

A systematic review of evidence on the employment impacts of changes to disability benefit, eligibility, assessment criteria, and income replacement levels.

PROTOCOL

July 2018

Philip McHale, Andy Pennington, Ben Barr

Department of Public Health and Policy, University of Liverpool

Citation: McHale P, Pennington A, Barr B (2018) **A systematic review of evidence on the employment impacts of changes to disability benefit, eligibility, assessment criteria, and income replacement levels.** – PROTOCOL. Liverpool: University of Liverpool.

Contact:

Philip McHale

Department of Public Health and Policy

University of Liverpool

Whelan Building

Brownlow Hill

Liverpool

L69 3GB

Email: hlpmchal@liverpool.ac.uk

Contents

1	Background.....	2
1.1	Purpose of the systematic review, and place within the THRIVE programme.....	2
1.2	Aim of the review.....	2
1.3	Review Questions.....	2
1.4	Operational definitions.....	3
1.5	Outputs from the review.....	3
2	Methods.....	5
2.1	Search strategy.....	5
2.2	Stage 1. Broad Search – relevant to overarching THRIVE programme research questions.....	5
2.3	Stage 2 study identification and inclusion.....	6
3	Data extraction.....	7
4	Validity assessment.....	8
5	Data synthesis.....	8
6	References.....	10
7	Appendices.....	12

1 Background

There is much interest in the influence that eligibility criteria and level of income replacement benefits for disabled people, have on the rate of employment in this population. Large increases in rates of claimants for such benefits across OECD countries (OECD, 2003), as well as a persistently high disability employment gap, have made this a pressing issue. One hypothesis is that relaxed eligibility of such disability benefits, combined with increased level compared to wages and other income replacement benefits, have an overall negative impact on employment rates, and are the driving force behind increasing numbers of claimants. A systematic review by Barr *et al.* (2010) argued that the evidence base for this hypothesis was focussed on the US context, and questioned the generalisability of findings. The review included studies from five OECD countries (UK, Canada, Norway, Sweden and Denmark) and found that the evidence base suggested no association between relaxed eligibility and employment rates, while increased level was associated with a small, negative effect on employment rates. Further research has investigated this policy question (for example, Marie and Vall Castello, 2012; Barr *et al.*, 2016), however other work has raised concerns with the lack of power and high potential for bias in much of the econometric literature (Ioannidis *et al.*, 2017). To our knowledge there has been no previous systematic review of the evidence across all OECD countries. Establishing the extent to which changes to disability benefits effect employment is crucial to enable policy makers to assess the benefits and harms of such policies.

1.1 Purpose of the systematic review, and place within the THRIVE programme

The purpose of this systematic review is to consider evidence on how increasing or decreasing the income replacement level of disability benefits, and/or making eligibility more or less strict, impacts on employment rates. The review will be limited to studies which analyse a change in either benefit eligibility or income replacement level, and a before and after assessment of change in employment outcomes.

The review is part of the broader Tackling Health Inequalities and Extending Working Lives (THRIVE) programme, a programme of international research that aims to advance understanding of the differential impacts of health inequalities on opportunities to work later in life and of strategies and policies for extending working life that take these health inequalities into consideration (<http://gtr.ukri.org/projects?ref=ES%2FN019261%2F1>).

1.2 Aim of the review

The aim of this systematic review is to identify, critically appraise and synthesize the existing evidence about the relationship between changes in the income replacement level and/or eligibility/assessment criteria of disability benefits, and employment rates in older working age individuals (50-65 years old) in OECD countries.

1.3 Review Questions

The systematic review will address the following questions:

RQ1. How do changes to the eligibility, assessment process and income replacement level of disability benefits effect employment rates of older working-age individuals in OECD countries?

RQ2. Do effects differ by age, sex, socioeconomic status (based on income, educational level, or occupation), or country welfare state model.

1.4 Operational definitions.

The review used the following working definitions of terms.

Macro-level policies and interventions.

