

Evaluation of the health-led employment trials

Research protocol

Institute for Employment Studies (IES)

RAND Europe (RAND)

National Centre for Social Research (NatCen)

Learning and Work Institute (LWI)

Professor Richard Dorsett, University of Westminster

Institute for Employment Studies

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

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

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

Copyright © 2017 Institute for Employment Studies

Contents

| | |
|---|-----------|
| Key trial contacts..... | 1 |
| Abbreviations..... | 2 |
| Summary | 3 |
| Background | 3 |
| <i>Aim of the research</i> | 3 |
| <i>Research design and methods</i> | 4 |
| <i>Reporting and dissemination</i> | 5 |
| 1 Introduction | 6 |
| 1.1 Policy context..... | 6 |
| 1.1.1 The economic and personal costs of sickness absence..... | 6 |
| 1.1.2 A fund to test new services to improve health and work | 7 |
| 1.2 Introduction to Individual Placement and Support (IPS) and the two trials | 7 |
| 1.2.1 Individual Placement and Support (IPS) | 7 |
| 1.2.2 The Trial in West Midlands Combined Authority (WMCA)..... | 8 |
| 1.2.3 The Trial in Sheffield City Region (SCR) | 9 |
| 1.3 Research questions, summary of trial design and trial rationale | 9 |
| 1.4 Roles and parties involved..... | 12 |
| 1.4.1 Design and commissioning of the new services | 12 |
| 1.4.2 Evaluation consortium | 12 |
| 2 Evaluation approach and method..... | 14 |
| 2.1 Randomisation and monitoring | 14 |
| 2.2 Data collection | 15 |
| 2.2.1 Baseline data collection | 15 |
| 2.2.2 Qualitative data collection | 16 |
| 2.2.3 <i>Management Information</i> | 19 |
| 2.2.3 Service user interviews..... | 20 |
| 2.2.4 Staff interviews..... | 21 |
| 2.3 Service user surveys..... | 22 |
| 2.4 Information from national databases | 28 |
| 2.5 Phase 3: analyses and reporting..... | 29 |
| 2.5.1 Analysis of impacts on employment, health and wellbeing using survey data and data from national database | 29 |
| 2.5.2 Analysis of interview findings | 29 |
| 2.5.3 <i>Economic analysis</i> | 30 |
| 2.5.4 <i>Reporting and dissemination</i> | 33 |
| 3 Statistics and data analysis | 35 |
| 3.1 Sample size..... | 35 |
| 3.2 Planned recruitment rate..... | 38 |
| 3.3 Statistical analysis plan..... | 40 |
| 3.3.1 Summary of baseline data and flow of patients | 40 |

| | | |
|-----------|--|------------|
| 3.3.2 | Primary outcome analysis..... | 40 |
| 3.3.3 | Secondary outcome analysis..... | 41 |
| 3.3.4 | Subgroup analyses | 41 |
| 3.3.5 | Adjusted analysis | 41 |
| 3.3.6 | Interim analysis and criteria for the premature termination of the trial..... | 42 |
| 3.3.7 | Procedure(s) to account for missing data | 42 |
| 4 | Trial setting..... | 43 |
| 5 | Agreement to participate in the trial | 45 |
| 5.1 | Agreement materials..... | 46 |
| 5.2 | Agreement to survey and interview..... | 47 |
| 6 | Disclosure protocol..... | 48 |
| 7 | Risk management and mitigation..... | 50 |
| 7.1 | Risks from agreement process and how they will be mitigated..... | 50 |
| 7.2 | Risks related to data protection, confidentiality and how they will be mitigated..... | 50 |
| 7.3 | Risks and benefits from survey and how they will be mitigated..... | 51 |
| 7.4 | Risks and benefits from qualitative interviews with service users and how risks will be mitigated..... | 52 |
| 7.5 | Risks and benefits from qualitative interviews with practitioners | 53 |
| 7.5 | Potential risks for researchers..... | 54 |
| 8 | Data management | 55 |
| 8.1 | Introduction to the data collected and used in the trial and evaluation | 55 |
| 8.2 | Data management by the evaluation consortium..... | 55 |
| 8.3 | How confidentiality of data will be ensured by consortium's working practices..... | 59 |
| 8.4 | Plans for data storage, destruction, archiving after the end of the trial | 59 |
| 9 | Research management, governance and peer review | 60 |
| 9.1 | Programme and project management..... | 60 |
| 9.2 | Peer review | 61 |
| 9.2.1 | Biographies for the peer review team | 61 |
| 10 | Research timetable | 63 |
| | Appendix A: Evaluation design and approach | 67 |
| | Phase 1: Design..... | 67 |
| | <i>Literature review and scoping</i> | <i>67</i> |
| | <i>Developing a Theory of Change for the trials</i> | <i>68</i> |
| | <i>Finalisation of outcomes</i> | <i>69</i> |
| | <i>Finalising evaluation design and data collection tools</i> | <i>69</i> |
| | Further details on the rationale for the selection of measures within the evaluation | 72 |
| | References included in the literature review..... | 84 |
| | Appendix B: Peer review commentary | 86 |
| | Appendix C: Detailed Data Flow..... | 89 |
| | References | 100 |

Key trial contacts

| | |
|---|--|
| Chief Investigator | Stephen Bevan Institute for Employment Studies City Gate 185 Dyke Road Brighton BN3 1TL UK stephen.bevan@employment-studies.co.uk +44(0)1273 763 424 |
| Funder | Work and Health Unit (Department for Work and Pensions and Department of Health and Social Care) and NHS England |
| Sponsor | Department for Work and Pensions, Department of Health and Social Care, NHS England |
| Co-investigators (IES) | Becci Newton Helen Gray Rosie Gloster |
| Co-investigators (RAND Europe) | Emma Disley Joanna Hofman Alex Sutherland |
| Co-investigators (NatCen) | Emily Tanner Julia Griggs |
| Co-investigators (Learning and Work Institute) | Tony Wilson Kathryn Ray Liz Davies |
| Co-investigators (University of Westminster) | Richard Dorsett |
| Organisation commissioning the intervention being evaluated | Wolverhampton CCG on behalf of the West Midlands Combined Authority (WMCA) Sheffield CCG on behalf of the Sheffield City Region (SCR) |

Abbreviations

| | |
|------------|---|
| Consortium | Evaluation consortium organisations |
| CCG | Clinical commissioning group |
| DHSC | Department of Health and Social Care |
| DWP | Department for Work and Pensions |
| GP | General Practitioner |
| IAPT | Improving Access to Psychological Therapy |
| IES | Institute for Employment Studies |
| IPS | Individual Placement Support |
| L&W | Learning and Work Institute |
| MSD | Musculoskeletal disorders |
| MSK | Musculoskeletal |
| NatCen | National Centre for Social Research |
| NHS | National Health Service |
| PIS | Participant information sheet |
| PSCD | Personal, sensitive, confidential data |
| QALYs | Quality Adjusted Life Years |
| RAND | RAND Europe |
| RAT | Random allocation tool |
| RCT | Randomised controlled trial |
| SCR | Sheffield City Region |
| WHU | Work and Health Unit |
| WMCA | West Midlands Combined Authority |

