who had at least five annual data points over 10 years). NB: one mental health centre used supported employment services, the others provided some sort of other support.<br><br>From seven New Hampshire mental health catchment areas, participants who met the following criteria: long-term psychotic illness (schizophrenia, schizoaffective disorder, or bipolar disorder), active substance use disorder within the past six months, and absence of mental retardation. |
| Study Population     | Steady work (avg. 5050±3993 hours over 10 years) and minimum work (inc. no work - (avg. 411±654 hours over 10 years) group of people with severe mental health disorders   |
| Study type           | Observational  |

| Study ID                       | Bush, 2009   |
|--------------------------------|--|
| Definition of employment       | <p>Competitive jobs included any paid position in the regular job market. For each job, participants indicated the number of weeks since the previous interview that they worked and the average number of hours that they worked weekly.</p> <p>A steady work group (N=51) included individuals whose work hours increased rapidly and then stabilised to average 5,060 hours per person over 10 years and a minimum work group (N=136 - combination of no-work group and the late-work group)</p>  |
| Length of follow-up            | 10 years   |
| Duration of employment         | Varied   |
| Duration of unemployment       | Varied   |
| Outcome type: health           | No   |
| Outcome type: financial        | Yes  |
| Outcome type: Other            | Yes  |
| Outcome metric                 | Annual costs of outpatient services and institutional stays, examining the differences in 10-year utilisation costs. Research psychiatrists established co-occurring diagnoses of severe mental illnesses and substance use disorders using the Structured Clinical Interview for DSM-III-R (28).  |
| Statistical analysis           | Latent-class growth analysis. A three-step analysis was conducted. Firstly, latent trajectory groups using hours in competitive jobs over 10 years were identified. Secondly, possible trajectory group differences with respect to baseline characteristics were tested. Thirdly, trajectory group differences were tested with respect to longitudinal time-varying covariates of interest – outpatient service utilisation, stays in institutions (hospitals, jails, or prisons), and costs – controlling for baseline differences identified in the second step. |
| Employment status at baseline  | <p>Minimum work group = 45/136 (34%) worked in the past year (any job)</p> <p>Steady work group = 39/51 (76%) worked in the past year (any job)</p>  |
| Employment status at follow-up | <p>At 10 year follow-up (has data for interim years also):</p> <p>Total competitive work hours (mean±SD):</p> <p>Minimum work group: 140±311</p> <p>Steady work group: 667±673</p>   |

| Study ID      | Bush, 2009  |
|---------------|---|
| Results       | Use of outpatient services for the steady work group declined at a significantly greater rate than it did for the minimum-work group, while institutional (hospital, jail, or prison) stays declined for both groups without a significant difference. The average annual cost of outpatient services and institutional days for the minimum-work group over 10 years of follow-up was \$31,108± 29,451 per year, more than two times that of the steady work group total of \$14,473±13,258. Over the 10 years, cost in the minimum-work group exceeded that in the steady work group by \$166,350 per person. |
| Study quality | Strong (++)   |

| Study ID                 | Butterworth, 2010   |
|--------------------------|---|
| Country                  | Australia   |
| Study Population (N)     | Overall study sample N = 7,155 (N men = 3,305, N women = 3,850), aged between 20-55<br><br>Study sample specifically for employment transitions: N = 693 (either moved from unemployment to another employment circumstance or remained unemployed)   |
| Study Population         | Unemployed general population (assesses quality of job)   |
| Study type               | Longitudinal  |
| Definition of employment | Employment status was coded as employed (part-time or full-time), unemployed and looking for work, or not participating in the labour force (NILF).<br><br>Full details of the construction and validation of the different measures of job quality are presented elsewhere. (Leach, 2010: Deriving an Evidence-Based Measure of Job Quality from the HILDA Survey; Butterworth, 2009: Financial hardship, socio-economic position and depression: Results from the PATH Through Life Survey) |
| Length of follow-up      | 12 months   |
| Duration of employment   | Not known   |
| Duration of unemployment | Varied/unknown  |
| Outcome type: health     | Yes   |
| Outcome type: financial  | No  |
| Outcome type: Other      | No  |

| Study ID                       | Butterworth, 2010   |
|--------------------------------|---|
| Outcome metric                 | Mental health --> measured by the five-item Mental Health Inventory (MHI), a subscale from the SF-36 general health survey. The MHI assesses symptoms of depression and anxiety (nervousness, depressed affect) and positive aspects of mental health (feeling calm, happy) in the past four weeks. The current analyses use the continuous mental health scale score, with higher scores representing better mental health – a score of <50 indicates mental health problems.  |
| Statistical analysis           | A lagged longitudinal regression model was conducted to assess moving from unemployment to some type of employment, or remaining unemployed, with mental health change scores as the dependent measure.   |
| Employment status at baseline  | N = 693 unemployed<br>**For study examining employment transitions  |
| Employment status at follow-up | Numbers not provided  |
| Results                        | <p>The health benefits of becoming employed were dependent on the quality of the job. Moving from unemployment into a high quality job led to improved mental health (mean change score of +3.3). However, the transition from unemployment to a poor quality job was more detrimental to mental health than remaining unemployed ( 5.6 vs 1.0).</p> <p>The results confirm that those who moved into optimal jobs showed significant improvement in mental health compared to those who remained unemployed. Those respondents who moved into poor quality jobs showed a significant worsening in their mental health compared to those who remained unemployed.</p> |
| Study quality                  | Strong (++)   |

| Study ID                 | Carlier, 2013  |
|--------------------------|--|
| Country                  | Netherlands  |
| Study Population (N)     | <p>N= 4,308</p> <p>Unemployed persons between 18-65 years receiving social benefits in the Netherlands.</p> <p>The Dutch Employment Centre generated a dataset of 70,121 persons, who were on social benefits for at least six months in 2006. From this dataset an age-stratified random sample was drawn of 20,847 persons on unemployment benefit (UB) or social security benefit (SSB). Four sequential questionnaires resulted in a 35%, 59%, 40%, and 49% response. Persons with at least two complete questionnaires were included in the study. This resulted in a study sample of 4,308 persons of which 2,604 persons participated twice, 871 persons three times, and 833 persons four times.</p> |
| Study Population         | Unemployed general population receiving social security benefits   |
| Study type               | Prospective study  |
| Definition of employment | Re-employment was based on self-reports in the questionnaire on having entered paid employment in the past six months.   |
| Length of follow-up      | Up to 18 months  |
| Duration of employment   | Not known  |
| Duration of unemployment | Varied – Persons who remained unemployed received consecutive questionnaires, whereas those entering paid employment had a maximum follow-up of six months after the transition into employment.   |
| Outcome type: health     | Yes  |
| Outcome type: financial  | No   |
| Outcome type: Other      | No   |
| Outcome metric           | <p>Quality of life --&gt; measured with question 'How would you rate your life in general in the last six months?' on a 10-point VAS scale. Those reporting less than 'six' were defined as having a poor quality of life.</p> <p>Self-rated health was measured by a slightly adjusted question, derived from SF-36, "In general, how would you define your health in the last six months?" A five-point scale was used, ranging from 'very good', 'good', 'not good/not bad', 'poor', to 'very poor' [26]. Those reporting poor or very poor were defined as having a poor health and others were classified as good health.</p>   |
| Statistical analysis     | <p>Generalised estimation equation modelling was performed to study the influence of re-employment on change in self-rated health and quality of life over time. Independent variables were: employment status, duration of the employment, time, sex, age, education, ethnic background, parenthood, marital status and type of benefit.</p> <p>Quality of life and health at baseline were included as independent variables in the models, and, hence an OR above one reflects that among</p>   |

| Study ID                       | Carlier, 2013  |
|--------------------------------|--|
|                                | those with a transition from unemployment into paid employment health and quality of life improved compared with those without any transition.   |
| Employment status at baseline  | All unemployed   |
| Employment status at follow-up | Varied - Persons who remained unemployed received consecutive questionnaires, whereas those entering paid employment had a maximum follow-up of six months after the transition into employment.   |
| Results                        | <p>Persons who became re-employed were 2.88 times more likely to change from poor to good health compared with those who stayed unemployed (95% CI 2.37–3.50). Up to a maximum of six months after re-employment, every month in paid employment after re-employment, the likelihood of improvement of health increased with 1.05 (95% CI 0.93–1.18).</p> <p>A similar effect of re-employment on quality of life was observed (Table 3). Re-employed persons were 1.76 times more likely to change from poor to good quality of life (95% CI 1.54–2.02) compared with persons who continued to be unemployed. The duration of re-employment was also positively associated with quality of life, increasing the likelihood of transition from poor to good quality of life with 1.12 (95% CI 1.02–1.23) with each month.</p> <p>Among re-employed persons, 60% improved, 40% did not change, and 4% worsened in self-rated health after the employment transition. Among persons who continued to be unemployed, 39% improved, 61% did not change and 9% worsened in self-rated health. For quality of life similar patterns were observed. Among re-employed persons 37% improved, 63% did not change and 8% worsened in quality of life, whereas for persons who continued to be unemployed 23% improved, 77% did not change and 8% improved.</p> <p>The beneficial effect of re-employment on health was more profound among men (OR 3.65 95% CI 2.60–5.12) than among women (OR 2.10 95% CI 1.62–2.71) (Table 4). The positive effect of re-employment on self-rated health and quality of life decreased with increasing age. In addition, among native Dutch persons (OR 4.01 95% CI 3.00–5.14) the increase in health was larger compared to non-native Dutch persons (OR 2.22 95% CI 1.52–3.22). Educational level of type of benefit did not influence the effect of re-employment on health or quality of life.</p> |
| Study quality                  | Strong (++)  |

| Study ID                       | Crowther 2010   |
|--------------------------------|---|
| Country                        | International study/Systematic Review   |
| Study Population (N)           | Specific inclusion criteria were that a majority of clients in the trial were: (a) aged 18-65; and (b) suffering from severe mental disorder defined as: schizophrenia and schizophrenia-like disorders; bipolar disorder; or depression with psychotic features.   |
| Study Population               | Unemployed population with severe mental health disorders   |
| Study type                     | Systematic review   |
| Definition of employment       | Full- or part-time position held by the client in an ordinary work setting, for which they were receiving payment at the market rate  |
| Length of follow-up            | Ranged from four to 48 months   |
| Duration of employment         | Not reported  |
| Duration of unemployment       | Not reported  |
| Outcome type: health           | Yes   |
| Outcome type: financial        | Yes   |
| Outcome type: Other            | No  |
| Outcome metric                 | Various:<br>(a) Global outcome scale<br>(b) Self-esteem (Rosenberg Scale)<br>(c) Quality of Life (Lehman's scale, Lehman 1983)<br>(d) Psychiatric symptoms (BPRS)   |
| Statistical analysis           | Random effects meta-analysis conducted but not for relevant outcomes.   |
| Employment status at baseline  | Non employed  |
| Employment status at follow-up | Employed/not in employment  |
| Results                        | Eighteen randomised controlled trials were included in the review. One trial, involving 256 patients, found that for supported employment versus standard community care, there was no difference in the number of hospital admissions between Supported Employment and control. For supported employment (all approaches) versus pre-vocational training, there were five included studies. One reported no significant differences at six, 12 and 18 months in: (a) Global outcome (GAS, eg at 18 months Supported Employment 45.8 (SE 1.43), control 46.0 (SE 1.78)); (b) Self-esteem (Rosenberg Scale, eg at 18 months Supported Employment 18.5 (SE 0.7), control 18.1 (SE 0.68), (c) Quality of Life (Lehman's scale, Lehman 1983, eg at 18 months Supported Employment 5 (SE 0.17), control 4.8 (SE 0.18)) or (d) Psychiatric symptoms (BPRS, eg at 18 months Supported Employment 39.2 (SE 1.19), control 41.1(SE 1.54)). |
| Study quality                  | Strong (++)   |