Any service or financial, regulatory/legal or educational tool implemented at any level of government (i.e. local, provincial, or national level) applied at the population level (rather than within specific employers, or organisations). We also include policies implemented by government-related agencies such as the Workers Compensation Boards or similar organisations in the US, Canada and Australia.

Disability benefits

Cash benefits paid by government-related agencies to people who are assessed as unable to work or have limited capacity to work due to disability. These benefits are paid to cover the loss of income that results from limited capacity to work. In our working definition we do not include benefits paid to cover the additional costs associated with having a disability.

Employment

Persons who during a specified period were engaging in some paid work either as an employee or self-employed (Based on ILO, 2016).

Income replacement level of disability benefits

The cash level of benefits paid to recipients. This may be measured in absolute terms or relative to average wages – i.e. as a replacement rate (based on UNRISD et al, 2015).

Disability

A physical or mental impairment that has a ‘substantial’ and ‘long-term’ negative effect on ability to do activities of daily living including employment.

Eligibility criteria

Criteria within a disability benefit scheme determining who is or is not entitled to receive the benefit. An individual who meets a set of qualifiers for the benefit will be entitled to receive it.

Disability assessment

A test or assessment used to determine whether an applicant’s disability or impairment limits their work capacity sufficiently to warrant receipt of disability benefits. Some assessment procedures incorporate education, employment experience and the state of the local labour market, as well as measures of impairment in determining eligibility.

1.5 Outputs from the review

- Registration with PROSPERO, an International Register of Systematic Review protocols (www.crd.york.ac.uk/prospero).

- A systematic review report (published on the University of Liverpool repository <https://livrepository.liverpool.ac.uk>).
- A publication in a peer reviewed academic journal.

2 Methods

The review will use standard systematic review methodology, as described in the Centre for Reviews and Dissemination’s guide to Systematic Reviews (CRD, 2009), and will be reported following PRISMA and PRISMA-Equity guidelines (Moher et al., 2009; Welch et al., 2012, 2016).

2.1 Search strategy

A comprehensive search strategy was developed by experienced systematic reviewers and information scientists. The aim of the search is to identify all evidence on how changes in the income replacement level or eligibility/assessment criteria of income replacement disability benefits impact on employment rates in older working age individuals in OECD countries.

The search strategy was designed to be implemented across two stages.

2.2 Stage 1. Broad Search – relevant to overarching THRIVE programme research questions.

During stage 1 (now complete), the electronic databases MEDLINE, EMBASE, PsycInfo, Econlit were searched to identify papers relevant to the broader THRIVE programme’s research questions (Box 1). An example of the MEDLINE search strategy is shown in Appendix 1.

The stage 1 search provided the basis for a number of systematic reviews focused on specific policy areas across the broader THRIVE work programme.

Box 1. Research questions from overarching THRIVE programme

- How do the effects of policies that aim to promote employment at older ages differ by socioeconomic group, gender and for different health conditions?
- Which policies are most likely to be effective at equitably extending the working lives people with disabilities?

The stage 1 search strategy reflected the PICOS criteria for the overarching THRIVE programme reviews outlined in Table 1.

Table 1. PICOS criteria for broader THRIVE programme search.

	Include	Exclude
Population / setting:	Includes population over 50 years with a disability from OECD countries (this includes studies that look at employment outcomes across the whole population).	Studies that only included persons younger than 50 years of age. All other countries.
Intervention:	Macro-level policies and interventions.	Meso and micro level only policies.
Comparison:	N/A	N/A
Outcome:	Change or variation in employment status or employment rates.	Studies not considering changes in employment status or rates.

Study design:	Any quantitative study design.	Qualitative studies.
---------------	--------------------------------	----------------------

Papers were initially screened against the criteria in Table 1. Based on an initial review of titles and abstracts, they were then mapped/coded based on their policy focus as follows:

Policy theme

- 1: Disability
- 2: Retirement
- 3: Both

Policy sub-themes under the disability theme:

- A: Return-to-work
- B: Wage subsidies
- C: Change or variation in disability benefit eligibility or assessment procedures
- D: Change or variation in disability benefit income replacement level.