Summary

Background

The health of the UK's working age population has become a focus of concern for policy-makers, healthcare professionals and employers in recent years. The 2016 Green Paper, Improving Working Lives, recognises that while work can have a number of health benefits, support for citizens who have long-term health problems or disabilities can help improve labour market participation rates, physical and psychological wellbeing and social inclusion. There are 4.6 million people of working age with a disability or long-term health condition currently out of work and, if more of these individuals are to be supported back to work it is likely that more integrated and innovative models of referral and service delivery will be needed. In 2015, the Work and Health Unit (WHU) – a joint unit between the Department for Health (DH), the Department for Work and Pensions (DWP) working with NHS England – was awarded a health and work innovation fund to develop, deliver and test new ways of working across health and work to improve individual economic, social and clinical outcomes within the objective of reducing the disability employment gap. This calls for innovation in how services are delivered and several 'complementary' interventions to exist next to each other.

Aim of the research

Using the Innovation Fund, a trial of new services to improve health and work outcomes is being implemented and will be tested in two geographic areas. In addition, an independent evaluation of this trial has been commissioned. These two local areas, and the name of the trial in each, are:

- Thrive into Work IPS – operated by the West Midlands Combined Authority (WMCA)
- Working Win – operated by Sheffield City Region (SCR)

The trial in each area will use variants of the Individual Placement and Support (IPS) model of supported employment. IPS is a well-evidenced voluntary employment programme for supporting people with severe and enduring mental health needs in secondary care settings to find paid employment. The trial will be assessing whether an IPS intervention has the impact of improving health, wellbeing and employment outcomes beyond what can be achieved with 'business as usual' (BAU or usual care) among participants:

- Who have been out of work for 4+ weeks due to a health condition or disability (West Midlands Combined Authority (WMCA))
- Who are in- or out of work due to mild/moderate mental health problems and/or mild/moderate musculoskeletal disorders (MSDs) (Sheffield City Region (SCR))

The trial also offers the opportunity to test whether being in work as a result of a supported employment intervention such as IPS is associated with improvements in a number of health and wellbeing outcomes among participants and whether it also supports improved self-management of health by guiding participants into 'appropriate' work which is matched well to their aspirations and which provides support and opportunities for fulfilment which might help to sustain both improved

4 - Evaluation of the health-led employment trials

health and employment outcomes. The national evaluation will aim to answer the following questions:

1. What impact, if any, does the provision of IPS type services to the selected client groups have upon them attaining and sustaining employment¹ and benefit receipt?
2. What impact, if any, does the provision of IPS type services to the selected client groups, have upon self-reported health, self-management of health and wider wellbeing, and upon health service usage?
3. What costs are incurred and what benefits arise from the provision of IPS type services to the selected client groups?
4. How are any impacts of the trial upon sustained employment, benefits receipts, health and wellbeing achieved? What is the causal pathway to the impacts that are achieved? How might poor or negative outcomes for some participants be explained? What system-level characteristics (e.g. stakeholder cooperation, relationships with employers, awareness among GPs) need to be in place if similar interventions are to be successful and adopted in other locations or settings?²

While there is good evidence that IPS is associated with positive outcomes for people living with severe and enduring mental illness, it is not well-evidenced for other health conditions and in other settings. The trial therefore has an underlying hypothesis that IPS will also work in wider health-settings, and for other health conditions.

Research design and methods

The national evaluation will be using a mixed methods design using a randomised controlled trial, a process evaluation and an economic evaluation.

- **Randomised Controlled Trial (RCT):** individuals agreeing to participate will be randomly allocated to either a treatment group (the new service) or a control group (support as usual). Data needed to answer the research questions will be collected on both groups from the following sources: data collected by Employment Specialists delivering the new service; health service usage data collected and held nationally by NHS Digital; data about earnings and employment and benefit held nationally by HMRC and DWP; and surveys of trial participants carried out in the course of the evaluation. The evaluation will use these data, in order to reach assessments of the impact, benefits and costs of the trials.
- **Process evaluation:** alongside the quantitative analysis, the evaluation will invite trial participants in treatment and control groups to complete an interim and follow-up user surveys, and will undertake qualitative research (interviews) with users, staff and other stakeholders such as employers in each area, to understand how the trial is operating as well as to unpick the causal pathway to outcomes based on theories of change that were agreed at design phase.
- **Economic evaluation:** our approach to economic evaluation seeks to answer the question whether the benefits of the IPS interventions exceed the costs. We will conduct a cost-benefit-analysis (CBA), i.e. a valuation of individual and wider social/economic impacts in monetary terms (including future discounted benefits), from which the costs for particular interventions can be subtracted to derive an 'analysis of value for money', e.g. benefit-cost ratios.

¹ The working definition of 'sustained' employment we have adopted is 13 weeks in competitive employment although this remains in discussion

² The causal pathways will be an interpretation drawn from the Theory of Change and process evaluation

5 - Evaluation of the health-led employment trials

The trial is intended to supply evidence to inform and support future policy making – nationally and in the two regions where it operates – about approaches and services that enable people with health conditions and disabilities to return to work and to achieve improved health and wellbeing.

Reporting and dissemination

The national evaluation will produce reports looking at the findings for each of the areas where the trial operates and, additionally, a synthesis report which draws together data and other findings from across sites. These reports will be made accessible to all key audiences for the research, including policy-makers, healthcare professionals, employment support professionals and employers. In addition to formal reports, articles, presentations and local workshops for practitioners will be delivered as part of a dissemination strategy once the evaluation is complete.

1 Introduction

1.1 Policy context

1.1.1 The economic and personal costs of sickness absence

The health of the UK's working age population has become a focus of concern for policy-makers, healthcare professionals and employers in recent years. The 2016 Green Paper, *Improving Working Lives*³, recognises that while work can have a number of health benefits, support for citizens who have long-term health problems or disabilities can help improve labour market participation rates, physical and psychological wellbeing and social inclusion. Although the rate of sickness absence in the UK, as measured by the number of working days lost, has fallen in recent years, the cost to the state, employers and the wider economy is estimated to be between £15-20 billion⁴. This includes the costs to employers and Government through lost production, the payment of statutory sick pay as well as the impact of reduced tax receipts and the costs of welfare benefits. There are 4.6 million people of working age with a disability or long-term health condition currently out of work⁵. Reducing the extent of sickness absence in the UK and in particular long-term sickness absence has been a policy priority for at least the last 10 years as the benefits from even a small reduction in sickness absence would be significant. The Black/Frost Review⁶ and the RAND report of 2014⁷ amongst others concluded that service delivery needs to be more integrated, new referral routes are required, and innovative solutions should be considered.