| Study ID                       | Curnock, 2016   |
|--------------------------------|---|
| Country                        | United Kingdom  |
| Study Population (N)           | For those transitioning from disability benefits to employment.<br><br>For mental health: N control = 1545, N treatment = 124 (total N = 1669)<br>For physical health: N control = 1547, N treatment = 124 (total N = 1671)<br><br>Where 'treatment' represents the number of people who moved from receiving disability benefits (IB or ESA) to employment |
| Study Population               | Unemployed general population receiving disability benefits   |
| Study type                     | Longitudinal  |
| Definition of employment       | None provided   |
| Length of follow-up            | 12 months   |
| Duration of employment         | Not provided  |
| Duration of unemployment       | Not provided  |
| Outcome type: health           | Yes   |
| Outcome type: financial        | No  |
| Outcome type: Other            | No  |
| Outcome metric                 | Outcome measures were based on the 'Short Form-12' (SF-12) self-completed questionnaire, with individual responses converted to physical and mental health summary scores on a scale from 0 (low functioning) to 100 (high functioning), with a mean population score of 50   |
| Statistical analysis           | A difference-in-difference approach was used to compare change between treatment and control groups in mental and physical health using the SF-12.  |
| Employment status at baseline  | N = 1669/1671 (mental/physical) All receiving disability benefits (IB or ESA)<br><br>*NB: for analysis concerning movement from receiving disability benefits (IB or ESA) to employment   |
| Employment status at follow-up | N = 124   |

| <b>Study ID</b> | <b>Curnock, 2016</b>  |
|-----------------|---|
| Results         | <p>Transitions from disability benefits (IB or ESA) to employment (n =124) were associated on average with an improvement in the SF-12 mental health score of 5.94 points (95% CI = 3.52 - 8.36), and an improvement in the physical health score of 2.83 points (95% CI = 0.85 - 4.81) compared with those remaining on disability benefits (n = 1545). Transitions to unemployed status (n = 153) were associated with a significant improvement in mental health but not physical health.</p> <p>The propensity score matching strategy sought to reduce the bias between groups for each covariate. Once the propensity score matching strategy was finalised the difference-in-difference estimates were re-calculated (DiD-PSM); the statistically significant transitions were unchanged from the main DiD analyses although each of the estimates reduced slightly.</p> <p>Note for table: Analysis 1 = transitions from disability benefits to employment.</p> |
| Study quality   | <b>Strong (++)</b>  |

| <b>Study ID</b>          | <b>Ferreira, 2015</b>  |
|--------------------------|--|
| Country                  | Portugal   |
| Study Population (N)     | 566 Portuguese men and women who were unemployed at the beginning of the study.  |
| Study Population         | Unemployed general population  |
| Study type               | Longitudinal   |
| Definition of employment | <p>The relation between configurations of unemployment history at the beginning of the study and re-employment three years later were examined.</p> <p>Two indicators of unemployment were used: (a) duration of unemployment at T1 (less than a month, one to six months, six months to one year, more than one year); [collapsed to short term=one year or less, and long term=more than one year]; (b) employed at T2 (or not).</p> |
| Length of follow-up      | Three years  |
| Duration of employment   | Two indicators of unemployment were used: (a) duration of unemployment at T1 (less than a month, one to six months, six months to one year, more than one year); [collapsed to short term=one year or less, and long term=more than one year]; (b) employed at T2 (or not).  |
| Duration of unemployment | Two indicators of unemployment were used: (a) duration of unemployment at T1 (less than a month, one to six months, six months to one year, more than one year); [collapsed to short term=one year or less, and long term=more than one year]; (b) employed at T2 (or not).  |
| Outcome type: health     | <b>Yes</b>   |
| Outcome type: financial  | <b>No</b>  |
| Outcome type: Other      | <b>No</b>  |

| <b>Study ID</b>                | <b>Ferreira, 2015</b>  |
|--------------------------------|--|
| Outcome metric                 | The cognitive aspect of subjective wellbeing was assessed using the Portuguese version of the Satisfaction with Life Scale; Positive and negative affect scales --> measure addresses both positive and negative affective components of wellbeing.  |
| Statistical analysis           | True change score models   |
| Employment status at baseline  | All unemployed   |
| Employment status at follow-up | Various  |
| Results                        | Individuals who reported a long period of unemployment at T1 but were re-employed at T2 showed meaningful gains in positive affect and life satisfaction compared to those who had a shorter history of unemployment and were re-employed. An examination of gender differences revealed that the women who were re-employed after a long history of unemployment showed the greatest relative gains in positive affect. |
| Study quality                  | Moderate (+)   |

| <b>Study ID</b>          | <b>Galić &amp; Šverko, 2008</b>  |
|--------------------------|--|
| Country                  | Croatia  |
| Study Population (N)     | N = 394 unemployed persons registered with the Croatian Employment Bureau who completed all three waves of questionnaires.                 |
| Study Population         | Unemployed general population  |
| Study type               | Longitudinal   |
| Definition of employment | None provided  |
| Length of follow-up      | Two to two-and-a-half years  |
| Duration of employment   | Not provided   |
| Duration of unemployment | Of the participants 35% were unemployed for less than six months, 32% from seven months to three years, and 26% for more than three years. |
| Outcome type: health     | Yes  |
| Outcome type: financial  | No   |
| Outcome type: Other      | No   |

| Study ID                       | Galić & Šverko, 2008  |
|--------------------------------|---|
| Outcome metric                 | Health was assessed with the SF-36 Health Survey, psychometrically sound multi-dimensional questionnaire tapping both physical and mental health concepts. The scores on all dimensions were transformed into a scale whose theoretical minimum is nought, and maximum 100 scores, with higher scores indicating better health. General psychological health was calculated as a mean score from Social Functioning, Role – Emotional, and Mental Health scales. In a similar vein, general physical health was computed as mean result from Physical Function, Role – Physical and Bodily Pain scales. Considering that focus was only on general psychological and physical health, results on General Health and Vitality scales weren't used in further analyses. |
| Statistical analysis           | In order to examine in more detail the effects of prolonged unemployment and re-employment on physical and psychological health, the health scores of the three groups obtained at the successive time points with the repeated-measures analysis of covariance (ANCOVA) were analysed. To control for the potentially confounding influence of the socio-demographic variables, age, gender, education and unemployment duration were used as the covariates.  |
| Employment status at baseline  | All unemployed  |
| Employment status at follow-up | N = 206 for continuously unemployed (UUU)<br>N = 53 for unemployed at baseline, then unemployed, then employed (UUE)<br>N = 113 for unemployed at baseline, then continuously employed (UEE)  |
| Results                        | ***The significant main effect for groups ( $F(1,358) = 4.19; p < .05$ ) means that three groups differ in their average level of psychological health, and the significant interaction effect ( $F(2,358) = 3.00; p < .05$ ) denotes that changes in psychological health across time are different for the three groups.  |
| Study quality                  | Moderate (+)  |

| Study ID             | Gebel, 2014  |
|----------------------|--|
| Country              | Germany  |
| Study Population (N) | The sample includes persons between the ages of 16-54 who have at least experienced one respective employment transition (treatment groups) or are continuously employed or unemployed (control groups). |
| Study Population     | General continuously employed, unemployed and long-term unemployed population  |
| Study type           | Quasi-experimental   |

| Study ID                       | Gebel, 2014   |
|--------------------------------|---|
| Definition of employment       | Permanent employment = open-ended contracts,<br>Temporary employment = contracts of a predefined limited duration.  |
| Length of follow-up            | 15 years  |
| Duration of employment         | For the treatment “re-employment” (unemployed at t and employed at t + 1) only persons who report one single change from the status of unemployment to employment within the period [t; t + 1] and no other kind of monthly activities such as inactivity or education have been used. These persons are compared to the control group of persons who are continuously unemployed in each month of the period [t; t + 1].   |
| Duration of unemployment       | As above  |
| Outcome type: health           | Yes   |
| Outcome type: financial        | No  |
| Outcome type: Other            | No  |
| Outcome metric                 | Psychological and mental health --> measured by a life satisfaction scale ranging from 0 (completely dissatisfied) to 10 (completely satisfied).  |
| Statistical analysis           | Difference-in-difference propensity score matching (DIDPSM) approach  |
| Employment status at baseline  | Observed 2594 transitions into unemployment (treatment “job loss”) (based on 578 persons) and 79,784 person spells of continuous employment (respective control group) (based on 13,497 persons). In addition, 2470 transitions out of unemployment (treatment “re-employment”) (based on 750 persons) and 4753 person spells of continuous unemployment (respective control group) (based on 1025 persons) were observed.  |
| Employment status at follow-up | Various   |
| Results                        | The results show that only psychological, not physical, health is causally affected by the respective employment transitions. Specifically, the effects of unemployment and re-employment are of similar size, highlighting the importance of re-employment in compensating unemployment's negative impact on psychological health. In contrast, health selection and confounding seem to be important determinants of the cross sectional association between unemployment and physical health. Contrary to expectations it was found that both employment transitions have effects of a similar size for permanent and temporary workers. |
| Study quality                  | Moderate (+)  |

| <b>Study ID</b>                | <b>Hao, 2008</b>  |
|--------------------------------|---|
| Country                        | USA   |
| Study Population (N)           | N = 7,830 individuals aged 55-66  |
| Study Population               | N = 7,830 individuals aged 55-66  |
| Study type                     |   |
| Definition of employment       | Paid or formal volunteering   |
| Length of follow-up            | 10 years (1992 baseline, 1996, 1998, 2000 and 2002 follow-up years)   |
| Duration of employment         | Not reported  |
| Duration of unemployment       | Not reported  |
| Outcome type: health           | Yes   |
| Outcome type: financial        | No  |
| Outcome type: Other            | No  |
| Outcome metric                 | Mental health --> measured by the Centre for Epidemiologic Studies-Depression Scale (CES-D). The eight items in the modified CES-D scale in the HRS were selected based on factor analysis results reported in Radloff (1977). Respondents were asked whether they had experienced eight specific symptoms in the past week, including (a) was depressed, (b) everything was an effort, (c) sleep was restless, (d) was (not) happy, (e) felt lonely, (f) (did not) enjoy life, (g) felt sad, and (h) could not get going. Dichotomous response categories were consistent from Wave 1 to Wave 4 and signified the presence (1) or absence (0) of a symptom. Higher scores indicate more depressive symptoms. |
| Statistical analysis           | Panel data analysis. Growth curve modeling (Raudenbush & Bryk, 2002) was used to estimate individual health trajectories. As one type of hierarchical modeling, growth curve models represent a dual level analysis of change process.  |
| Employment status at baseline  | At baseline (N= 7,830), 56% of those at least 55 years old engaged in paid work, with 16% of respondents having a part-time job, and another 40% having full-time work. For formal volunteering, the figures were lower. The percentages were 21% and 12%, respectively, for low-level (<100 hr) and high-level (>100 hr) volunteering. As for concurrent participation in both paid and volunteer work, 21% of older adults fell into this category.   |
| Employment status at follow-up | Over the course of the study, respondents reported a steady decrease in involvement in paid work, dropping from 56% to 37%. Volunteer participation was relatively constant over time, ranging from 32-34%.   |
| Results                        |   |
| Study quality                  | Moderate (+)  |

| <b>Study ID</b>                | <b>Hoare, 2010</b>  |
|--------------------------------|---|
| Country                        | Australia   |
| Study Population (N)           | N = 115 (male = 59, female = 56)<br><br>115 unemployed participants (59 males) with a mean age of 38.81 years (SD = 14.49, range 17-64) registered with Government-funded employment agencies in South East Queensland, Australia.  |
| Study Population               | Unemployed general population   |
| Study type                     | Longitudinal  |
| Definition of employment       | Not provided  |
| Length of follow-up            | Six months  |
| Duration of employment         | Fifty-eight participants were employed and 57 were unemployed at Time 2 --> implies duration of employment of six months  |
| Duration of unemployment       | Fifty-eight participants were employed and 57 were unemployed at Time 2. Approximately 28% of participants had not worked at all in the past. Of the 57 people who were unemployed at the six-month follow-up, three had been employed at Time 1 (ie part-time or casually), seven had been doing volunteer/unpaid work at Time 1, 46 had not been working at Time 1, and one person had been studying at Time 1.   |
| Outcome type: health           | Yes   |
| Outcome type: financial        | Yes   |
| Outcome type: Other            | No  |
| Outcome metric                 | Financial hardship --> question on ease of living on net fortnightly income ranging from one (extremely easy) to six (extremely difficult).<br>Financial strain --> The Latent and Manifest Benefits Scale (LAMB), includes six items that measure access to manifest benefit of employment. LAMB is also used to measure access to the latent benefits (collective purpose, social contact, status, activity, and time structure).<br>Mental health --> GHQ-12 |
| Statistical analysis           | Study investigates perceived access to the latent benefits during unemployment and re-employment.<br><br>Analysis consisted of two-way ANOVAs with Time as a repeated measures factor   |
| Employment status at baseline  | 115: all unemployed   |
| Employment status at follow-up | N employed = 58<br>N unemployed = 57  |