Policy sub-themes under the retirement theme:

- E: Retirement age
- F: Pension generosity

Papers could be coded with having multiple policy sub-themes.

Papers meeting the stage 1 inclusion criteria, and coded as relevant to this systematic review (i.e. theme 1; and policy sub-themes C and D), were then allocated for further screening within stage 2.

2.3 Stage 2 study identification and inclusion

The papers from stage 1 will be assessed for inclusion against the criteria for this systematic review, shown in Table 2.

Studies will be screened over two steps. First, a random sample of 20% the same title and abstracts will be independently screened by two reviewers. This will be followed by a screening ‘calibration’ exercise to ascertain levels of agreement and ensure consistency of approaches by reviewers. Once a level of agreement/consistency is reached (on >90% of include/exclude decisions), the remaining titles and abstracts will be screened by one reviewer. Second, full-text copies of all papers included during title and abstract screening will be obtained and independently screened for inclusion by two reviewers. During screening, any queries and disagreements will be resolved by discussion, or by recourse to a third reviewer.

Table 2: Inclusion/Exclusion criteria for systematic review

<i>PICOS inclusion / exclusion criteria</i>		
	Include	Exclude
Population / settings	Older working age population (aged 50-65 years), in OECD countries.	Persons younger than 50 years of age, or older than 65 years of age. All other countries.
Intervention / exposure	Changes in the income replacement level, eligibility and/or assessment approaches of disability benefits.	Changes to other forms of disability benefits. Changes to other forms of income replacement benefits. All other types of benefits.
Comparison	Either comparisons with the same	Studies that only included a

	population prior to the policy introduction (e.g. as in before and after and interrupted time series studies) or comparison over time between populations experiencing the policy change and those who have not.	population were there was variation in exposure.
Outcomes	Effect on the probability of being in employment. Time off work or not in employment.	Volunteer work. Length of time on disability /sickness benefits.
Study designs	Studies that include data pre and post policy exposures including: Controlled intervention studies Before and after studies Interrupted Time series studies Difference in differences Panel regression studies.	Studies that do not include data pre and post policy exposures.
<i>Publication characteristics inclusion / exclusion criteria</i>		
	Include	Exclude
Publication types	Primary studies from peer-reviewed literature. Papers published or in-press. Working papers.	Any work that is not a primary research study, including letters, editorials, commentaries, conference proceedings, books and book chapters, meeting abstracts, lectures and addresses. Previous reviews and meta-analyses are not eligible, but relevant reviews will be used to identify relevant primary studies.
Year of publication	1990-2018	Prior to 1990
Language	English language	Non-English language

The reference lists of all included studies will then be hand-searched to identify further studies of interest (i.e. ‘backward citation searching’ or ‘snowballing’), and information on unpublished and in-progress research will be requested from key researchers in the field. Forward citation searches of included studies will also be conducted via Web of Science.

3 Data extraction

Data from each included study will be extracted into pre-designed and piloted forms. Forms will be completed by one reviewer and checked for accuracy by another. Periodically during the process, a random selection will be considered independently by 2 reviewers (that is, double assessed) for at least 20% of the studies. Data to be extracted are shown in Box 2.

Box 2: Data to be extracted

1. Study population, age, sex, country, other selection criteria
2. Sample size
3. Year study started
4. Duration of study
5. Study design
6. Method of data collection (survey, registry, administrative data)
7. Description of intervention/ change or difference in benefits policy – including direction of change in benefit level
8. Outcomes measured and definitions
9. Missing data/ response rates
10. Type of analysis
11. Other variables controlled for in multivariate analysis
12. Results
13. Limitations identified by authors
14. Authors conclusions
15. Journal Type (Economic or Health)

4 Validity assessment

All papers identified for inclusion in the review will be assessed using a validity assessment framework from the Barr et al, 2010 review (see Appendix 2). Validity assessments will be completed by one reviewer and checked for accuracy by another. Periodically during the process, a random selection will be considered independently by 2 reviewers (that is, double assessed) for at least 20% of the studies. Any queries and disagreements will be resolved by discussion, or by recourse to a third reviewer.