There are two main conditions that disproportionately lead to people struggling to maintain or gain employment: mental illness and musculoskeletal disorders. Around one in six working age people in England has a mental health condition at a given point in time. Of these, the majority has either depressive disorders, anxiety disorders, or a mixture of the two conditions (McManus et al., 2009). Almost a quarter of Jobseeker's Allowance claimants and more than 40 per cent of incapacity benefits⁸ claimants have a mental health problem. The Mental Health Task Force recently concluded that current service provision is insufficient to meet the demand (Mental Health Task Force, 2016). So, more needs to be done, and given the pressure on government finances, may need to be done differently and more efficiently.

Musculoskeletal disorders comprise about 55 per cent of all work-related illness and are the second most commonly identified cause of long-term absence for manual workers (44 per cent)⁹.

³ *Improving Lives: The Health, Work and Disability Green Paper*, Department of Work and Pensions & Department of Health, Cm 9342, October 2016.

⁴ Ibid

⁵ Op Cit.

⁶ Black C and Frost D, (2011), *Health at work – an independent review of sickness absence*, London: Department of Work and Pensions.

⁷ Van Stolk, C, Hofman J, Hafner M and Janta B (2014), *Psychological Wellbeing and Work: Improving service provision and outcomes*, Cambridge: RAND Europe.

⁸ Most typically Employment and Support Allowance

⁹ Zheltoukhova K, Bevan S and Reich A, *Fit for Work? Musculoskeletal Disorders and the Australian Labour Market*, London: The Work Foundation, 2012.

7 - Evaluation of the health-led employment trials

We also know that the risk of employees leaving the workplace is even greater where a mental health issue co-occurs with a physical health issue.¹⁰

There is substantial evidence that shows ‘good’ work is beneficial for physical and mental health, whereas unemployment and long-term sickness absence often have a harmful impact (Marmot and Bell, 2012). The Black Review – Working for a Healthier Tomorrow¹¹ – recognised that there is strong evidence that work, health and wellbeing are closely linked and need to be addressed together. The costs to society, individuals, government and the economy are substantial.

1.1.2 A fund to test new services to improve health and work

The RAND report for the Department for Work and Pensions (DWP) in 2014¹² identified some key limitations of how services are delivered to encourage individuals with common mental health issues to stay in or to gain employment. These included: low rates of diagnosis or referral to specialist health and employment support; the silo working of services that tackle either the mental health or the employment need as unconnected issues; delayed service provision that leads to both health and employment problems worsening; a one-size fits all solution which is unlikely to work for all. The report also made some specific recommendations on which approaches could be piloted by DWP and the Department of Health and Social Care (DHSC). This included: extending a service called Individual Placement and Support (IPS) into primary care settings, and expanding the use of modalities such as telephone, online, and group therapy in different settings.

In 2015, the Work and Health Unit (WHU) – a joint unit between the DH, the DWP, working with NHS England – was awarded a health and work innovation fund to develop, deliver and test new ways of working across health and work to improve individual economic, social and clinical outcomes within the objective of halving the disability employment gap. This calls for innovation in how services are delivered and several ‘complementary’ interventions to exist next to each other.

Using the Fund, a trial of new services to improve health and work outcomes is being implemented and will be tested in two geographic areas. In addition, an independent evaluation of this trial has been commissioned. These two local areas, and the name of the trial in each, are:

- Thrive into Work IPS – operated by the West Midlands Combined Authority (WMCA)
- Working Win – operated by Sheffield City Region (SCR)

1.2 Introduction to Individual Placement and Support (IPS) and the two trials

1.2.1 Individual Placement and Support (IPS)

Individual Placement and Support (IPS) is a well-evidenced voluntary employment programme for supporting people with severe and enduring mental health needs in secondary care settings to find paid employment. The trials will test whether IPS is effective in wider settings and with different groups of people.

The IPS approach is based on eight key principles:

- It aims to get people into competitive employment

¹⁰ <https://www.gov.uk/government/publications/health-and-wellbeing-at-work-survey-of-employees>.

¹¹ Working for a healthier tomorrow: Dame Carol Black's Review of the health of Britain's working age population, Department of Work and Pensions & Department of Health, October 2008.

¹² Op Cit

8 - Evaluation of the health-led employment trials

- It is open to all those who want to work
- It tries to find jobs consistent with people's preferences
- It works quickly
- It brings employment specialists into clinical teams
- Employment specialists develop relationships with employers based upon a person's work preferences
- It provides time unlimited, individualised support for the person and their employer
- Access to specialist benefits counselling is included.

IPS defines competitive employment as a job that any person can apply for regardless of disability status. These jobs may be full or part time and can include self-employment. Workers in these positions should earn at least minimum wage, and receive similar wages and benefits as their co-workers. Volunteering, training, and work placements are not considered to be outcomes but may, in specific cases, be activities that help an individual to get a competitive, paid job.

As noted, the effect of IPS on the health and employment outcomes of participants is well evidenced for supporting people with severe and enduring mental health needs in secondary care settings (although these do not cover the participant groups that the WMCA and SCR trials will attract). As such, a fidelity scale has been developed which measures the degree to which IPS interventions implement evidence based practice. The two local sites will be responsible for assessing the fidelity of their IPS interventions although the planned process evaluation will also explore this.

1.2.2 The Trial in West Midlands Combined Authority (WMCA)

The service that will be trialled in the West Midlands is called 'Thrive Into Work'. On a voluntary basis, people can be referred to service by their clinician and it is open to anyone with a health condition who has been out of work for 4+ weeks due to these health conditions. The service involves Employment Specialists providing support to service users to conduct a rapid job search, tailored to their individual aspirations and skills. Once in work, Employment Specialists continue to support both employee and employer as appropriate. In parallel, service users access health treatment offered by NHS clinicians. Aside from working with service users, Employment Specialists also engage with clinicians to change the culture of the NHS to be much more focused on employment as a driver of health and wellbeing.

The service will be commissioned by Wolverhampton CCG on behalf of the West Midlands Combined Authority, DWP, DH, and NHS England.

This service will follow the principles of time-limited IPS, which adapts principle vii, above, to limit the length of time that an employment specialist will support a client to nine months, or four months after the client enters work. This has been shown to allow more people to use the service without impacting outcomes (Burns et al, 2015). The provider may exceed this time limit on a case-by-case basis where individuals are actively seeking employment or in need of continued in-work support. However, the maximum service duration for all individuals will be 12 months from initial engagement.