| Study ID      | Hoare, 2010   |
|---------------|---|
| Results       | <p>Participants who gained employment were better off financially, reported greater access to social contact and time structure and had significant improvements in their mental health at Time 2.<br/>Participants that remained unemployed showed no change over time.</p> <p>For Financial hardship, there was a significant two-way interaction with <math>F(1,113) = 28.03, p &lt; .01, \eta^2 = .20</math>. Analysis of simple main effects showed that the scores for those who were re-employed improved significantly (95% CIs for T1 and T2 scores were 4.19 – 4.81 and 2.93 – 3.59 respectively) while the scores for those who remained unemployed remained stable (95% CIs for T1 and T2 scores were 4.11 – 4.74 and 4.07 – 4.73 respectively).</p> <p>Financial strain demonstrated a significant two-way interaction with <math>F(1,113) = 41.06, p &lt; .01, \eta^2 = .27</math>. Analysis of simple main effects also showed that the scores for those who were re-employed improved significantly (95% CIs for T1 and T2 scores were 32.92 – 37.08 and 22.03 – 27.11 respectively) while the scores for those who remained unemployed remained stable (95% CIs for T1 and T2 scores were 31.27 – 35.46 and 30.90 – 36.02 respectively).</p> <p>For the measures of access to the latent benefits of employment, there were significant two-way interactions for Social Contact with <math>F(1,113) = 9.51, p &lt; .01, \eta^2 = .08</math>, and Time Structure with <math>F(1,113) = 22.60, p &lt; .01, \eta^2 = .17</math>. Analysis of simple main effects for Social Contact showed that the scores for those who were re-employed improved significantly (95% CIs for T1 and T2 scores were 19.30 – 24.25 and 23.66 – 28.59 respectively) while the scores for those who remained unemployed remained stable (95% CIs for T1 and T2 scores were 18.19 – 23.18 and 17.88 – 22.86 respectively). Analysis of simple main effects for Time Structure showed that the scores for those who were re-employed improved significantly (95% CIs for T1 and T2 scores were 20.38 – 25.59 and 29.49 – 34.24 respectively) while the scores for those who remained unemployed remained stable (95% CIs for T1 and T2 scores were 24.93 – 30.19 and 25.46 – 30.26 respectively).</p> <p>Mental Health (GHQ), there was also a significant two-way interaction with <math>F(1,113) = 22.97, p &lt; .01, \eta^2 = .17</math>. Analysis of simple main effects showed that the scores for those who were re-employed improved significantly (95% CIs for T1 and T2 scores were 14.25 – 17.61 and 8.58 – 12.25 respectively) while the scores for those who remained unemployed remained stable (95% CIs for T1 and T2 scores were 12.27 – 15.66 and 12.92 – 16.62 respectively).</p> |
| Study quality | Weak (-)  |

| Study ID                      | Huber, 2009  |
|-------------------------------|--|
| Country                       | Germany  |
| Study Population (N)          | N= 21,000 initially, down to 13,914 after drop-outs<br><br>The evaluation sample was restricted to individuals who entered welfare within 12 months before interview, discarded individuals stating not to receive welfare benefits at interview 1 (246 obs.), being younger than 26 (1486), or having missing values in the outcomes (183) and pre-transition outcomes (182). The evaluation sample consists of 2,849 individuals, for whom three welfare states are considered: remaining on welfare (henceforth W), finding employment (E), and programme participation (P). The sample is not drawn randomly from the population of welfare recipients, but is stratified. Stratification is based on the following individual characteristics: age (15-24 / 25-49 / 50-64), children under age three are in the household, and being a single parent. |
| Study Population              | Unemployed general population receiving social security benefits   |
| Study type                    | Prospective study  |
| Definition of employment      | In employment or in a programme for at least one month between interviews one and two.<br><br>Work capacity per day (scale, 1: <3 hours, 4: 8+ hours)<br>Capability of working six or more hours per day (binary)  |
| Length of follow-up           | Not clear  |
| Duration of employment        | At least one month   |
| Duration of unemployment      |  |
| Outcome type: health          | Yes  |
| Outcome type: financial       | No   |
| Outcome type: Other           | No   |
| Outcome metric                | General health is covered by the assessment of overall health on a scale from one (very good) to five (poor) and the capacity to work up to a specific amount of hours per day (1: less than three hours, 2: three to less than six hours, 3: six to less than eight hours, 4: eight hours and more). Based on these variables indicator variables for 'very good, good, or satisfactory health' and of 'being capable of working six or more hours per day' were constructed.   |
| Statistical analysis          | Survey information linked to administrative records. Propensity score matching estimates. Whereas state W (remaining on welfare) is defined as receiving welfare over the whole transition period, E (finding employment) and P (programme participation) only require to be employed or in a programme, respectively, for at least one month.   |
| Employment status at baseline | Welfare recipients   |

| <b>Study ID</b>                | <b>Huber, 2009</b>  |
|--------------------------------|---|
| Employment status at follow-up | Varied  |
| Results                        | <p>The empirical results (inserted Table 2) suggest that the transition to work has a positive impact on health. While the increase in overall health is positive but not statistically significant, there is a large and significant positive effect on the daily work capacity: the probability of having a work capacity of six or more hours per day increases by 11% from 0.79 to 0.9. Moreover, the number of symptoms is reduced significantly by 0.19. The prevalence of mental symptoms decreases by 8% (from 0.30 to 0.22), which may be driven by the decline in nerval problems, anxieties, and sleeping problems. Furthermore, working individuals are less likely to feel lethargic and depressed by 8% compared to their matched counterparts. In contrast, physical symptoms do not seem to be affected much, at least for the short follow-up period considered.</p> <p>The positive effects are mainly driven by males and individuals with bad initial health conditions and are largest for males with poor health. In contrast, the effects of welfare-to-work programmes, including subsidised jobs, are ambiguous and statistically insignificant for most outcomes.</p> |
| Study quality                  | Moderate (+)  |

| <b>Study ID</b>          | <b>Jacobs, 2016</b>   |
|--------------------------|---|
| Country                  | USA   |
| Study Population (N)     | N = 2,126 low-income urban women with at least one child (aged 0-4 or 10-14) (11% lost to follow up)  |
| Study Population         | Low-income urban women with at least one child  |
| Study type               | Longitudinal  |
| Definition of employment | Respondents were asked if they worked for pay in the last week. Responses were dummy coded (0) no and (1) yes.  |
| Length of follow-up      | Two years   |
| Duration of employment   | Not provided  |
| Duration of unemployment | Not provided  |
| Outcome type: health     | Yes   |
| Outcome type: financial  | No  |
| Outcome type: Other      | No  |
| Outcome metric           | Brief Symptom Inventory (BSI-18) to assess mental health or symptoms of psychological distress. Response categories for each item ranged from one (not at all) to five (extremely). These items consistently demonstrate adequate validity, reliability, and principal components in previous research. |

| Study ID                       | Jacobs, 2016   |
|--------------------------------|--|
| Statistical analysis           | In lieu of lagged endogenous dependent variable models, change score models to assess two-year changes in distress were employed.  |
| Employment status at baseline  | All unemployed   |
| Employment status at follow-up | 62.1% remained unemployed<br>36.7% acquired employment (6.7% with child care conflict, 31.2% without)  |
| Results                        | The bars for those respondents who remained unemployed or who acquired employment with child care conflict tend to extend above zero. Bars extending above zero indicate average increases in distress symptoms over the study period. The bars for those respondents who acquired employment without child care conflict tend to extend below zero. Bars extending below zero indicate average reductions in distress symptoms over the study period. |
| Study quality                  | Moderate (+)   |

| Study ID                 | Reine, 2008   |
|--------------------------|---|
| Country                  | Sweden  |
| Study Population (N)     | N = 1,044 (96.4% initial response rate, N women = 497, N men = 547)<br><br>This group consisted of people who were in an unstable labour market position for a year or more, between 25-29, and who had acquired a permanent job one year before and at the time of the investigation.  |
| Study Population         | Precariously employed and unemployed general population   |
| Study type               | Longitudinal  |
| Definition of employment | An unstable labour market position was defined as being in unemployment, occasional jobs or labour market programmes. The variable "labour market position" was defined independent of having part-time or full-time employment.<br><br>1. Reference group (without permanent employment): a) $\geq$ one year of an unstable labour market position between ages 25-29 and b) unstable labour market position for the last year prior to the investigation and c) being in an unstable labour market position at the time of the investigation at age 30 (n = 73).<br>2. Obtained permanent employment: a) $\geq$ one year of an unstable labour market position between ages 25-29 and b) having permanent employment one year prior to and c) at the time of the investigation at age 30 (n = 72).<br>3. Constant permanent employment: a) $<$ one year of an unstable labour market position between ages 25-29 and b) having permanent employment one year prior to and c) having permanent employment at the time of the investigation at age 30 (n = 572). Those who did not belong to any of these three groups were excluded from the analyses (n = 322). |

| Study ID                       | Reine, 2008   |
|--------------------------------|---|
| Length of follow-up            | 14 years (ages 16, 18, 21 and 30)   |
| Duration of employment         | Not clear   |
| Duration of unemployment       | Not unemployment but precarious employment  |
| Outcome type: health           | Yes   |
| Outcome type: financial        | No  |
| Outcome type: Other            | No  |
| Outcome metric                 | The outcome variable psychological symptoms at age 30 were measured with a well-known and validated scale. The scale consisted of questions about symptoms during the last year with the answer alternatives "yes" (coded as 1) or "no" (coded as 0). The index of psychological symptoms included six items on restlessness, concentration problems, nervousness, palpitations, anxiety and other nervous symptoms. The range of the index was from nought to six, with higher values corresponding to more psychological problems. The proportion over the cut-off point ( $\geq 1$ ) was defined as those with psychological symptoms. Those below the cut-off point ( $<1$ ) were defined as not having psychological symptoms. |
| Statistical analysis           | <p>A multivariate logistic regression was performed in order to analyse whether the transition from an unstable labour market position to permanent employment was associated with psychological symptoms.</p> <p>A separate bivariate analysis was performed to analyse the association between psychological symptoms at age 30 and the variable labour market positions.</p>   |
| Employment status at baseline  | <p>Men: n = 42 for those without permanent employment, n = 46 for those who obtained permanent employment and n = 324 for those with permanent employment.</p> <p>Women: n = 31 for those without permanent employment, n = 26 for those who obtained permanent employment and n = 248 for those with permanent employment.</p>   |
| Employment status at follow-up | A separate analysis of labour market activities among those without permanent employment at age 25, ie, during the first year of the exposure period for the variable "unstable labour market position", showed that 32% were unemployed, 15% had occasional jobs, 14% were in labour market programmes while the rest were in studies or in permanent employment (data not shown in the tables). As regards the excluded group, the majority had permanent jobs (62%), while 25% were studying.  |

| <b>Study ID</b> | <b>Reine, 2008</b>  |
|-----------------|---|
| Results         | <p>Association was found between the lower probability of psychological symptoms and obtaining permanent employment (OR = 0.35, 95% CI 0.19–0.63) as well as having permanent employment (OR = 0.22, 95% CI 0.10–0.51).</p> <p>A separate bivariate analysis was performed to analyse the association between psychological symptoms at age 30 and the variable labour market positions (data not shown in the tables). For those who obtained permanent employment the odds ratio for psychological symptoms was 0.26 (95% CI 0.16–0.43) and for those in permanent employment the odds ratio was 0.19 (95% CI 0.09–0.41).</p> |
| Study quality   | Strong (++)   |

| <b>Study ID</b>          | <b>Schneider, 2009</b>   |
|--------------------------|--|
| Country                  | United Kingdom   |
| Study Population (N)     | <p>N = 174</p> <p>Although only one of the agencies was organisationally linked to a provider of mental health services nearly all the respondents had taken medication and/or consulted a mental health professional in the previous two years.</p> |
| Study Population         | Unemployed population with severe mental health disorders  |
| Study type               | Pre-post study   |
| Definition of employment | No specific definition of employment given.  |
| Length of follow-up      | 12 months  |
| Duration of employment   | 32 people who went to work did so for a mean of 27 hours per week (standard deviation 14, median 27, mode 35 hours)  |
| Duration of unemployment | Duration of unemployment for those who remained unemployed not given.  |
| Outcome type: health     | No   |
| Outcome type: financial  | Yes  |
| Outcome type: Other      | Yes  |
| Outcome metric           | Service use and frequency --> costs analysed from government perspective (excluding earnings) and a societal perspective (excluding welfare benefits and taxes)  |