5 Data synthesis

Extracted data will be collated in a structured database. Both harvest plots (Ogilvie et al. 2008) and meta-analyses will be used to synthesise the data. Harvest plots will be created to display and summarise the results of all of the included studies and explore variation within subgroups of study based on type of policy intervention, age, sex, socioeconomic status (based on income, educational level, or occupation), and country welfare state model. In a harvest plot each reported effect size is represented by a single bar. The height of the bars are used to indicate the validity score of the studies from which the associations arose, so that the strength of the evidence can be visualised, and greater weight given to the most methodologically robust and reliable studies.

Meta-analyses will be conducted in R using an inverse variance random-effects model on combined results. Where necessary, standard methods will be used to calculate elasticities (for studies investigating changes in income replacement level) and risk differences (for changes in eligibility/ assessment approaches) with confidence intervals (Higgins et al., 2011). Where studies present multiple models the model including the highest number of control variables will be used.

Statistical heterogeneity will be assessed by applying the I^2 statistic with values of 30 to 60%, 50 to 90% and 75 to 100% used to denote moderate, substantial and considerable levels of heterogeneity, respectively (Higgins et al., 2011). Random-effects meta-regression will be used to investigate differences in effect between socioeconomic, age and gender groups and

between studies based on their validity assessment scores. Publication bias will be assessed using funnel plots.

The estimates of effect size derived from the metanalysis will be used to estimate the statistical power of the original studies using the methods outlined by Ioannidis et al. (2017).

6 References

- Barr B, Clayton S, Whitehead M, et al. (2010) To what extent have relaxed eligibility requirements and increased level of disability benefits acted as disincentives for employment? A systematic review of evidence from countries with well-developed welfare systems. *J Epidemiol Community Health*; 64: 1106–14.
- Barr B, Taylor-Robinson D, Stuckler D, Loopstra R, Reeves A, Wickham S, et al. (2016) Fit-for-Work or Fit-for-Unemployment? Does the Reassessment of Disability Benefit Claimants Using a Tougher Work Capability Assessment Help People into Work? *J Epidemiol Community Health*; 70(5): 452–58. <https://doi.org/10.1136/jech-2015-206333>.
- CRD (2009) *Systematic Reviews - CRD's guidance for undertaking systematic reviews in health care*. York: Centre for Reviews and Dissemination, University of York.
- Equalities Act 2010, s 6(1)(a-b).
www.legislation.gov.uk/ukpga/2010/15/pdfs/ukpga_20100015_en.pdf.
- Higgins JPT, Green S, eds. (2011) *Cochrane Handbook for Systematic Reviews of Interventions* Version. 5.1.0. London: The Cochrane Collaboration.
www.handbook.cochrane.org.
- ILO (2016) *Employment (definition)*. Geneva: International Labour Organization.
www.insee.fr/en/metadonnees/definition/c1159.
- Ioannidis J, Stanley T, Doucouliagos H (2017) The Power of Bias in Economics Research. *The Economic Journal*; 127(605): F236–65. <https://doi.org/10.1111/ecoj.12461>.
- Marie O, Vall Castello J (2012) Measuring the (income) effect of disability insurance adequacy on labour market participation. *J Public Econ*; 96: 198–210.
- Mays N, Pope C, Popay J (2005) Systematically reviewing qualitative and quantitative evidence to inform management and policy-making in the health field. *J Health Serv Res Policy*; 10(1)S1:6-S1:20.
- Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009) Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med*; 6(7): e1000097. <https://doi.org/10.1371/journal.pmed.1000097>.
- OECD (2003) *Transforming Disability into Ability Policies to Promote Work and Income Security for Disabled People*. Paris: Organisation for Economic Co-operation and Development. www.oecd.org/els/emp/transformingdisabilityintoability.htm.
- Ogilvie D, Fayter D, Petticrew, Sowden A, Thomas S, Whitehead M, et al. (2008) The harvest plot: A method for synthesising evidence about the differential effects of interventions. *BMC Med Res Methodol*; 8(8). <https://doi.org/10.1186/1471-2288-8-8>.

Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, et al. (2006) Guidance on the Conduct of Narrative Synthesis in Systematic Reviews - A Product from the ESRC Methods Programme. Lancaster University: Lancaster. <https://tinyurl.com/oovelag>.

UNRISD, ILO, ECLAC, OHCHR, ESCAP, ESCWA et al. (2015) Level of Benefits. <http://socialprotection-humanrights.org/framework/principles/level-of-benefits>.

Welch V, Petticrew M, Petkovic J, Moher D, Waters E, White H, et al. (2016) Extending the PRISMA Statement to Equity-Focused Systematic Reviews (PRISMA-E 2012): Explanation and elaboration. *J Clin Epidemiol*; 70:68-89.

Welch V, Petticrew M, Tugwell P, Moher D, O'Neill J, Waters E, et al. (2012) PRISMA-Equity 2012 Extension: Reporting Guidelines for Systematic Reviews with a Focus on Health Equity. *PLoS Med* 9(10): e1001333. <https://tinyurl.com/mt998e4>.

7 Appendices

Appendix 1. Example of MEDLINE search strategy

Database: MEDLINE: Epub Ahead of Print, In-Process & Other Non-Indexed Citations, via Ovid MEDLINE® Daily and Ovid MEDLINE® 1946 to Present, searched 12th April 2017.

-
- 1 "Organisation for Economic Co-Operation and Development"/ (75)
 - 2 "organi?ation for economic co-operation and development".ti,ab. (645)
 - 3 "organi?ation for economic cooperation and development".ti,ab. (594)
 - 4 OECD.ti,ab. (3149)
 - 5 exp Australia/ (122518)
 - 6 Australia\$.ti,ab. (111765)
 - 7 Austria/ (17785)
 - 8 Austria\$.ti,ab. (14308)
 - 9 Belgium/ (16392)
 - 10 Belgium.ti,ab. (12895)
 - 11 Belgian?.ti,ab. (7427)
 - 12 exp Canada/ (142416)
 - 13 Canada.ti,ab. (67730)
 - 14 Canadian?.ti,ab. (50673)
 - 15 Chile/ (11422)
 - 16 Chile\$.ti,ab. (13740)
 - 17 Czech Republic/ (6155)
 - 18 Czech?.ti,ab. (9620)
 - 19 Denmark/ (42875)
 - 20 Denmark.ti,ab. (22492)
 - 21 (Danes or Danish).ti,ab. (23980)
 - 22 Estonia/ (2098)
 - 23 Estonia\$.ti,ab. (2707)
 - 24 exp Europe/ (1270372)
 - 25 Europe\$.ti,ab. (231006)
 - 26 Finland/ (31707)
 - 27 Finland.ti,ab. (21345)
 - 28 Finns.ti,ab. (1449)
 - 29 France/ (86824)
 - 30 France.ti,ab. (52433)
 - 31 French.ti,ab. (59949)
 - 32 exp Germany/ (143064)
 - 33 German\$.ti,ab. (116613)
 - 34 Greece/ (16308)
 - 35 Greece.ti,ab. (12128)
 - 36 Greek?.ti,ab. (11073)
 - 37 Hungary/ (17305)
 - 38 Hungar\$.ti,ab. (13862)
 - 39 Iceland/ (3956)
 - 40 Iceland\$.ti,ab. (5447)
 - 41 Ireland/ (15372)
 - 42 Northern Ireland/ (4531)
 - 43 Ireland.ti,ab. (16610)
 - 44 Irish.ti,ab. (7594)
 - 45 Israel/ (26345)
 - 46 Israel\$.ti,ab. (26906)