Details of the services planned by WMCA are contained in its Trial Protocol.

1.2.3 The Trial in Sheffield City Region (SCR)

SCR will trial an employment service co-located within primary care and co-ordinated with the patient's health support, with the service known as Working Win. The service will be implemented in Barnsley, Bassetlaw, Doncaster, Rotherham and Sheffield later this year, which is a complete Sustainability Transformation Plan (STP) geography. The service provides employment support to individuals with mild to moderate mental health and/or musculoskeletal (MSK) conditions and who are either unemployed and seeking work or who are in work but struggling at work as a result of their health or who are off-sick. Individuals will be voluntary participants of the service.

The SCR service is based on a modified form of Individual Placement and Support (IPS) service and IPS fidelity scale, and has been developed through an on-going process of intensive co-design and engagement with local health partners and service users.

Service users will be given the support of trained Employment Specialists – a personal caseworker for their core employment support needs and a co-ordinator to wider health and support needs. These case workers will be responsible for supporting both in and out of work participants, drawing on a range of activities/approaches to support their clients as required.

Referrals will come mainly from key elements of the local health system – GP practices, Improving Access to Psychological Therapy (IAPT) teams, physiotherapists and pain management teams – as well as from existing social prescribing caseloads and via self-referral. Employment specialists will be co-located inside existing health teams such as IAPT services, physiotherapists, pain management, GP practices and will attend these primary care team meetings to seek to better integrate patients' work and health support. Employment specialists will meet with clients in these accessible frontline locations.

SCR plans to lead marketing and communications activity for which posters and leaflets are appended to this submission. It will also lead a social media campaign based on the following messages:

- We're trialling a new type of job support for people who have mild to moderate mental health condition, and/or a physical health condition. This is a clinical research trial, designed to identify whether such support can positively improve the health of patients.
- We know that being in good work can support good health. That's why we're delivering dedicated and personalised support, helping people to find work, or to stay in work if they're struggling.

Details of the services planned by SCR are contained in its Trial Protocol.

1.3 Research questions, summary of trial design and trial rationale

Primarily, the trial will be assessing whether an Individual Placement and Support (IPS) intervention has an impact of improving health, wellbeing and employment outcomes beyond what can be achieved with 'business as usual' (BAU or usual care) among participants:

- Who have been out of work for 4+ weeks due to a health condition or disability (West Midlands Combined Authority (WMCA))
- Who are in- or out of work due to mild/moderate mental health problems and/or mild/moderate musculoskeletal disorders (MSDs) (Sheffield City Region (SCR))

However, the trial also offers the opportunity to test whether being in work as a result of a supported employment intervention such as IPS is associated with improvements in a number of health and wellbeing outcomes among participants and whether it also supports improved self-

10 - Evaluation of the health-led employment trials

management of health by guiding participants into 'appropriate' work which is matched well to their aspirations and which provides support and opportunities for fulfilment which might help to sustain both improved health and employment outcomes.

While IPS is well-evidenced in respect of supporting and impacting on the outcomes of individuals with severe and enduring mental health needs, it is not well-evidenced for other needs and in other settings. The trials therefore have an underlying hypothesis that IPS will also work in wider health-settings, and for other health conditions.

As such, the evaluation will focus on three core areas of impact:

- Impact of employment and health outcomes - does IPS-type support increase employment and improve health outcomes relative to BAU/Usual Care?
- Process impact - what are the 'process' components of the service which contribute most to any positive health and employment outcomes measured in the evaluation?
- Economic impact – what are the economic impacts of the trial interventions, over and above BAU/Usual Care, including on the 'whole system' costs and benefits across agencies and stakeholders?

In focusing on these aspects of impact, the evaluation will aim to answer the following questions:

1. What impact, if any, does the provision of IPS type services to the selected client groups have upon them attaining and sustaining employment¹³ and benefit receipt?
2. What impact, if any, does the provision of IPS type services to the selected client groups, have upon self-reported health, self-management of health and wider wellbeing, and upon health service usage?
3. What costs are incurred and what benefits arise from the provision of IPS type services to the selected client groups?
4. How are any impacts of the trial upon sustained employment, benefits receipts, health and wellbeing achieved? What is the causal pathway to the impacts that are achieved? How might poor or negative outcomes for some participants be explained? What system-level characteristics (e.g. stakeholder cooperation, relationships with employers, awareness among GPs) need to be in place if similar interventions are to be successful and adopted in other locations or settings?¹⁴

These research questions will be evaluated using a randomised controlled trial, a process evaluation and an economic evaluation.

- **Randomised Controlled Trial (RCT):** individuals who agree to participate will be randomly allocated to either a treatment group (the new service) or a control group (support as usual). Data needed to answer the research questions will be collected on both groups from the following sources: data collected by Employment Specialists delivering the new service; health service usage data collected and held nationally by NHS Digital; data about earnings and employment and benefit held nationally by HMRC and DWP; and surveys of trial participants carried out in the course of the evaluation. The evaluation will use these data, in order to reach assessments of the impact, benefits and costs of the trials.

¹³ The working definition of 'sustained' employment we have adopted is 13 weeks in competitive employment although this remains in discussion

¹⁴ The causal pathways will be an interpretation drawn from the Theory of Change and process evaluation

11 - Evaluation of the health-led employment trials

- **Process evaluation:** alongside the quantitative analysis, the evaluation will invite trial participants in treatment and control groups to complete an interim and follow-up user surveys, and will undertake qualitative research (interviews) with users, staff and other stakeholders such as employers in each area, to understand how the trial is operating as well as to unpick the causal pathway to outcomes based on the theories of change¹⁵ that were agreed at design phase.
- **Economic evaluation:** our approach to economic evaluation seeks to answer the question whether the benefits of the IPS interventions exceed the costs. To this end, we propose a cost-benefit-analysis (CBA), i.e. a valuation of individual and wider social/economic impacts in monetary terms (including future discounted benefits), from which the costs for particular interventions can be subtracted to derive an 'analysis of value for money', e.g. benefit-cost ratios. This allows one to say that for every pound spent on IPS-type interventions for this group, the return on investment (benefit) is x pounds.

The trial is intended to supply evidence to inform and support future policy making – nationally and in the two regions where it operates – about approaches and services that enable people with health conditions and disabilities to return to work and to achieve improved health and wellbeing.