| Study ID                       | Schneider, 2009   |
|--------------------------------|---|
| Statistical analysis           | <p>At baseline and follow up interviews people were asked to recall their use in the preceding three months of an exhaustive list of services. If they had used a service, they were asked to judge how much or for how long. With this information and published generic costs (Netten &amp; Curtis, 2005) average costs per week at baseline and a year later were calculated. The rationale behind this approach has been well established (Netten &amp; Beecham, 1993).</p> <p>In the analysis of service costs, inputs are aggregated by the providing agency to make inferences more robust. Unlike the other costs reported here, employment-related input was calculated by the project grant administrator on the basis of data from each agency.</p> <p>Descriptive statistics were used to profile the study participants. Mean costs at baseline and follow-up were estimated for the sample as a whole and for each group (remained unemployed, working less than one year, working more than one year).</p> |
| Employment status at baseline  | 141 – all unemployed  |
| Employment status at follow-up | N employed = 32<br>N unemployed = 77  |

| Study ID      | Schneider, 2009  |
|---------------|--|
| Results       | <p><b>**Tables with figures in report</b></p> <p>People who had entered work within the past year were the only group whose total service related costs, other than supported employment, fell significantly (<math>p=0.002</math>). This was due principally to the reduction in mental health service costs (<math>p&lt;0.001</math>). The group of people in longer-term employment also saw their mental health service use fall significantly over the course of the study, but the drop in their overall service use did not quite attain statistical significance (<math>p=0.067</math>). By contrast, the costs of supported employment increased significantly for people who had worked for less than one year (<math>p=0.04</math>) while they declined for those who remained unemployed (<math>p&lt;0.001</math>) as well as for those who had been working for longer (<math>p=0.002</math>). Figure 1 profiles the differences in costs of mental health services and supported employment for each sub-group. There was no correlation between the time with the agencies and supported employment costs at either time point (baseline Pearson's <math>r=-0.061</math>, follow-up <math>r=-0.065</math>). Neither was there any correlation between time with the agencies and other costs aggregated (baseline <math>r=0.069</math>, follow-up <math>r=0.041</math>). Hours worked per week did correlate negatively with supported employment costs at baseline (<math>r=-0.0259</math>, <math>p=0.085</math>) and at follow up (<math>r=-0.0265</math>, <math>p=0.078</math>) but these correlations did not attain statistical significance.</p> <p>Income from benefit entitlements went down across both groups, but (contrary to what one might expect) did not decline significantly more for people entering work (<math>p=0.72</math>, independent t-test). The mean entitlement to allowances such as help with housing, carers' allowance and child benefit did drop for those entering work, but this did not attain statistical significance (<math>p=0.16</math>). Since some benefits (eg housing benefit) are income related, this may reflect the fact that the new workers include some low earners, and that more than half of them worked part-time.</p> |
| Study quality | Strong (++)  |

| Study ID                 | Schuring 2011  |
|--------------------------|--|
| Country                  | Netherlands  |
| Study Population (N)     | 2754 persons on social security benefits who were capable of full-time employment and who were referred by the Employment Centre of the City of Rotterdam, The Netherlands, to one of four, re-employment training centres in the area for re-employment training. |
| Study Population         | Unemployed general population receiving social security benefits   |
| Study type               | Longitudinal   |
| Definition of employment | Ending social security benefits for at least three months because of starting with paid employment.  |
| Length of follow-up      | Six months   |
| Duration of employment   | Cox Proportional Hazards analysis  |

| <b>Study ID</b>                | <b>Schuring 2011</b>   |
|--------------------------------|--|
| Duration of unemployment       | At least six months  |
| Outcome type: health           | Yes  |
| Outcome type: financial        | No   |
| Outcome type: Other            | No   |
| Outcome metric                 | Health related quality of life measured with Short Form-36 health survey (SF-36).  |
| Statistical analysis           | Linear regression analysis.  |
| Employment status at baseline  | Unemployed   |
| Employment status at follow-up | Employed/not in employment   |
| Results                        | The response at follow-up was 53% (965/1829). Among the re-employed subjects General Health, Physical Function, Social Functioning, Vitality, Mental Health, Bodily Pain, and role-limitations due to emotional or physical problems improved during the follow-up period, with an effect size varying from 0.11 to 0.66. The largest relative improvements were observed for Mental Health, Social Functioning, and role-limitations due to emotional or physical problems, whereas Physical Function showed the smallest relative improvement. For those subjects who remained unemployed the effect sizes varied from -0.04 to 0.06, indicating that their health status remained virtually unchanged during the six months follow-up period. |
| Study quality                  | Strong (++)  |

| <b>Study ID</b>          | <b>Steele 2013</b>   |
|--------------------------|--|
| Country                  | United Kingdom   |
| Study Population (N)     | The analysis sample contained 8784 men who contributed 69,576 person-year observations between 1991 and 2009 on transitioning from unemployment to employment via the British Household Panel Survey |
| Study Population         | Unemployed male general population   |
| Study type               | Observational  |
| Definition of employment | Any paid employment (as opposed to unemployment or economic inactivity)  |
| Length of follow-up      | 16 years   |
| Duration of employment   | Not reported   |
| Duration of unemployment | Not reported   |
| Outcome type: health     | Yes  |

| <b>Study ID</b>                | <b>Steele 2013</b>  |
|--------------------------------|---|
| Outcome type: financial        | No  |
| Outcome type: Other            | No  |
| Outcome metric                 | Mental health effects: Psychological distress and anxiety using the General Health Questionnaire  |
| Statistical analysis           | Simultaneous equations model in which each outcome may directly affect the other over time  |
| Employment status at baseline  |   |
| Employment status at follow-up |   |
| Results                        | Moving from unemployment to employment was strongly associated with an improvement in mental health, whereas becoming unemployed was detrimental. However, these associations were attenuated by unmeasured confounders. After adjustment for indirect selection, the increased distress and anxiety associated with becoming unemployed decreased from 2.5 (95% confidence interval = 2.2 to 2.7) to 2.2 (2.0 to 2.5). (A change of 2.5 equates to half a standard deviation on the 12-point scale.) The improvement with moving from unemployment to employment was also weakened slightly (from -2.1 [-2.4 to -1.7] to -1.8 [-2.1 to -1.5]). |
| Study quality                  | Strong (++)   |

| <b>Study ID</b>          | <b>Thomas 2007</b>  |
|--------------------------|---|
| Country                  | United Kingdom  |
| Study Population (N)     | Participants of the British Household Panel Survey from 1991 to 2000. These included individuals who contributed at least two consecutive years to the study and whose complete data were available for the analyses. |
| Study Population         | Unemployed general population   |
| Study type               | Longitudinal  |
| Definition of employment | Paid employment (full-time or part-time)  |
| Length of follow-up      | 10 years  |
| Duration of employment   | Since last follow up  |
| Duration of unemployment | Not reported  |
| Outcome type: health     | Yes   |
| Outcome type: financial  | No  |
| Outcome type: Other      | No  |

| <b>Study ID</b>                | <b>Thomas 2007</b>   |
|--------------------------------|--|
| Outcome metric                 | The outcome used in this study was psychological distress, measured by the 12-item General Health Questionnaire (GHQ-12). Individuals with a GHQ-12 score of >3, which is the average for the sample, were classified as cases of psychological distress.  |
| Statistical analysis           | Random-effects logistic regression   |
| Employment status at baseline  | Unemployed   |
| Employment status at follow-up | Employed or unemployed   |
| Results                        | Both men and women who returned to work from unemployment were less likely to experience distress (OR 0.52 and 0.68, respectively) than those who remained in non-employment. However, interaction testing found that the beneficial effects of transitions from unemployment to employment were limited to those with distress while unemployed (OR 0.40 (95% CI 0.25 to 0.64) for men; OR 0.62 (95% CI 0.35 to 1.10) for women). The beneficial effect of transitions from family care to paid employment on mental health was also limited to women who experienced psychological distress before the transition (OR 0.48 (95% CI 0.24 to 0.96)). |
| Study quality                  | <b>Strong (++)</b>   |

| <b>Study ID</b>          | <b>van de Noordt 2014</b>  |
|--------------------------|--|
| Country                  | International study/Systematic Review  |
| Study Population (N)     | Review included longitudinal studies of adult populations in transition from unemployment to employment. |
| Study Population         | Unemployed general population  |
| Study type               | Systematic review  |
| Definition of employment | Not reported. Inclusion of only longitudinal studies suggestive of long-term employment.                 |
| Length of follow-up      | Not reported   |
| Duration of employment   | Not reported   |
| Duration of unemployment | Not reported   |
| Outcome type: health     | <b>Yes</b>   |
| Outcome type: financial  | <b>No</b>  |
| Outcome type: Other      | <b>No</b>  |

| <b>Study ID</b>                | <b>van de Noordt 2014</b>  |
|--------------------------------|--|
| Outcome metric                 | Mental health effects:<br>Depression – Hopkins Symptom Checklist, Beck Depression Inventory, CES-Depression Scale, Rosenberg’s Depressive Affect Scale and the Edinburgh Postnatal Depression Scale.<br><br>Psychological distress – the General Health Questionnaire (GHQ), CES-Depression Scale  |
| Statistical analysis           | Meta-analyses (both random and fixed effect models)  |
| Employment status at baseline  | Unemployed   |
| Employment status at follow-up | Employed or unemployed   |
| Results                        | Thirty-three prospective studies were included, of which 23 were of high quality. Strong evidence was found for a protective effect of employment on depression and general mental health. Pooled effect sizes showed favourable effects on depression (OR=0.52; 95% CI 0.33 to 0.83) and psychological distress (OR=0.79; 95% CI 0.72 to 0.86). A relative risk (RR) of 0.45 (p<0.05) and an OR of 0.26 (95% CI 0.10 to 0.64) after a follow-up of two and five years, respectively, for the risk on depression for re-employed compared with still unemployed people. Insufficient evidence was found for general health, physical health and mortality due to lack of studies or inconsistent findings. |
| Study quality                  | <b>Strong (++)</b>   |

| <b>Study ID</b>          | <b>van Rijn 2015</b>   |
|--------------------------|--|
| Country                  | International study/Systematic Review  |
| Study Population (N)     | Long-term unemployed persons with severe mental health problems.   |
| Study Population         | Unemployed population with severe mental health disorders  |
| Study type               | Systematic review  |
| Definition of employment | A job in the open labour market that anyone could hold with a minimum wage or higher for at least 30 days. |
| Length of follow-up      | Not reported   |
| Duration of employment   | At least 30 days   |
| Duration of unemployment | Not reported   |
| Outcome type: health     | <b>Yes</b>   |
| Outcome type: financial  | <b>No</b>  |
| Outcome type: Other      | <b>No</b>  |

| Study ID                       | van Rijn 2015  |
|--------------------------------|--|
| Outcome metric                 | <p>Various:</p> <p>Functioning – measured using validated scales like Global Assessment of Functioning (GAF) and Short Form-36 health survey (SF-36).</p> <p>Mental health – Positive and Negative Syndrome Scale (PANSS), Hospital Anxiety and Depression Scale (HADS), Rosenberg self-esteem scale, Brief Psychiatric Rating Scale (BPRS), and Chinese General Self-efficacy Scale (CGSS)</p> <p>Quality of life – the Manchester Short Assessment of Quality of Life (MANSA), the Lancashire Quality of Life Profile, Lehman’s Quality of life interview, the Wisconsin Quality of Life Index or the WHO Quality of Life Measure</p>  |
| Statistical analysis           | Random effect meta-analysis  |
| Employment status at baseline  | Unemployed   |
| Employment status at follow-up | Competitive employment or unemployed   |
| Results                        | <p>Functioning – nine studies described functioning as an outcome measure. The pooled effect size was <math>-0.01</math> (95% CI <math>-0.13</math> to <math>0.11</math>). Within these studies, the percentage of competitive employed ranged from 13-55% in the intervention groups compared to 6-28% in the control groups. The meta-regression analysis showed no influence of study characteristics on the effect size of functioning.</p> <p>Mental health – 15 studies described mental health as an outcome measure. Six studies, describing 12 comparisons, provided enough information to calculate effect sizes ranging from <math>-0.58</math> to <math>1.39</math>. Six studies measured psychiatric symptoms and the pooled effect size was <math>0.20</math> (95% CI <math>-0.23</math> to <math>0.62</math>). Within these studies, the percentage of competitive employment ranged from 13-55% in the intervention groups compared to 2-28% in the control groups. The meta-regression analysis showed that studies with over 50% of male participants had a lower effect size of mental health (<math>\beta -1.43</math>, 95% CI <math>-2.12</math> to <math>-0.74</math>).</p> <p>Quality of life – 13 studies described quality of life as an outcome measure. Seven studies, describing eight comparisons, provided enough information to calculate effect sizes, ranging from <math>0.00</math> to <math>0.99</math> with a pooled effect size of <math>0.28</math> (95% CI <math>0.04</math> to <math>0.52</math>). Within these studies, the percentage of competitive employment ranged from 13-74% in the intervention groups compared to 2-68% in the control groups.</p> |
| Study quality                  | Moderate (+)   |