47 Italy/ (78352)
48 Italian?.ti,ab. (41879)
49 Italy.ti,ab. (51074)
50 Japan/ (115143)
51 Japan\$.ti,ab. (180597)
52 Korea/ (16506)
53 Korea\$.ti,ab. (55797)
54 Latvia/ (1119)
55 Latvia\$.ti,ab. (1338)
56 Luxembourg/ (637)
57 Luxembourg\$.ti,ab. (862)
58 Mexico/ (32626)
59 Mexic\$.ti,ab. (51949)
60 Netherlands/ (58567)
61 Netherlands.ti,ab. (39876)
62 Dutch.ti,ab. (30732)
63 New Zealand/ (34676)
64 New Zealand\$.ti,ab. (45041)
65 exp North America/ (1415726)
66 North America\$.ti,ab. (44204)
67 Norway/ (33970)
68 Norway.ti,ab. (25610)
69 Norwegian?.ti,ab. (16114)
70 Poland/ (45883)
71 Poland.ti,ab. (19811)
72 Polish.ti,ab. (14465)
73 Portugal/ (10054)
74 Portugal.ti,ab. (9125)
75 Portuguese.ti,ab. (10141)
76 Slovakia/ (2350)
77 Slovak\$.ti,ab. (4507)
78 Slovenia/ (2171)
79 Slovenia\$.ti,ab. (3230)
80 exp South America/ (127995)
81 Spain/ (64165)
82 Spain.ti,ab. (44310)
83 Spanish.ti,ab. (41461)
84 Sweden/ (65376)
85 Sweden.ti,ab. (38445)
86 Swede?.ti,ab. (38907)
87 Swedish.ti,ab. (32152)
88 "Scandinavian and Nordic Countries"/ (4714)
89 Switzerland/ (31977)
90 Switzerland.ti,ab. (19385)
91 Swiss.ti,ab. (27605)
92 Turkey/ (29272)
93 Turkey.ti,ab. (29572)
94 Turk?.ti,ab. (2228)
95 exp United Kingdom/ (342668)
96 United Kingdom.ti,ab. (31705)
97 England/ (83223)
98 England.ti,ab. (41925)
99 English.ti,ab. (138452)
100 Scotland/ (24079)
101 Scotland.ti,ab. (14596)