- In particular, the study aims to fill the following evidence gaps:
 - Scale: The IPS model has been tested in the UK mainly through small-scale feasibility pilots. The large sample in these trials will have greater statistical power to detect effects if they exist (and allow for subgroup analysis).'
 - Broader range of health conditions: The evidence of the effectiveness of IPS with participants beyond the traditional mental health groups, such as people with spinal injuries (Ottomanelli et al, 2014) and chronic back pain (Coole et al, 2013) is only slowly emerging. The programme has previously been used for people with severe and enduring mental health conditions but the trials will test the IPS support model with people who have either mild to moderate mental illness or musculoskeletal disorders which have an impact on their ability to find work. The trial in SCR will also test the model with a sub-sample of people in work, as well as out of work – a group not usually included in IPS-type interventions.
 - Adaptations to the model: Certain adaptations of the IPS fidelity model, such as time-limited support, have been tested only with small samples. This trial will contribute to building the evidence base for the IPS-LITE model.
 - Relationship to non-employment outcomes has not been widely explored. This trial will examine the effect of IPS on health outcomes, self-management and wider wellbeing.
 - Use of linked administrative data: these trials will be the first in the UK to analyse health and employment outcomes using national administrative data.

¹⁵ Please refer to Appendix A for further details of the approach to their design. Theory of Change is an approach used to map the connections between activities and outcomes within an intervention, to generate hypotheses about how the intervention will achieve the desired change. Using this approach within evaluation facilitates a more systematic focus on explaining how and why an intervention works (or does not work).

1.4 Roles and parties involved

1.4.1 Design and commissioning of the new services

The services are funded by NHS England, the Department for Work and Pensions (DWP), and the Department of Health and Social Care (DHSC).

West Midlands Combined Authority (WMCA)

In the West Midlands, this programme is a key element of the Thrive West Midlands action plan developed by the West Midlands Mental Health Commission in 2016,¹⁶ although the focus of this trial will be broader than mental health. The programme is championed by the West Midlands Combined Authority, with procurement led by Wolverhampton CCG in conjunction with Arden & GEM Commissioning Support Unit (CSU). The overall project sponsor is Sarah Norman, CEO of Dudley Council, and governed by the Health and Wellbeing Board of the Combined Authority. Strategic leadership will be provided by Sean Russell, Director of Implementation for the West Midlands Mental Health Commission within the Combined Authority. The development of this service specification has been informed by the commissioning and policy guidance within the 'The Five Year Forward View' (2014), 'The Five Year Forward View for Mental Health' (2016) and the 'General Practice Forward View' (2016). This is to support people of all ages to access high quality mental and physical health care, and primary care and community based recovery care pathways including specialist support for people with mental health difficulties.

Sheffield City Region (SCR)

Sheffield City Region (SCR) has a commitment to inclusive growth and enabling all residents to benefit from the employment opportunities in the city-region. The South Yorkshire and Bassetlaw STP places a strong emphasis on preventive care and the important role of wider social determinants to this preventative care, with tackling health-related unemployment central to this. The SCR trial has been led by a co-design group made up of the SCR co-design lead and 1-2 health colleagues from each of the five areas who are, for example, Directors of Public Health, Directors of Health Improvement, CCG Director of Quality and Care, and a WHU representative. The SCR trial will be performance managed by SCR and its commissioning will be led by Sheffield CCG on behalf of the 5 CCGs areas involved in the trial. The local project sponsor is Dr Ruth Adams, Deputy Executive Director, SCR and the SCR Skills, Employment and Education Board will be the governance body with lead responsibility for trial performance management, feeding into the overall Combined Authority Board. A regular steering group is to be established within SCR that brings together SCR and local authority partners with health partners from the STP and CCGs as well as national partners in order to review the progress of the trial and opportunities for future partnership working.

1.4.2 Evaluation consortium

The evaluation is being conducted by independent researchers appointed by the Department of Work and Pensions.

The consortium is led by the **Institute for Employment Studies** (IES), which is leading on programme and project management and on the economic evaluation. IES is an independent, apolitical, international centre of research and consultancy in public employment policy and HR issues. It works closely with employers in all sectors, government departments, agencies, professional bodies and associations. IES is a focus of knowledge and practical experience in

¹⁶ <https://www.wmca.org.uk/what-we-do/mental-health-commission/>

13 - Evaluation of the health-led employment trials

employment and training policy, the operation of labour markets, and HR planning and development. IES is a not-for-profit organisation.

There are four other partners in the Consortium:

- **The Learning and Work Institute (L&W)** was formed in January 2016, following the merger of the Centre for Economic and Social Inclusion and NIACE. L&W's combined organisation brings a wealth of expertise in programme evaluation, randomised trials, community-based learning and support for hard-to-reach groups. L&W researches what works, develop new ways of thinking and help to implement new approaches. It is leading on the process evaluation including agreeing and testing theories of change for the intervention.
- **National Centre for Social Research (NatCen)** is the UK's largest independent social research agency. NatCen are not-for-profit and focused on delivering high quality social research that can improve people's lives. NatCen's clients span central government, the third sector, universities, local government and the private sector. NatCen is a full-service agency with a strong track record in delivering surveys, qualitative research, secondary analysis and evaluation. NatCen carries out many of the UK's most highly regarded surveys and many challenging qualitative studies with groups regarded as 'harder to reach'. NatCen is leading on the survey component of the evaluation.
- **RAND Europe** is a not-for-profit research institute whose mission is to help improve policy and decision-making through research and analysis. As the European arm of the RAND Corporation, RAND Europe shares its mission and values. Over the last 20 years RAND Europe's work has steadily expanded to cover a wide range of policy-relevant topics from innovation and technology policy, through to health and social issues. Cutting across topical areas are the methods-focused groups, including evaluation and impact measurement, and choice modelling and valuation. The mix of subject expertise, contextual understanding and innovative methodologies creates insightful yet robust analysis. RAND Europe is leading on the ethical approval process for the evaluation design and supporting the impact, process and economic evaluations.
- **Richard Dorsett** is professor of Economic Evaluation at the University of Westminster Business School (WBS). Established in 1997, WBS is an internationally-facing, professional and research-engaged business school, with a mission to facilitate the development of the business and management careers of our students in a complex and uncertain professional world. WBS does this by drawing on its applied research, practitioner expertise and management development experience to offer a wide range of business and management programmes. Richard is leading on the impact evaluation.

2 Evaluation approach and method

The evaluation of the trials will have a clear focus on three interlinked core elements:

- Evaluation of impact
- Process evaluation
- Economic evaluation

The task of the national evaluation will be to balance each of these components and to ensure that linkages between them are captured and exploited to inform to over-arching research objectives. The evaluation will collect, isolate and analyse **quantitative** data about the impact and effectiveness of the IPS intervention on trial participants and develop a thorough understanding, through strong theory-of-change based process evaluation of the **qualitative** aspects of the operation of the trials at local level. Together, these elements may allow conclusions to be drawn about the transferability and the scalability of these interventions in other locations and settings.