## I. In depth analysis of health scales and ratings

|  |   |
|--|---|
| <b>Name</b>                                  | <b>36-item Short Form Survey (SF-36)</b>  |
| <b>Category</b>                              | Mental and physical health  |
| <b>Description</b>                           | The Short Form-36 (SF-36) is a 36-item questionnaire that measures Quality of Life (QoL) across eight domains, which are both physically and emotionally based. The eight domains are: Physical Function; Role limitations due to physical health; Role limitations due to emotional problems; Bodily Pain; Vitality; Social Functioning; General Health; Mental Health. A single item is also included that identifies perceived change in health, making the SF-36 a useful indicator for change in QoL over time and treatment. The SF-36 has been widely validated for numerous professions and patient groups and can be administered by clinicians or by the patient at home. |
| <b>Scale</b>                                 | 0-100 range<br>No normative values or cut-off scores are, therefore, presented. However, the scoring of the SF-36 indicates that 0% in a domain represents the poorest possible QoL and 100% indicates full QoL (the best possible result). Taking this into account, it is easy to see that higher scores on the SF-36 indicate better QoL.  |
| <b>Sources</b>                               | <a href="http://www.rand.org/health/surveys_tools/mos/36-item-short-form.html">http://www.rand.org/health/surveys_tools/mos/36-item-short-form.html</a><br><a href="http://www.measuringimpact.org/s4-sf-36">http://www.measuringimpact.org/s4-sf-36</a>  |
| <b>References</b>                            | Zvonimar, 2008; Carlier, 2013; Schuring, 2011   |
| <b>Quality</b>                               | <b>Strong (++)</b>  |
| <b>Reason for quality rating</b>             | Validated questionnaire; used in a number of studies  |
| <b>Questionnaires</b>                        | <a href="http://www.shcdenver.com/Portals/902/web-content/files/JamesGenuario/JG-health%20questionnaire.pdf">http://www.shcdenver.com/Portals/902/web-content/files/JamesGenuario/JG-health%20questionnaire.pdf</a>   |
| <b>Include in model - Yes/No with reason</b> | Yes – validated questionnaire, used in a number of included studies. There is a robust methodology to convert findings into utilities, subsequently QALYs, thus reporting on health changes and any other associated benefits.  |

|                                       |   |
|---------------------------------------|---|
| <b>Name</b>                           | <b>Swedish Living Conditions Surveys</b>  |
| Category                              | Mental health   |
| Description                           | The Swedish Living Conditions Surveys give information on living conditions among different groups in Swedish society. The areas covered include housing, income, health, leisure, civic activities, social relationships, employment and security. A number of the questions are the same as those in surveys of other EU countries, so it is possible to make comparisons with other countries within the EU.   |
| Scale                                 | Specific to the paper that utilised results from the survey: The scale consisted of questions about symptoms during the last year with the answer alternatives "yes" (coded as 1) or "no" (coded as 0). The index of psychological symptoms included six items on restlessness, concentration problems, nervousness, palpitations, anxiety and other nervous symptoms. The range of the index was from nought to six, with higher values corresponding to more psychological problems. The proportion over the cut-off point ( $\geq 1$ ) was defined as those with psychological symptoms. Those below the cut-off point ( $<1$ ) were defined as not having psychological symptoms. |
| Sources                               | <a href="http://www.scb.se/en/finding-statistics/statistics-by-subject-area/living-conditions/living-conditions/living-conditions-surveys-ulfslc/">http://www.scb.se/en/finding-statistics/statistics-by-subject-area/living-conditions/living-conditions/living-conditions-surveys-ulfslc/</a>   |
| References                            | Reine, 2008   |
| Quality                               | Moderate (+)  |
| Reason for quality rating             | Based only on the Swedish population but likely to be somewhat applicable to the UK   |
| Questionnaires                        | <a href="#">Unable to find</a>  |
| Include in model - Yes/No with reason | No – unable to convert into useable metric to calculate health benefits.  |

|                                       |  |
|---------------------------------------|--|
| <b>Name</b>                           | <b>Hopkins Symptom Checklist (HSCL)</b>  |
| Category                              | Mental health  |
| Description                           | The HSCL is comprised of 58 items representative of the symptom configurations commonly observed among outpatients. It is scored on five underlying symptom dimensions – somatization, obsessive-compulsive, interpersonal sensitivity, anxiety and depression – which have been identified in repeated factor analyses. |
| Scale                                 | 58-232   |
| Sources                               | <a href="http://onlinelibrary.wiley.com/doi/10.1002/bs.3830190102/abstract">http://onlinelibrary.wiley.com/doi/10.1002/bs.3830190102/abstract</a>  |
| References                            | Van de Noordt, 2014  |
| Quality                               | Moderate (+)   |
| Reason for quality rating             | Number of journals discussing the validity and reliability of the scale  |
| Questionnaires                        | <a href="http://www.interactionscounseling.com/downloads/HopkinsSymptomChecklist.pdf">http://www.interactionscounseling.com/downloads/HopkinsSymptomChecklist.pdf</a>  |
| Include in model - Yes/No with reason | No – included in the Van Rijn review as part of a pooled analysis where results were not in a useable format and could not be converted.   |

|                                       |   |
|---------------------------------------|---|
| <b>Name</b>                           | <b>Beck Depression Inventory</b>  |
| Category                              | Mental health   |
| Description                           | The Beck Depression Inventory Second Edition (BDI-II) is a 21-item self-report instrument intended to assess the existence and severity of symptoms of depression as listed in the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders Fourth Edition (DSM-IV; 1994). BDI has been used for 35 years to identify and assess depressive symptoms, and has been reported to be highly reliable regardless of the population. |
| Scale                                 | 0-63<br>17-20 --> borderline depression<br>21-30 --> moderate depression<br>31-40 --> severe depression<br>> 40 --> extreme depression  |
| Sources                               | <a href="http://academicdepartments.musc.edu/family_medicine/rcmar/beck.htm">http://academicdepartments.musc.edu/family_medicine/rcmar/beck.htm</a>   |
| References                            | Van de Noordt, 2014   |
| Quality                               | Strong (++)   |
| Reason for quality rating             | Been in operation for 35 years and been reported to be highly reliable regardless of population   |
| Questionnaires                        | <a href="http://www.hr.ucdavis.edu/asap/pdf_files/Beck_Depression_Inventory.pdf">http://www.hr.ucdavis.edu/asap/pdf_files/Beck_Depression_Inventory.pdf</a>   |
| Include in model - Yes/No with reason | No – included in the Van de Noordt review as part of a pooled analysis.   |

|  |  |
|--|--|
| <b>Name</b>                                  | <b>Centre for Epidemiologic Studies Depression Scale Revised (CESD-R)</b>  |
| <b>Category</b>                              | Mental health  |
| <b>Description</b>                           | <p>The CESD-R is a screening test for depression and depressive disorder. The CESD-R measures symptoms defined by the American Psychiatric Association' Diagnostic and Statistical Manual (DSM-V) for a major depressive episode.</p> <p>It is a 20-item measure that asks caregivers to rate how often over the past week they experienced symptoms associated with depression, such as restless sleep, poor appetite, and feeling lonely. Response options range from nought to three for each item (0 = Rarely or None of the Time, 1 = Some or Little of the Time, 2 = Moderately or Much of the time, 3 = Most or Almost All the Time). Scores range from 0 to 60, with high scores indicating greater depressive symptoms. The CES-D also provides cut-off scores (eg 16 or greater) that aid in identifying individuals at risk for clinical depression, with good sensitivity and specificity and high internal consistency (Lewinsohn, Seeley, Roberts, &amp; Allen, 1997).</p> |
| <b>Scale</b>                                 | <p>0-60<br/>Higher scores (<math>\geq 16</math>) indicating the presence of more symptamatology.</p> <p>American Psychology Association suggests threshold of <math>\geq 16</math> as a sign of people at risk of clinical depression</p> <p><a href="https://www.ncbi.nlm.nih.gov/pubmed/21312347">https://www.ncbi.nlm.nih.gov/pubmed/21312347</a> --&gt; suggests a threshold of <math>\geq 24</math></p>   |
| <b>Sources</b>                               | <a href="http://cesd-r.com/">http://cesd-r.com/</a>  |
| <b>References</b>                            | Hao, 2008; Van de Noordt, 2014   |
| <b>Quality</b>                               | <b>Strong (++)</b>   |
| <b>Reason for quality rating</b>             | The scale is well known and remains as one of the most widely used instruments in the field of psychiatric epidemiology.   |
| <b>Questionnaires</b>                        | <a href="https://www.outcometracker.org/library/CES-D.pdf">https://www.outcometracker.org/library/CES-D.pdf</a>  |
| <b>Other information</b>                     | <a href="http://www.apa.org/pi/about/publications/caregivers/practice-settings/assessment/tools/depression-scale.aspx">http://www.apa.org/pi/about/publications/caregivers/practice-settings/assessment/tools/depression-scale.aspx</a>  |
| <b>Include in model - Yes/No with reason</b> | <p>No – included in the Van de Noordt review as part of a pooled analysis.</p> <p>With regards to results reported in Hao, these will not be included in the model due to the narrow population (people aged 55-66).</p>   |

|                                       |   |
|---------------------------------------|---|
| <b>Name</b>                           | <b>Rosenberg Self-esteem Scale</b>  |
| Category                              | Mental health   |
| Description                           | The Rosenberg self-esteem scale (RSES), developed by sociologist Dr. Morris Rosenberg ,[1] is a self-esteem measure widely used in social-science research. It uses a scale of 0-30 where, and a score less than 15 may indicate a problematic low self esteem.           |
| Scale                                 | 0 - 30 range<br>0 - 15 = problematic low self-esteem  |
| Sources                               | <a href="http://psych.ut.ee/~jyri/en/Schmitt-Allik_JPSP2005.pdf">http://psych.ut.ee/~jyri/en/Schmitt-Allik_JPSP2005.pdf</a>   |
| References                            | Crowther, 2010; Van de Noordt, 2014; Van Rijn, 2015   |
| Quality                               | Strong (++)   |
| Reason for quality rating             | The RSES has been translated and adapted to various languages, such as Persian, French, Chinese, Italian, German, Portuguese, and Spanish. The scale is extensively used in cross-cultural studies in up to 53 different nations.   |
| Questionnaires                        | <a href="http://fetzer.org/sites/default/files/images/stories/pdf/selfmeasures/Self_Measures_for_Self-Esteem_ROSENBERG_SELF-ESTEEM.pdf">http://fetzer.org/sites/default/files/images/stories/pdf/selfmeasures/Self_Measures_for_Self-Esteem_ROSENBERG_SELF-ESTEEM.pdf</a> |
| Other information                     | Van de Noordt names it the Rosenberg Depressive Effect Scale but nothing online exists for this.  |
| Include in model - Yes/No with reason | Possible - scale does not measure a validated clinical definition therefore but might be included as extra narrative of the effects of a transition.  |