102 Scottish.ti,ab. (8205)
 103 Wales/ (13313)
 104 Wales.ti,ab. (20375)
 105 Welsh.ti,ab. (1782)
 106 Great Britain.ti,ab. (6699)
 107 British.ti,ab. (43578)
 108 exp United States/ (1249418)
 109 United States.ti,ab. (191899)
 110 America\$.ti,ab. (333505)
 111 Developed Countries/ (19822)
 112 developed countr\$.ti,ab. (22653)
 113 or/1-112 (4094258)
 114 Fiscal Policy/ (1)
 115 Policy/ (1602)
 116 Social Control Policies/ (771)
 117 Organizational policy/ (13586)
 118 Government Programs/ (4407)
 119 Government Regulations/ (19762)
 120 Health Policy/ (58786)
 121 Employment, Supported/ (1062)
 122 Insurance, Disability/ (1366)
 123 Insurance, Health/ (32523)
 124 Legislation/ (1674)
 125 "Legislation as Topic"/ (15921)
 126 Pensions/ (3647)
 127 Retirement/ (8768)
 128 Sick Leave/ (4875)
 129 Social Security/ (7359)
 130 Workers' Compensation/ (7309)
 131 or/114-130 (168990)
 132 (change? or changing or eligib\$ or entitlement or generosity or increas\$ or introduction or reduc\$ or reform\$ or restrict\$).ti,ab. (8216224)
 133 131 and 132 (45424)
 134 (benefit? adj5 (change? or changing or eligibility or entitlement or generosity or increas\$ or introduction or reduc\$ or reform\$ or restrict\$)).ti,ab. (27240)
 135 (compensation adj5 (change? or changing or eligibility or entitlement or generosity or increas\$ or introduction or reduc\$ or reform\$ or restrict\$)).mp. (2945)
 136 (insurance adj5 (change? or changing or eligibility or entitlement or generosity or increas\$ or introduction or reduc\$ or reform\$ or restrict\$)).ti,ab. (4559)
 137 (legislat\$ adj5 (change? or changing or eligibility or entitlement or generosity or increas\$ or introduction or reduc\$ or reform\$ or restrict\$)).mp. (4695)
 138 (policies adj5 (change? or changing or eligibility or entitlement or generosity or increas\$ or introduction or reduc\$ or reform\$ or restrict\$)).mp. (7609)
 139 (policy adj5 (change? or changing or eligibility or entitlement or generosity or increas\$ or introduction or reduc\$ or reform\$ or restrict\$)).mp. (16816)
 140 (program\$ adj5 (change? or changing or eligibility or entitlement or generosity or increase? or introduction or reduc\$ or reform\$ or restrict\$)).mp. (44811)
 141 (retirement? adj3 age? adj5 (change? or changing or eligibility or entitlement or generosity or increas\$ or introduction or reduc\$ or reform\$ or restrict\$)).ti,ab. (190)
 142 (retirement? adj3 benefit? adj5 (change? or changing or eligibility or entitlement or generosity or increas\$ or introduction or reduc\$ or reform\$ or restrict\$)).ti,ab. (37)
 143 (sickness absence? adj5 (change? or changing or eligibility or entitlement or generosity or increas\$ or introduction or reduc\$ or reform\$ or restrict\$)).ti,ab. (457)
 144 (sickness benefit? adj5 (change? or changing or eligibility or entitlement or generosity or increas\$ or introduction or reduc\$ or reform\$ or restrict\$)).ti,ab. (34)

145 (social insurance adj5 (change? or changing or eligibility or entitlement or generosity or
 increas\$ or introduction or reduc\$ or reform\$ or restrict\$)).ti,ab. (64)
 146 (wage subsid\$ adj5 (change? or changing or eligibility or entitlement or generosity or increas\$
 or introduction or reduc\$ or reform\$ or restrict\$)).ti,ab. (4)
 147 or/134-146 (104468)
 148 133 or 147 (139455)
 149 (disab\$ adj3 retirement?).ti,ab. (298)
 150 (early adj3 (retire? or retiring or retirement?)).ti,ab. (1147)
 151 (earning? adj5 (employment or job? or career? or occupation\$ or profession\$)).ti,ab. (498)
 152 Employment/ (42077)
 153 employment.ti,ab. (46706)
 154 "exit from work\$.ti,ab. (41)
 155 health status indicators/ (22417)
 156 health status/ (71404)
 157 health status disparities/ (11263)
 158 Income/ (25313)
 159 (labo? force adj3 participat\$.ti,ab. (887)
 160 (labo?r market adj3 participat\$.ti,ab. (200)
 161 long-term disab\$.ti,ab. (2192)
 162 longterm disab\$.ti,ab. (19)
 163 (outcome? adj3 (employment or job? or career? or occupation\$ or profession\$)).ti,ab. (3261)
 164 retirement/ (8768)
 165 Return to Work/ (1287)
 166 (return\$ adj3 work\$.ti,ab. (10406)
 167 RTW.ti,ab. (728)
 168 self-employ\$.ti,ab. (1289)
 169 short-term disab\$.ti,ab. (232)
 170 (stay\$ adj2 work\$.ti,ab. (310)
 171 unemployment/ (6133)
 172 (work\$ adj3 participation).ti,ab. (1866)
 173 Work/ (19508)
 174 or/149-173 (238077)
 175 113 and 148 and 174 (9088)
 176 limit 175 to yr="1990 -Current" (8140)