In recognition of its status as the 'gold standard' for impact estimation, the evaluation has been designed as a randomised control trial (RCT). This is with the aim of achieving the highest quality evidence on effectiveness. The challenge is to be able to estimate how IPS participants' outcomes would have looked had they not participated. The difficulty arises from the fact that participants are likely to be very different from non-participants. An RCT addresses this by making participation random, with the consequence that such differences are removed. While other evaluation approaches might be possible in principle, they rely on stronger assumptions than an RCT. Furthermore, these assumptions are not verifiable. Consequently, the resulting estimates are often less convincing and more open to challenge. Careful design is critical to RCTs, and this is described in detail in chapter 3.

This following chapter builds on the design phase (see appendix A) to describe how the evaluation will be implemented. During this phase, participants will be recruited to the trial, data will be collected and process interviews will be carried out. Details on each of these aspects are presented below.

2.1 Randomisation and monitoring

At the first meeting with a service user, once eligibility has been assessed, the Employment Specialist will administer the agreement process and those agreeing to participate in the trial will be randomised into the control or intervention group. The trial will open for recruitment in Spring 2018 and close to recruitment in Summer 2019. In WMCA, it is the intention to recruit 2,650 in the treatment arm, 2,650 in the control arm. In SCR, it is the intention to recruit 3,750 in the treatment arm, 3,750 in the control arm.

2.1.1 Randomisation software tool

Randomisation will be carried out during face-to-face meetings between the referred individual and the Employment Specialist. The Employment Specialist will perform the randomisation using a

15 - Evaluation of the health-led employment trials

bespoke random allocation tool (RAT) which is an adapted version of the tool produced by Behaviour Insights Team for the Islington IPS trial. The RAT will be served on the internet, and has been designed to collect a baseline assessment of participants (treatment and control) which includes personal and sensitive information. Server space is being commissioned that will provide the security that is needed to hold such data. The RAT has been designed to transfer data to the safe haven (ONS) every 24 hours.

Randomisation will be stratified by site and, within SCR, by whether an individual is in work or otherwise at the time of referral. It will use a permuted block approach to prevent the outcome of randomisation being predictable therefore avoiding Employment Specialists being able to affect the randomisation outcome. Trial participants will be immediately informed of their allocated group, and the RAT will automatically send a confirmatory email to the email address they have given. For those allocated to the intervention group, the Employment Specialist and the client will immediately arrange the first IPS session. For those allocated to the control group, signposting information to the BAU support will be provided.

2.1.2 Randomisation ratio

Randomisation will be at the level of the individual. Trial participants will be allocated according to a randomisation ratio that may alter over the course of the trial to allow for caseload management. To begin with, a randomisation ratio of 50% will be used. Beyond this point, the randomisation ratio will be adjusted if there is a risk of insufficient capacity to deliver the IPS treatment. Monitoring data collected on an on-going basis throughout the course of the trial will provide the required information on Employment Specialist capacity. Where the projected inflow over the course of the following would result in Employment Specialist capacity being exceeded, the randomisation ratio will be adjusted to allocate a higher proportion of trial participants to the control condition, sufficient to bring projected caseload into line with projected capacity. To be effective in managing caseload, this adjustment will need to take place at the site level. Consequently, there may be variation across sites in the randomisation ratio in place at any time. Furthermore, as noted earlier, the SCR trial will include both in- and out-of-work participants. The aim is to achieve a 30/70 split between in- and out-of-work IPS participants. The randomisation ratio will be allowed to vary across these two groups in SCR in order to achieve that target.

Randomisation will be monitored to ensure that (1) the randomisation tool successfully assigns the specified proportion of individuals to the treatment arm and (2) that the characteristics of the treatment and control arms will be produced in order to confirm that two similar-looking groups are identified. At this stage, we expect that randomisation reports will be produced weekly in the first six weeks of the trial – to allow formative assessment and adjustments to be made as necessary, reducing to fortnightly and then, three months after trial start, monthly.

2.2 Data collection

2.2.1 Baseline data collection

Employment Specialists in the local sites will record management information on those randomised for the trial. There are several reasons why this is important. First, it allows the characteristics of the participant population to be understood. Second, it allows important subgroups to be identified. Third, it is required as means of achieving data linkage. Fourth, it will be incorporated into the impact analysis as a means of achieving improved statistical power. Fifth, it provides an insight into the nature of survey non-response and a possible means of addressing it.

The data collected at baseline will cover the following items:

- Personal data (first name; surname; postcode; date of birth; national insurance number and National Health Service number). These items will be required to be able to obtain data on those randomised from national databases and to carry out the process study and surveys.

16 - Evaluation of the health-led employment trials

- Demographic information and background characteristics, including employment status and experience; benefit receipt, job search activity, baseline health and wellbeing and barriers to work. Items such as this will be required to carry out any analysis of the impact of the trials on different subgroups of participants and to determine whether the characteristics of the intervention and control groups were similar at the time of randomisation. They will also be used to assess changes in circumstances over time and to draw the sample for the process study.
- Information on the referral (such as when the referral was received and the source), when randomisation occurred and information on participation in any other trials. Such data will be used to explore differences in referral routes, or between Employment Specialists, and to identify any potential contamination from other trials running concurrently. It will also be used to determine when to survey participants, as the interim and final surveys will be conducted at fixed points after randomisation.

Baseline information will be used as a benchmark to judge progress as a result of the trials. For example, Employment Specialists will be asked to record responses to a small number of questions on health, employment, job search and wellbeing. These include EuroQol-5D-5L which covers mobility, self-care, usual activities, pain/discomfort and anxiety/depression; the Short Warwick-Edinburgh Mental Well-being Scale (SWEMWBS) which covers both feeling and functioning aspects of mental wellbeing; the Office for National Statistics Personal Well-being Question on life satisfaction (ONS-1); a question on whether the individual is currently in paid employment, self-employed, unemployed, in education/training, inactive, a carer or retired; and the Job Search Self-Efficacy Scale (JSSE) which asks about self-efficacy relating to finding employment. The same questions will then be repeated on interim and final surveys so that the impact of the service can be assessed.

This management information collected by Employment Specialists will be the primary source of baseline data. The baseline information (other than the personal data) will then be combined with the on-going management information, interim and final surveys and national administrative data to create a pseudo-anonymised dataset which gives a full picture of individual circumstances and characteristics at baseline, trial participation and outcomes. Whilst the management information will be the primary source of baseline data, the national administrative databases will also provide information on the characteristics of participants at the time that they are randomised, as well as historical information on their circumstances prior to randomisation.

2.2.2 Qualitative data collection

The process study is designed to support the impact study by explaining how and why the intervention worked (or did not work). This will include examination of whether it was more or less effective for different client groups and in different locations, as well as which elements of support appeared to be most effective in achieving outcomes. This will enable the identification of the trials' critical success factors which will support its transfer to other contexts.