|                           |  |
|---------------------------|--|
| <b>Name</b>               | <b>Edinburgh Postnatal Depression Scale</b>  |
| Category                  | Mental health  |
| Description               | The Edinburgh Postnatal Depression Scale (EPDS) is a 10-item questionnaire that was developed to identify women who have postpartum depression.[1] Items of the scale correspond to various clinical depression symptoms, such as guilt feeling, sleep disturbance, low energy, anhedonia, and suicidal ideation. Overall assessment is done by total score, which is determined by adding together the scores for each of the 10 items. Higher scores indicate more depressive symptoms.[2] The EPDS may be used within 8 weeks postpartum and it also can be applied for depression screening during pregnancy.[3] |
| Scale                     | 0 - 30 scale<br>Possible depression: >= 10   |
| Sources                   | <a href="http://www.fresno.ucsf.edu/pediatrics/downloads/edinburghscale.pdf">http://www.fresno.ucsf.edu/pediatrics/downloads/edinburghscale.pdf</a>  |
| References                | Van de Noordt, 2014  |
| Quality                   | Strong (++)  |
| Reason for quality rating | Validated questionnaire; used in a number of studies   |
| Questionnaires            | <a href="http://www.fresno.ucsf.edu/pediatrics/downloads/edinburghscale.pdf">http://www.fresno.ucsf.edu/pediatrics/downloads/edinburghscale.pdf</a>  |

|  |  |
|--|--|
| <b>Name</b>                              | <b>Edinburgh Postnatal Depression Scale</b>  |
| Other information                        |  |
| Include in model<br>- Yes/No with reason | No - – included in the Van Rijn review as part of a pooled analysis where results were not in a useable format and could not be converted. |

|                           |  |
|---------------------------|--|
| <b>Name</b>               | <b>12-item General Health Questionnaire (GHQ-12)</b>   |
| Category                  | Mental health  |
| Description               | The General Health Questionnaire (GHQ) is a screening device for identifying minor psychiatric disorders in the general population and within community or non-psychiatric clinical settings such as primary care or general medical out-patients. Suitable for all ages from adolescent upward – not children, it assesses the respondent’s current state and asks if that differs from his or her usual state. It is, therefore, sensitive to short-term psychiatric disorders but not to long-standing attributes of the respondent. GHQ-12: a quick, reliable and sensitive short form – ideal for research studies. This version of the GHQ is very quick to administer and score as it contains only 12 questions. It has comparable psychometric properties to the longer versions even though it only takes around two minutes to complete. Given its speed of administration, it is often used in research studies where it is impractical to administer a longer form. |
| Scale                     | 1/2 (max score 12)<br>Higher scores indicate a higher level of distress  |
| Sources                   | <a href="https://www.gi-assessment.co.uk/products/general-health-questionnaire-ghq/">https://www.gi-assessment.co.uk/products/general-health-questionnaire-ghq/</a>  |
| References                | Hoare, 2010; Steele, 2013; Thomas, 2007; Van de Noordt, 2014, Van Rijn, 2015   |
| Quality                   | <b>Strong (++)</b>   |
| Reason for quality rating | Validated questionnaire; used in a number of studies   |
| Questionnaires            | <a href="#">Unable to find</a>   |

| Name                                  | <b>12-item General Health Questionnaire (GHQ-12)</b>  |
|---------------------------------------|---|
| Other information                     | <p>Paper discussing threshold levels for GHQ-12 and MHI-5:<br/> <a href="http://hqlo.biomedcentral.com/articles/10.1186/1477-7525-2-23">http://hqlo.biomedcentral.com/articles/10.1186/1477-7525-2-23</a></p> <p>The GHQ-12, a 12-item questionnaire, is used to identify people likely to have mental problems, and who, therefore, may need health care. Cut-off point <math>\geq 2</math> was used [1]. The MHI-5 measures general mental health [5] and is part of the Short Form-36 health survey (SF-36), a questionnaire for measuring health-related quality of life [10]. The MHI-5 has a score of 0 to 100, where a score of 100 represents optimal mental health. The mean scores as well as the prevalence of mental problems were reported. Because no formal cut-off point is determined, the cut-off score for the MHI-5 (<math>\leq 72</math>) that generated the most comparable prevalence of mental problems as the GHQ-12 cut-off point <math>\geq 2</math> was used.</p> |
| Include in model - Yes/No with reason | No – although a validated questionnaire used in a number of included studies, it was not possible to convert scores into useable metrics in the tool.   |

| Name                                  | <b>10-point Visual Analogue Scale (VAS)</b>  |
|---------------------------------------|--|
| Category                              | Quality of Life  |
| Description                           | A Visual Analogue Scale (VAS) is a measurement instrument that tries to measure a characteristic or attitude that is believed to range across a continuum of values and cannot easily be directly measured.[1] It is often used in epidemiologic and clinical research to measure the intensity or frequency of various symptoms |
| Scale                                 | 0-10   |
| Sources                               | <a href="http://www.physio-pedia.com/Visual_Analogue_Scale">http://www.physio-pedia.com/Visual_Analogue_Scale</a>  |
| References                            | Carlier, 2013  |
| Quality                               | <b>Weak (-)</b>  |
| Reason for quality rating             | Subjective visual scale – difficult to generalise the results  |
| Questionnaires                        |  |
| Other information                     |  |
| Include in model - Yes/No with reason | No – scale does not measure a validated clinical definition, therefore, it has been excluded.  |

|                                       |  |
|---------------------------------------|--|
| <b>Name</b>                           | <b>12-item Short Form Survey (SF-12)</b>   |
| Category                              | Mental and physical health   |
| Description                           | The SF-12v2™ Health Survey is a 12-item subset of the SF-36v2™ that measures the same eight domains of health. It is a brief, reliable measure of overall health status. It is useful in large population health surveys and has been used extensively as a screening tool     |
| Scale                                 | 0-100  |
| Sources                               | <a href="http://qol.thoracic.org/sections/instruments/pt/pages/sf12.html">http://qol.thoracic.org/sections/instruments/pt/pages/sf12.html</a>  |
| References                            | Curnock, 2016;   |
| Quality                               | <b>Strong (++)</b>   |
| Reason for quality rating             | Validated questionnaire  |
| Questionnaires                        | <a href="http://health.utah.gov/oph/publications/2001hss/sf12/SF12_Interpreting.pdf">http://health.utah.gov/oph/publications/2001hss/sf12/SF12_Interpreting.pdf</a>  |
| Include in model - Yes/No with reason | No – validated questionnaire, used in a number of our included studies. SF-12 is derived from the SF-36, however, the algorithm used to convert SF-36 scores into a useable metric (in this case, the SF-6D which is used to convert scores into QALYs). is unable to be used. |

|                                       |   |
|---------------------------------------|---|
| <b>Name</b>                           | <b>Manchester Short Assessment of Quality of Life (MANSA)</b>   |
| Category                              | Quality of Life   |
| Description                           | Based on experiences and empirical evidence gained in studies using the Lancashire Quality of Life Profile (LQLP), the Manchester Short Assessment of Quality of Life (MANSA) has been developed as a condensed and slightly modified instrument for assessing quality of life. Its properties have been tested in a sample of community care patients. |
| Scale                                 |   |
| Sources                               | <a href="http://europepmc.org/abstract/med/10443245">http://europepmc.org/abstract/med/10443245</a>   |
| References                            | Van Rijn, 2015  |
| Quality                               | <b>Weak (-)</b>   |
| Reason for quality rating             | Not able to find concrete information on the scale  |
| Questionnaires                        | <a href="#">Unable to find</a>  |
| Include in model - Yes/No with reason | No – included in the Van Rijn review as part of a pooled analysis where results were not in a useable format and could not be converted.  |

| Name                                  | Lancashire Quality of Life Profile  |
|---------------------------------------|---|
| Category                              | Quality of Life   |
| Description                           | The Lancashire Quality of Life Profile is an instrument for obtaining a detailed quality of life profile, encompassing general quality of life as well as satisfaction with different life domains and subjective as well as interviewer-assessed quality of life. Devised for people with schizophrenia and other psychoses (Oliver, Huxley, Bridges & Mohamad, 1996) it was derived from Lehman's quality of life interview (Lehman, Ward & Linn, 1982). Besides a global rating of quality of life, the original LQOLP comprises nine life domains: living situation, family, social relationships, leisure activities, work/education, finances, personal safety, health, and religion. |
| Scale                                 |   |
| Sources                               | <a href="http://link.springer.com/referenceworkentry/10.1007%2F978-94-007-0753-5_1593">http://link.springer.com/referenceworkentry/10.1007%2F978-94-007-0753-5_1593</a>   |
| References                            | Van Rijn, 2015  |
| Quality                               | Moderate (+)  |
| Reason for quality rating             | Only one study that uses this scale   |
| Questionnaires                        | <a href="#">Unable to find</a>  |
| Include in model - Yes/No with reason | No – included in the Van Rijn review as part of a pooled analysis where results were not in a useable format and could not be converted.  |

| Name                                  | Lehman's Quality of Life Profile  |
|---------------------------------------|---|
| Category                              | Quality of Life   |
| Description                           | An instrument that evaluates both subjective and objective components of quality of life. The QoLI is a structured questionnaire, with three types of reply for each life domain: dichotomous replies (yes, no), open responses to reveal objective information such as type of residence, and replies located on a seven-point Likert scale. The original QoLI contains 143 items and takes about 45 minutes to score. |
| Sources                               | <a href="http://www.rees-france.com/en/IMG/pdf/ART-859_QoI_Internal_structure.pdf">http://www.rees-france.com/en/IMG/pdf/ART-859_QoI_Internal_structure.pdf</a>   |
| References                            | Van Rijn, 2015  |
| Quality                               | Moderate (+)  |
| Reason for quality rating             | Only one study that uses this scale   |
| Questionnaires                        | <a href="#">Unable to find</a>  |
| Include in model - Yes/No with reason | No – included in the Van Rijn review as part of a pooled analysis where results were not in a useable format and could not be converted.  |

|                                       |  |
|---------------------------------------|--|
| <b>Name</b>                           | <b>Wisconsin Quality of Life interview</b>   |
| Category                              | Quality of Life  |
| Description                           | The Wisconsin Quality of Life Index (W-QLI) is a multidimensional model for measuring quality of life. This model assumes that quality of life is comprised of nine dimensions: life satisfaction; occupational activities; psychological wellbeing; physical health; social relations; economics; activities of daily living; symptoms; patient's own goals. The W-QLI takes into account the different perspectives of the patient, the patient's family, and the clinician, both currently and over time. |
| Scale                                 |  |
| Sources                               | <a href="http://wqli.fmhi.usf.edu/">http://wqli.fmhi.usf.edu/</a>  |
| References                            | Van Rijn, 2015   |
| Quality                               | Moderate (+)   |
| Reason for quality rating             | Only one study uses this scale   |
| Questionnaires                        | <a href="http://wqli.fmhi.usf.edu/docs/knowledge-base/wqli-manual/WQLI%20Questionnaire%20Client%20Coding&amp;Scoring.pdf">http://wqli.fmhi.usf.edu/docs/knowledge-base/wqli-manual/WQLI Questionnaire Client Coding&amp;Scoring.pdf</a>  |
| Other information                     |  |
| Include in model - Yes/No with reason | No – included in the Van Rijn review as part of a pooled analysis where results were not in a useable format and could not be converted.   |

| Name                                  | WHO Quality of Life  |
|---------------------------------------|--|
| Category                              | Quality of Life  |
| Description                           | WHOQOL-BREF is an abbreviated generic Quality of Life Scale developed through the World Health Organization. The SeaQoL Group is the location of the US WHOQOL Center and distributes the US English version of the WHOQOL-100 and WHOQOL-BREF. The WHOQOL-Bref, still in field trials, is a subset of 26 items taken from the WHOQOL-100. |
| Scale                                 | 0-100  |
| Sources                               | <a href="http://depts.washington.edu/seaqol/docs/WHOQOL-BREF%20with%20scoring%20instructions_Updated%2001-10-14.pdf">http://depts.washington.edu/seaqol/docs/WHOQOL-BREF%20with%20scoring%20instructions_Updated%2001-10-14.pdf</a>  |
| References                            | Van Rijn, 2015   |
| Quality                               | Strong (++)  |
| Reason for quality rating             | Reputable organisation (World Health Organization); asks a wide variety of questions.  |
| Questionnaires                        | <a href="http://www.who.int/mental_health/media/en/76.pdf?ua=1">http://www.who.int/mental_health/media/en/76.pdf?ua=1</a>  |
| Other information                     |  |
| Include in model - Yes/No with reason | No – included in the Van Rijn review as part of a pooled analysis where results were not in a useable format and could not be converted.   |