Appendix 2. Validity Assessment framework

Table A1. Validity assessment framework based on Barr et al., 2010

Criteria	Rationale	Score
Unit of analysis	There were three types of analysis units used in the studies, aggregate (ecological), individual or repeated measures on the same individuals (panel). Panel data was seen as being the most robust as it allows for unmeasured confounding factors to be accounted for where these do not vary within individuals over time. Ecological studies were seen at the least robust as ecological bias can occur where aggregate data are used to make inferences about individuals.[1]	3: Longitudinal (panel) data 2: Individual data (repeated cross section) 1: Ecological (aggregate data)
Comparison approach	Studies either investigate changes over time before and after an intervention or using a difference in differences approach. Studies that look at changes in the same group over time will overcome bias to a certain extent; however, the results will be at risk of being influenced by other secular trends. The more robust approach will be where a policy has changed over time for one group and this is compared with another group that is unaffected by the change (a difference in differences approach).	3: Difference in Differences 2: Interrupted time series 1: Other method
Sample selection	Studies either use: (1) nationally recognised surveys based on random sampling, (2) non-random but representative data, for example administrative data from a scheme with universal coverage, or (3) a non-random sample not representative of the rest of the population such as administrative data from a scheme without universal coverage.	3: Nationally recognised survey, based on random sampling 2: Non-random sample that is representative 1: Non-random sample that is not representative
Number of time points of data	A large number of time points enables more robust analysis that better accounts for long term trends in exposed and unexposed groups.	3: >5 time points – with at least 2 after policy start 2: 3-5 time points – with at least 2 after policy start 1: Only one time point after policy start.
Response /follow up bias	Response / follow up rate of >80% is low risk, 60-79% is moderate risk, and <60% or not reported response rate is high risk. Measures taken to adjust for response bias/ attrition using weights reduce risk.	3: Response & follow up rate>80% 2. Response & follow up rate>60% & <80% & data weighted for non-response / loss to follow up. 1: Response / follow up rate <80% / or not reported & data not weighted for non-response / loss to follow up.
Exogeneity of policy exposure	The potential for bias will depend on the extent to which variation in exposure to the policy change is likely to be exogenous (unlikely to be associated with confounders – nearly random)	3: Policy variation is as good as random, un targeted roll out / arbitrary eligibility criteria. 2: Policy variation depends on administrative decisions unlikely to be associated with

		<p>outcomes. E.g. different jurisdictions.</p> <p>1: Policy variation relates to targeting /uptake / differential adoption of policy – likely to be associated with outcomes. E.g. targeting areas with poor initial outcomes.</p>
Confounding	<p>The potential for confounding factors to bias the results will depend on:</p> <p>2. Whether measured confounders were adequately adjusted for in the analysis (Age, Sex, Health status, Labour market conditions, wage, education or occupation.)</p> <p>3. Whether methods were used to account for unobserved confounders (e.g. fixed effects)</p>	<p>3: Most time varying confounders controlled for and unobserved time invariant confounders.</p> <p>2: Most time varying confounders controlled for and/or unobserved time invariant confounders controlled for.</p> <p>1: Missing important time varying confounders controlled, and unobserved time invariant confounders not controlled for.</p>
Sample size / power	<p>The likelihood of the analysis resulting in biased estimates will also depend on the power of the study. This will depend primarily on the sample size.</p>	<p>3: Priori power calculations performed indicating sufficient power / large sample size >500 observations.</p> <p>2: No power calculations – sample size 100-500</p> <p>1: No power calculations – sample size <100</p>
Statistical methods	<p>Uses appropriate methods adequately taking into account the distribution and heteroskedasticity of the data.</p>	<p>3: Appropriate statistical technique was used.</p> <p>1: Inappropriate statistical technique was used</p>
<p>1 Greenland S. Ecologic versus individual-level sources of bias in ecologic estimates of contextual health effects. <i>International Journal of Epidemiology</i> 2001; 30:1343-50.</p>		