The Theories of Change inform the process study design since they make explicit what support activities need to be delivered in order to achieve the desired outcomes (for clients, but also for health professionals and employers) and set out a series of hypotheses about the causal pathways by which the long-term outcomes will be achieved. The process study research is designed to explore whether and how these pathways operate in practice and how they vary for different clients and across different locations.

There are three key areas of exploration for the process study:

1. Whether **implementation and delivery** reflects intended design, as captured in the Theories of Change – by exploring participant, provider and partner views on fidelity, adaptations and what worked.

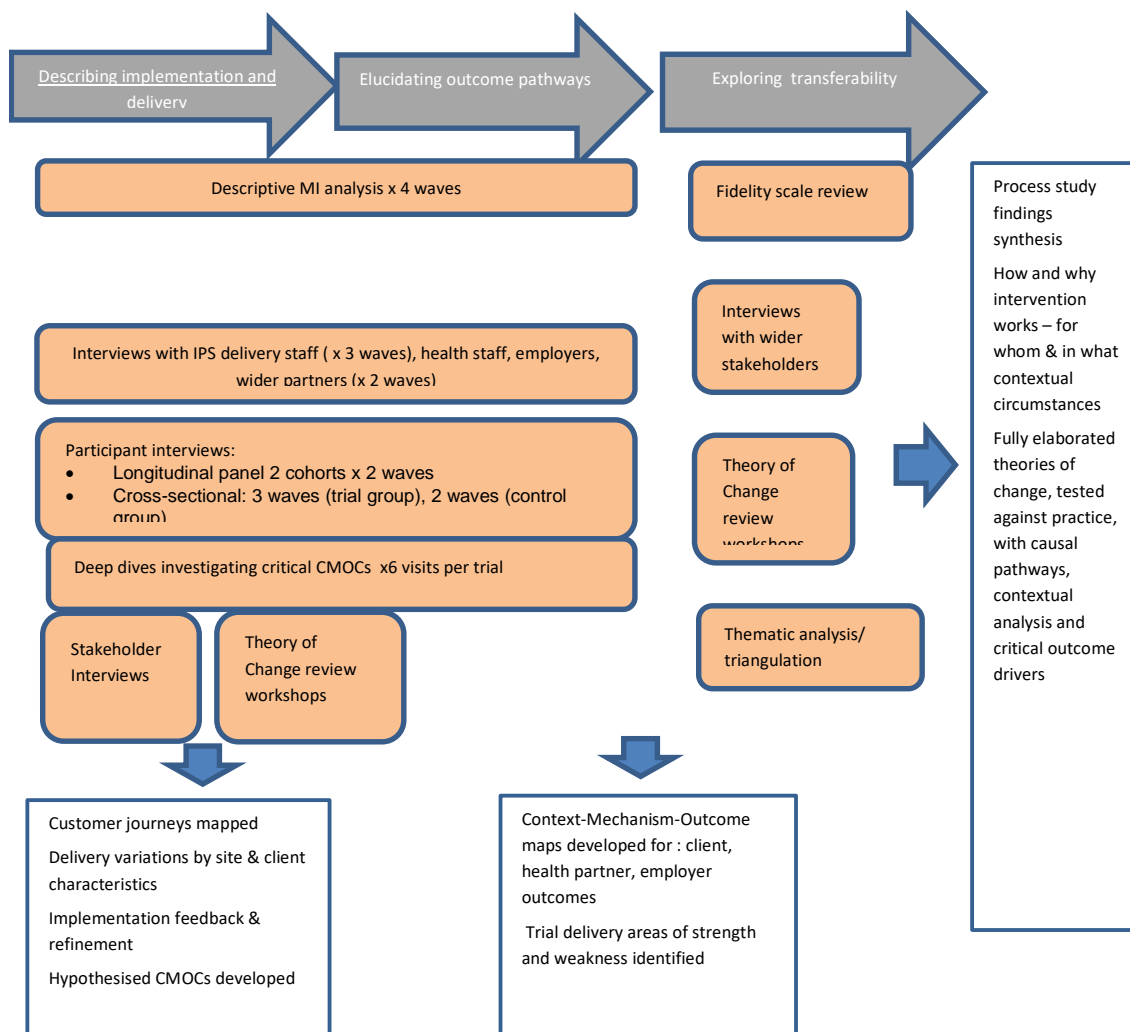
17 - Evaluation of the health-led employment trials

2. Pathways between **activities, intermediate and longer-term outcomes** as envisaged in the different Theories of Change (including variation across sub/groups and delivery sites), and the contribution of different modes of support and contextual factors.
3. **Transferability** – the potential to identify key learning, what works best, and what is transferable to other contexts.

To explore these issues, we propose a mixed methods design. The following strands of activity are planned as outlined below; however, exact timings and numbers of planned waves of activity may be amended during the delivery of the evaluation in order to flexibly respond to emerging data and intelligence from the trials:

- Descriptive MI analysis at four points across the delivery phase, exploring activities undertaken and outcomes achieved by client characteristics and delivery location;
- Depth interviews with IPS staff, core health teams, wider health staff, employers and other delivery partners, exploring implementation, delivery and outcomes, most likely undertaken in multiple waves over the delivery phase;
- Participant research which we expect to include two longitudinal panels of up to 12 participants to explore outcome pathways as they unfold for participants with varying characteristics, and multiple waves of cross-sectional interviews exploring journeys through the support for participants with different outcomes;
- 'Deep dives' focusing on emerging good practice, comprising site visits with observations, participant focus groups and interviews;
- Depth interviews with wider stakeholders and Theory of Change review workshops to explore contextual issues and transferability.

Figure 2.1: Process study design



Source: Evaluation Consortium 2017

These multiple methods of data collection ensure that:

- The views and experiences of all of the different groups that design, deliver or participate in the trial activities are captured;
- Variations in delivery both over multiple locations and over time are captured, through successive waves of fieldwork and sampling each respondent group by location.

2.2.3 Management Information

Management information (MI) collected by the trials will contribute to the process and impact evaluation. The types of information we expect to be collected have been informed by the activities and assumptions outlined in the Theories of Change. An indicative list of MI is given below, but this list will be refined and agreed with the trial areas, informed by the practicalities of data systems, Employment Specialist resource for data capture and the information needs for trial delivery. In addition to the baseline data noted above, the MI will capture information as follows.

Information to be recorded at each IPS meeting is likely to include:

- Meeting date
- Mode of support
- Referrals made to external services (including health services)
- Support received from external services
- Jobs applied for since previous meeting
- Number of job interviews since previous meeting
- Number of jobs started since previous meeting
- Dates job started
- Sector/occupation of each job
- Job characteristics e.g. part-time or full-time and contract type.