|                                       |  |
|---------------------------------------|--|
| <b>Name</b>                           | <b>Brief Psychiatric Rating Scale (BPRS)</b>   |
| Category                              | Mental health  |
| Description                           | The Brief Psychiatric Rating Scale (BPRS) is a widely used instrument for assessing the positive, negative, and affective symptoms of individuals who have psychotic disorders, especially schizophrenia. It has proven particularly valuable for documenting the efficacy of treatment in patients who have moderate to severe disease. The rater should enter a number ranging from one (not present) to seven (extremely severe). Nought is entered if the item is not assessed. There are 24 conditions that are considered including depression, anxiety, emotional withdrawal and guilt. |
| Scale                                 | 24-168 (if all 24 disorders considered)  |
| Sources                               | <a href="http://farmacologiaclinica.info/scales/BPRS/">http://farmacologiaclinica.info/scales/BPRS/</a>  |
| References                            | Crowther, 2010; Van Rijn, 2015   |
| Quality                               | Strong (++)  |
| Reason for quality rating             | One of the oldest, widely used scales to measure psychotic symptoms. It was first published in 1962.   |
| Questionnaires                        | <a href="http://www.public-health.uiowa.edu/icmha/outreach/documents/bprs_expanded.pdf">http://www.public-health.uiowa.edu/icmha/outreach/documents/bprs_expanded.pdf</a>  |
| Other information                     | See training powerpoint (s:drive - saved as 'bprs24training' --> Suggests a score of >= 4 in each categorisation is considered 'moderate' symptoms   |
| Include in model - Yes/No with reason | No – included in the Van Rijn review as part of a pooled analysis where results were not in a useable format and could not be converted. Also mentioned in Crowther paper but difficult to interpret results.  |

| Name                                  | Positive and Negative Syndrome Scale (PANSS)  |
|---------------------------------------|---|
| Category                              | Subjective wellbeing and mental health  |
| Description                           | The PANSS is based on findings that schizophrenia comprises at least two distinct syndromes. The positive syndrome consists of productive symptoms, while the negative syndrome consists of deficit features. This distinction is useful when developing treatment plans because focus can be on the type of symptoms the patient is experiencing. It is also useful when studying the effects of medication (eg in clinical drug trials) because it allows which type of symptoms are being affected to be determined. |
| Scale                                 | Positive scale --> 7 items; min. score = 7, max score = 49<br>Negative scale --> 7 items; min score = 7, max score = 49<br>General psychopathology scale --> 16 items; min score = 16, max score = 112  |
| Sources                               | <a href="http://www.pearsonclinical.co.uk/Psychology/AdultMentalHealth/AdultPsychiatricDisorder/PositiveandNegativeSyndromeScale(PANSS)/PositiveandNegativeSyndromeScale(PANSS).aspx">http://www.pearsonclinical.co.uk/Psychology/AdultMentalHealth/AdultPsychiatricDisorder/PositiveandNegativeSyndromeScale(PANSS)/PositiveandNegativeSyndromeScale(PANSS).aspx</a>   |
| References                            | Van Rijn, 2015  |
| Quality                               | Moderate (+)  |
| Reason for quality rating             | Only one study uses this scale  |
| Questionnaires                        | <a href="http://egret.psychol.cam.ac.uk/medicine/scales/PANSS">http://egret.psychol.cam.ac.uk/medicine/scales/PANSS</a>   |
| Include in model - Yes/No with reason | No – included in the Van Rijn review as part of a pooled analysis where results were not in a useable format and could not be converted.  |

| Name                                  | Hospital, Anxiety and Depression Scale (HADS)   |
|---------------------------------------|---|
| Category                              | Mental health   |
| Description                           | The Hospital Anxiety and Depression Scale (HADS) is a valid and reliable self-rating scale that measures anxiety and depression in both hospital and community settings. HADS gives clinically meaningful results as a psychological screening tool and can assess the symptom severity and caseness of anxiety disorders and depression in patients with illness and the general population. |
| Scale                                 | 0-21  |
| Sources                               | <a href="https://www.gi-assessment.co.uk/products/hospital-anxiety-and-depression-scale-hads/">https://www.gi-assessment.co.uk/products/hospital-anxiety-and-depression-scale-hads/</a>   |
| References                            | Van Rijn, 2015  |
| Quality                               | Strong (++)   |
| Reason for quality rating             | Validated questionnaire; used widely in the UK  |
| Questionnaires                        | <a href="http://www.scalesandmeasures.net/files/files/HADS.pdf">http://www.scalesandmeasures.net/files/files/HADS.pdf</a>   |
| Include in model - Yes/No with reason | No – included in the Van Rijn review as part of a pooled analysis where results were not in a useable format and could not be converted.  |

|                                       |   |
|---------------------------------------|---|
| <b>Name</b>                           | <b>Brief Syndrome Inventory (BSI-18)</b>  |
| Category                              | Mental health   |
| Description                           | <p>BSI or Brief Symptom Inventory is an instrument that evaluates psychological distress and psychiatric disorders in people. BSI collects data reported by patients for the evaluation. The test can be used for areas such as patient progress, treatment measurements, and psychological assessment.</p> <p>The Brief Symptom Inventory-18 (BSI-18; Derogatis, 2001) is a widely used self-report questionnaire that measures general psychological distress. It is the briefest and latest version in a series of instruments designed by Derogatis (Derogatis, 1983; Derogatis &amp; Melisaratos, 1983; Derogatis, 2001) to measure general distress. The questionnaire consists of 18 descriptions of physical and emotional complaints; respondents are asked to indicate on a scale from nought (not at all) through four (very much) to what extent they are troubled by the complaints.</p> |
| Scale                                 | 4-72  |
| Sources                               | <p><a href="http://www.statisticssolutions.com/brief-symptom-inventory-bsi/">http://www.statisticssolutions.com/brief-symptom-inventory-bsi/</a></p> <p><a href="#">An Evaluation of the Brief Symptom Inventory–18 Using Item ...</a></p>  |
| References                            | Jacobs, 2016  |
| Quality                               | Moderate (+)  |
| Reason for quality rating             | Validated questionnaire; only one study uses this scale   |
| Questionnaires                        | <a href="#">Unable to find</a>  |
| Include in model - Yes/No with reason | No – Population (low-income urban women with at least one child) is beyond scope of work and this is the only study that used the BSI-18.   |

|                                       |   |
|---------------------------------------|---|
| <b>Name</b>                           | <b>Comprehensive Quality of Life Scale (ComQoL)</b>   |
| Category                              | Quality of Life   |
| Description                           |   |
| Scale                                 | Seven domains   |
| Sources                               | <a href="http://www.acqol.com.au/instruments/comqol-scale/">http://www.acqol.com.au/instruments/comqol-scale/</a>                           |
| References                            | Beyer, 2010   |
| Quality                               | Moderate (+)  |
| Reason for quality rating             | Only one study uses this scale  |
| Questionnaires                        | <a href="http://www.acqol.com.au/instruments/comqol-scale/comqol-s5.pdf">http://www.acqol.com.au/instruments/comqol-scale/comqol-s5.pdf</a> |
| Other information                     |   |
| Include in model - Yes/No with reason | No – included in only one study that has a weak study quality and has, therefore, been excluded.  |

| Name                                  | Latent and Manifest Benefits Scale (LAMB)  |
|---------------------------------------|--|
| Category                              | Financial strain and the latent benefits of employment   |
| Description                           | The latent and manifest benefits [LAMB] of work scale by Muller et al. (2005) served as measure of the latent and manifest benefits. This scale consists of six subscales, each containing six items, using a seven-point Likert scale format, with scale endpoints expressing high or low access to the benefit |
| Scale                                 | Six items  |
| Sources                               | <a href="https://www.sheffield.ac.uk/polopoly_fs/1.366424!/file/Selenko_Batinic_Paul_2011_Does_latent_deprivation.pdf">https://www.sheffield.ac.uk/polopoly_fs/1.366424!/file/Selenko_Batinic_Paul_2011_Does_latent_deprivation.pdf</a>  |
| References                            | Hoare, 2010  |
| Quality                               | Moderate (+)   |
| Reason for quality rating             | Only one study uses this scale   |
| Include in model - Yes/No with reason | No – included in only one study that has a weak study quality and has, therefore, been excluded.   |

| Name                                  | Positive and Negative Affect Schedule (PANAS)   |
|---------------------------------------|---|
| Category                              | Subjective wellbeing  |
| Description                           | <p>The Positive and Negative Affect Schedule (PANAS) comprises two mood scales, one that measures positive affect and the other which measures negative affect. Used as a psychometric scale, the PANAS can show relations between positive and negative affect with personality stats and traits. Ten descriptors are used for each PA scale and NA to define their meanings.</p> <p>Participants in the PANAS are required to respond to a 20-item test using a five-point scale that ranges from very slightly or not at all (1) to extremely (5).</p> |
| Scale                                 | "Positive scale --> 10 items; min. score = 10, max score = 50<br>Negative scale --> 10 items; min score = 10, max score = 50  |
| Sources                               | <a href="https://booksite.elsevier.com/9780123745170/Chapter%203/Chapter_3_Worksheet_3.1.pdf">https://booksite.elsevier.com/9780123745170/Chapter%203/Chapter_3_Worksheet_3.1.pdf</a>   |
| References                            | Ferreira, 2015  |
| Quality                               | Moderate (+)  |
| Reason for quality rating             | Only one study uses this scale  |
| Questionnaires                        | <a href="https://booksite.elsevier.com/9780123745170/Chapter%203/Chapter_3_Worksheet_3.1.pdf">https://booksite.elsevier.com/9780123745170/Chapter%203/Chapter_3_Worksheet_3.1.pdf</a>   |
| Include in model - Yes/No with reason | No – scale does not measure a validated clinical definition but might be included as extra narrative of the effects of a transition.  |

|                                       |   |
|---------------------------------------|---|
| <b>Name</b>                           | <b>Global Assessment Scale (GAS)</b>  |
| Category                              | Overall functioning   |
| Description                           | The Global Assessment Scale (GAS) is a rating scale for evaluating the overall functioning of a subject during a specified time period on a continuum from psychological or psychiatric sickness to health. |
| Scale                                 |   |
| Sources                               | <a href="https://www.ncbi.nlm.nih.gov/pubmed/938196">https://www.ncbi.nlm.nih.gov/pubmed/938196</a>   |
| References                            | Crowther, 2010  |
| Quality                               | Weak (-)  |
| Reason for quality rating             | Only one study uses this scale; not much information online   |
| Questionnaires                        |   |
| Other information                     |   |
| Include in model - Yes/No with reason | No – used in only one study and not much information available.   |

|                                       |  |
|---------------------------------------|--|
| <b>Name</b>                           | <b>Global Assessment of Functioning (GAF)</b>  |
| Category                              | Overall functioning  |
| Description                           | An individual's overall functioning level. Impairments in psychological, social and occupational/school functioning are considered, but those related to physical or environmental limitations are not. The scale ranges from 0 (inadequate information) to 100 (superior functioning). Starting at either the top or the bottom of the scale, go up/down the list until the most accurate description of functioning for the individual is reached. |
| Scale                                 | 0-100<br><=50 Serious symptoms (the closer to 0, the closer to the likelihood of self harming).  |
| Sources                               | <a href="https://www.omh.ny.gov/omhweb/Childservice/mrt/global_assessment_functioning.pdf">https://www.omh.ny.gov/omhweb/Childservice/mrt/global_assessment_functioning.pdf</a>  |
| References                            | Van Rijn, 2015   |
| Quality                               | Moderate (+)   |
| Reason for quality rating             | Comprehensive questionnaire; only one study uses this scale  |
| Questionnaires                        | <a href="https://www.omh.ny.gov/omhweb/Childservice/mrt/global_assessment_functioning.pdf">https://www.omh.ny.gov/omhweb/Childservice/mrt/global_assessment_functioning.pdf</a>  |
| Other information                     | <a href="http://www.albany.edu/counseling_center/docs/GAF.pdf">http://www.albany.edu/counseling_center/docs/GAF.pdf</a>  |
| Include in model - Yes/No with reason | No – included in the Van Rijn review as part of a pooled analysis. In this case, it is pooled with the SF-36; given SF-36 reliability, this will be the preferred measure in the tool.   |

|  |   |
|--|---|
| <b>Name</b>                                  | <b>Mental Health Inventory (MHI-5)</b>  |
| <b>Category</b>                              | Mental health   |
| <b>Description</b>                           | <p>The Mental Health Inventory (MHI) is a method for evaluating mental health issues such as anxiety, depression, behavioural control, positive effect, and general distress. This instrument helps in the measure of overall emotional functioning.</p> <p>The Mental Health Inventory includes 38 items in which the respondent uses a six-point Likert-style response and it can generally be done without help. The test takes approximately five-10 minutes to administer. There is an abbreviated version of the MHI, which has only five items. The MHI-5 has a score of 0 to 100, where a score of 100 represents optimal mental health.</p> <p>The Medical Outcomes Study 36-Item Short Form health survey (SF-36, which includes the MHI-5)</p> <p>The MHI-5 used in the SF-20 has been retained without modification in the SF-36 – the MHI-5 is the Mental Health component of the SF-36.</p> <p>--&gt; <a href="http://www.espalibrary.eu/media/filer_public/f5/31/f531ce0b-0584-476b-b46a-abe3df14d0ec/ware-mc1992.pdf">http://www.espalibrary.eu/media/filer_public/f5/31/f531ce0b-0584-476b-b46a-abe3df14d0ec/ware-mc1992.pdf</a></p> |
| <b>Scale</b>                                 | 0-100   |
| <b>Sources</b>                               | <a href="http://www.statisticssolutions.com/mental-health-inventory-mhi/">http://www.statisticssolutions.com/mental-health-inventory-mhi/</a>   |
| <b>References</b>                            | Butterworth, 2010   |
| <b>Quality</b>                               | Strong (++)   |
| <b>Reason for quality rating</b>             | Validated questionnaire; sub-questionnaire of the SF-36   |
| <b>Questionnaires</b>                        | <a href="http://calmhsa.org/wp-content/uploads/2016/02/MHI-5English.pdf">http://calmhsa.org/wp-content/uploads/2016/02/MHI-5English.pdf</a>   |
| <b>Other information</b>                     | <p>Paper discussing threshold levels for GHQ-12 and MHI-5:</p> <p><a href="http://hqlo.biomedcentral.com/articles/10.1186/1477-7525-2-23">http://hqlo.biomedcentral.com/articles/10.1186/1477-7525-2-23</a></p>   |
| <b>Include in model - Yes/No with reason</b> | Yes – validated questionnaire, used in a number of included studies. It is also the mental health component of the SF-36.   |

| Name                                  | Satisfaction With Life Scale  |
|---------------------------------------|---|
| Category                              | Quality of Life   |
| Description                           | The SWLS is a short five-item instrument designed to measure global cognitive judgments of satisfaction with one's life. The scale usually requires only about one minute of a respondent's time.         |
| Scale                                 | 5-35  |
| Sources                               | <a href="https://internal.psychology.illinois.edu/~ediener/SWLS.html">https://internal.psychology.illinois.edu/~ediener/SWLS.html</a>   |
| References                            | Ferreira, 2015  |
| Quality                               | Moderate (+)  |
| Reason for quality rating             | Only one study uses this scale  |
| Questionnaires                        | <a href="https://internal.psychology.illinois.edu/~ediener/Documents/Understanding%20SWLS%20Scores.pdf">https://internal.psychology.illinois.edu/~ediener/Documents/Understanding%20SWLS%20Scores.pdf</a> |
| Other information                     |   |
| Include in model - Yes/No with reason | No – scale does not measure a validated clinical definition but might be included as extra narrative of the effects of a transition.  |

## J. Conversion of SF-36 Scores to QALYs

An algorithm derived by Ara & Brazier (2009) was used to convert aggregated SF-36 scores into QALYs. The algorithm, below, predicts a cohort level preference-based SF-6D score using the eight mean health dimension scores from the SF-36, when patient level data is not available. This score, once calculated, is subsequently used as a QALY value.

$$SF - 6D = 34.31814 + 0.0994 * PF + 0.0215 * RP + 0.1083 * BP + 0.0140 * GH + 0.0479 * VT + 0.1001 * SF + 0.0394 * RE + 0.1269 * MH)/100$$

Table 12 shows the change in SF-36 score from baseline to follow-up of the population of interest in the Schuring (2010) paper.

**Table 12: Change in SF-36 scores for included studies**

| SF-36 domain            | SF-36 incremental gain |                                     |                                    |  |
|-------------------------|------------------------|-------------------------------------|------------------------------------|--|
|                         | Schuring et al. (2010) | Butterworth (2011) Good Quality Job | Butterworth (2011) Bad Quality Job | Galić & Šverko, (2008) (average scores provided) |
| Physical Function_(PF)  | 9.2                    |                                     |                                    |  |
| Role – Physical_(RP)    | 20                     |                                     |                                    |  |
| Bodily Pain_(BP)        | 11.3                   |                                     |                                    |  |
| General Health_(GH)     | 7                      |                                     |                                    |  |
| Vitality_(VT)           | 7.8                    |                                     |                                    |  |
| Social Functioning_(SF) | 14.2                   |                                     |                                    | 3.55   |
| Role – Emotional_(RE)   | 22.7                   |                                     |                                    | 3.55   |
| Mental Health_(MH)      | 11                     | 3.3                                 | -5.6                               | 3.55   |

Using the numbers as provided, the SF-36 score, and thus, the QALY gains, as stipulated from the results of the Schuring (2010) paper is as follows:

$$SF - 6D = 34.31814 + (0.0994 * 9.2) + (0.0215 * 20) + (0.1083 * 11.3) + (0.0140 * 7) + (0.0479 * 7.8) + (0.1001 * 14.2) + (0.0394 * 22.7) + (0.1269 * 11)/100$$

$$SF - 6D = 0.0675159$$

This equates to a QALY gain of 0.0675 per person, by moving from unemployment to employment. This figure has been used in the tool to subsequently calculate the monetary value of a QALY gain per person of £4,050 (calculated using the value of £60,000 per QALY).

Butterworth (2010) only reported changes in the mental health component of the SF-36, the MHI-5. This, therefore, equates to a 0.00042 QALY gain per person, by moving from unemployment to employment. Galić & Šverko, (2008) reported on the three mental health components of the SF-36, which equated to a QALY gain of 0.0095 per person QALY gain.

### K. Mental health threshold methodology

Using the weighted average of three UK nationwide SF-36 scores provided in Bowling et. al. (1999) as baseline (Table 13), we use the methodology provided by Ritvo et. al. (2007) to create the Mental Component Summary (MCS) for our target population. The MCS formula is as follows:

$$MCS = 50 + (AGG_{MENT} * 10)$$

**Table 13: UK SF-36 Scores**

|                         | Health Survey for England (HSE) 1996 (16+) |                  | Oxford (Central England) Healthy Life Survey 1991-1992 (18-64) |                  | British ONS Survey 1992 (16+) |                  | Weighted average mean SF-36 scores |
|-------------------------|--|------------------|--|------------------|-------------------------------|------------------|------------------------------------|
|                         | Mean baseline SF-36 scores                 | Number of people | Mean baseline SF-36 scores                                     | Number of people | Mean baseline SF-36 scores    | Number of people |                                    |
| Physical Function_(PF)  | 81   | 16054            | 88   | 8801             | 90                            | 2025             | 84                                 |
| Role – Physical_RP      | 80   | 15155            | 86   | 9058             | 84                            | 2018             | 82                                 |
| Bodily Pain_(BP)        | 77   | 16160            | 82   | 10105            | 83                            | 2022             | 79                                 |
| General Health_GH       | 69   | 16016            | 74   | 8990             | 74                            | 2017             | 71                                 |
| Vitality_(VT)           | 63   | 16139            | 61   | 8998             | 65                            | 2018             | 62                                 |
| Social Functioning_(SF) | 85   | 16167            | 88   | 9124             | 89                            | 2020             | 86                                 |
| Role emotional_RE       | 84   | 16026            | 83   | 8067             | 88                            | 1919             | 84                                 |
| Mental Health_(MH)      | 75   | 16127            | 74   | 8930             | 77                            | 2019             | 75                                 |

The methodology transforms the raw baseline SF-36 scores to a 0-100 scale that is normally distributed by standardising the eight SF-36 domains, weighting and aggregating these scales and transforming them into the MCS. In order to calculate the MCS, an aggregate mental health scale score (AGG\_MENT) is created using the following formula:

$$AGG_{MENT} = (PF_z * -0.22999) + (RP_z * -0.12329) + (BP_z * -0.09731) + (GH_z * -0.01571) + (VT_z * 0.23534) + (SF_z * 0.26876) + (RE_z * 0.43407) + (MH_z * 0.48581)$$

Where the subscript z represents the z-score for each dimension. The formulae and results for calculations of the z-score are presented in Table 14 and plugged into the formula above to calculate an aggregate mental health scale score.

**Table 14: SF-36 z-score calculations**

|      | Formula                  | Z score (Schuring 2010) |
|------|--------------------------|-------------------------|
| PF_Z | (PF-84.52404)/22.89490   | 0.382038524             |
| RP_Z | (RP - 81.19907)/33.79729 | 0.625105747             |
| BP_Z | (BP - 75.49196)/23.55879 | 0.628586161             |
| GH_Z | (GH - 72.21316)/20.16964 | 0.280475937             |
| VT_Z | (VT - 61.05453)/20.86942 | 0.442859728             |
| SF_Z | (SF - 83.59753)/22.37642 | 0.755284135             |
| RE_Z | (RE - 81.29467)/33.02717 | 0.311942786             |
| MH_Z | (MH - 74.84212)/18.01189 | 0.604123929             |

The MCS for Schuring (2010) is thus:

$$\text{MCS} = 50 + (0.5055989 * 10) = 50.055989$$

As the MCS is normally distributed, a threshold approach was used to calculate the number of people who, from moving from unemployment to employment, no longer have a common mental health disorder. The threshold mean value of 44.8 provided in Kelly et. al. (2008) was used; this rate is defined as the misclassification rate as it attempts to provide the most appropriate cut-point that avoids classifying someone who does not have CMD(s) with CMD(s) when they do not and vice versa.

The MCS is then used to calculate the baseline proportion of people with CMDs at baseline, and the change in the proportion of people with CMDs using findings from the shortlisted reports.

**Table 15 - Mean MCS scores and change in proportion with CMD(s) compared with baseline**

|  | Mean MSC score | Standard Deviation | % with CMD | Change in proportion compared to baseline |
|--|----------------|--------------------|------------|---|
| <b>Baseline</b>                            | 50.47          | 10                 | 28.5%      |   |
| <b>Schuring 2010</b>                       | 55.06          | 10                 | 15%        | -13.3%                                    |
| <b>Butterworth 2011 (good quality job)</b> | 51.36          | 10                 | 25.6%      | -2.9%                                     |
| <b>Butterworth 2011 (poor quality job)</b> | 48.95          | 10                 | 34%        | 5.3%                                      |
| <b>Galić 2008 UUE → T3 - T2</b>            | 53.61          | 10                 | 19%        | -9.6%                                     |
| <b>Galić 2008 UUE → T3 - T1</b>            | 51.75          | 10                 | 24%        | -4.2%                                     |
| <b>Galić 2008 UEE → T3 - T1</b>            | 51.18          | 10                 | 26%        | -2.4%                                     |
| <b>Galić 2008 UEE → T2 - T1</b>            | 51.62          | 10                 | 25%        | -3.8%                                     |

The change in the proportion compared with baseline allows calculation of the number of people crossing the threshold relative to the proportion below the threshold. The results are shown in the table below (note that the results for Galić 2008 are averaged across all four time periods).

**Table 16: Proportion of mental health population improving**

|  | Proportion of mental health population improving |
|--|--|
| Schuring (2010)                          | 46.56%   |
| Butterworth (2011)<br>– good quality job | 10.32%   |
| Butterworth (2011)<br>– poor quality job | -18.70%  |
| Galić (2008)                             | 17.50%   |

The above proportions are multiplied by the proportion of people returning to work and the proportion of people with a CMD in the selected population of interest. Using the proportion of the mental health population improving, and using results as presented in Schuring (2010), in a cohort of the population of interest who gain sustainable employment = 314, (as is the case if the population of interest is Bedford, and 5% of this population are assumed to return to work), the number of people improving is:

$$\begin{aligned} &\text{Proportion of mental health population improving} \\ &= \text{Proportion of mental health population improving} \\ &\quad * \text{Number of people returning to work} * \text{Baseline proportion of people with CMD(s)} \end{aligned}$$

$$\text{Proportion of mental health population improving} = 46.56\% * 314 * 18.9\% = 28$$

The results therefore show that, in a cohort of 314 people returning to work, 28 of these people will see an improvement in the state of their mental health.