In addition, if Employment Specialists have sufficient time and an appropriate mechanism for capturing the following data can be agreed with the trials we will aim to collect the following:

- Support focus of meeting (linked to the Theory of Change)
- Support need referral designed to address
- Start and end dates of external support
- Intensity or frequency of external support
- Methods used for job search

Two types of analysis will be conducted using the management information at different points of the evaluation:

- **Basic descriptive analysis:** Basic descriptive analysis will take place at regular interviews (nominally every six months) during the evaluation to inform on-going monitoring of trial implementation and progress on outcomes, including variation across local sites and by participant characteristics. This will also inform sampling for qualitative research.
- **Full analysis:** A more comprehensive analysis of the management information will take place in August 2018 and March 2020, which will investigate more fully variations within the data alongside analysis of survey data and qualitative research findings.

The time between randomisation and first IPS session will also be monitored (for those assigned to the treatment arm).

2.2.3 Service user interviews

The in-depth interviews with service users will explore the range of outcomes they have experienced as a result of engaging with the trials and the perceived linkages between activities, intermediate and longer-term outcomes. They will also explore their experiences of engaging with the service and the different support activities and identify what approaches worked well for them. Purposive sampling will allow variations to be explored across trial locations and respondent characteristics. The sampling criteria (drawn from the MI) will include: location of support/delivery site, user demographic characteristics and barriers, activities/support undertaken and outcomes achieved. We expect that interviews will last approximately 45 minutes and service users who take part in an in-depth interview will be offered a £20 gift voucher as a 'thank you' in recognition of their time and contribution.

Respondents will be given a choice of telephone or face to face interviews, and we anticipate that the majority (c80%) will opt for a telephone interview because of ease of arranging. Our experience also suggests that telephone interviews may be the preferred mode for some participants with mental health and some physical health issues. However, where participants have health conditions that make taking part by phone challenging, we will accommodate face to face interviews.

The following types of interviews will be undertaken with service users (*the numbers given are for each individual trial*):

Longitudinal panel of service users – The panel will include between up to 24 service users (in two cohorts), selected using management information to provide a broad spread based on trial location, demographic and work and health characteristics. This is in order to ensure that the interviews capture all relevant variation in support needs and experiences. Each participant in the longitudinal panels will take part in 2 interviews – the first within 2 months of the start of their support, and then at +6 months. This enables the research to explore the causal pathways that link support received to attitudinal and behavioural changes, as they are occurring.

Panel attrition is a risk if not properly managed. We will reduce the risk through diligent panel maintenance, for example, keeping in touch emails at two-monthly intervals. Using a £20 gift voucher as an incentive at each wave should also encourage continued participation. We would also propose to refresh the panel at each wave, by replacing any members who have been lost to attrition with new research participants.

At this stage, two cohorts of longitudinal panels are planned, one to begin in 2018 and one in 2019, in order to explore the extent to which any in-programme changes alter people's overall experiences. Most interviews will take place by phone, but face-to-face (on site or in agreed public places) can be accommodated if requested by the participant.

Cross-sectional in-depth interviews – We also expect to undertake approximately 36 one-off interviews with service users across the duration of the trials and 18 with the control group.

Treatment group interviews will be undertaken in three cohorts of 12 between 2018 and 2019, with cohorts 2 and 3 will be supplemented by interviews with 9 control group participants. A purposive sampling approach¹⁷ will be used to capture participants at different points in their customer journey (at each time point) and to explore the experiences of participants who have achieved different types of outcomes as a result of the trial (including soft (or intermediate) outcomes, as well as work outcomes). The sample will be drawn from MI and the interim survey, and informed by analysis of this data. Most interviews will take place by phone, but face-to-face can be accommodated (on site or in agreed public places) if requested by the participant.

Deep dives – we also expect to conduct up to 30 interviews, across the trials' duration with service users as part of 'deep dives' focusing on hypothesised key Context-Mechanism-Outcome configurations in greater detail. Mixed methods will be used as appropriate to the topics selected, which may include observations of support (subject to the agreement of staff and service users to take part in the qualitative research), interviews and group discussions with frontline staff and/or partners, interviews or focus groups with service users and additional MI analysis. Topics for the deep dives will be selected on the basis of issues emerging from the analysis, and the topics will determine the locations of the site visits and any sampling criteria used to select participants. We expect to work with provider staff in sites to help us arrange these site visits. They will take place at 2 time points: in Spring/Summer 2019 and Spring/Summer 2020.

2.2.4 Staff interviews

In-depth interviews will also take place with delivery staff, health professionals, employers, stakeholders and other partners involved in the trials. These interviews will again explore the range of intermediate outcomes experienced by service users including variations by sub-groups/sites and linkages between activities, intermediate and longer-term outcomes for users. In addition, they will also explore whether and how 'systems level' changes are occurring as captured in the Theories of Change, including changes in their own and others' attitudes and behaviour.

The following interviews with stakeholders and staff are expected (approximate numbers are given for each trial):

In-depth interviews with delivery staff – up to 36 in-depth interviews with IPS delivery staff will be undertaken over the course of the process study. These interviews will be completed in batches of 12 in Jul 2018, Jan 2019 and Sep 2019, in order to explore staff's experiences and views across the duration of trial delivery and understand the reasons for any changes. These interviews will include IPS delivery staff e.g. Employment Specialists and their managers, as well as staff from core health teams supporting participants. The delivery staff selected will be identified in partnership with the trial leads and local partners and stakeholders, to ensure all relevant variation in implementation and delivery context is captured (for example differences in the nature of co-location/integration). We would aim to ensure at least 6 respondents are interviewed at two points in time to explore changes in attitudes, views and behaviours over the course of the trials.

In-depth interviews with delivery partners and local stakeholders – Approximately 10 local stakeholders and other delivery partners are expected to be interviewed at the start and end of the trial to explore their views on the trial delivery, contextual influences and expected impacts. This will include local strategic stakeholders, as well as key referral partners, and other partners supporting delivery.

In-depth interviews with employers – We expect to engage 20 employers who have been involved with the trials (for example by employing trial participants or by providing work experience opportunities) in interviews. We expect these will be completed in two waves in Jul 2018 and Sep 2019, in order to explore employer views and experiences across the duration of trial delivery and understand the reasons for any changes.

¹⁷ Purposive sampling will allow variations to be explored across trial locations and respondent characteristics. The sampling criteria (drawn from the MI) will include: location of support/delivery site, user demographic characteristics and barriers, activities/support undertaken and outcomes achieved.

Deep dive interviews – During the deep dive visits, interviews and/or discussion groups may take place with IPS delivery staff, health professionals and other partners, or interviews with strategic stakeholders or employers, as appropriate for the topics under investigation (see section 6.2). We anticipate up to 30 staff interviews could be undertaken in this way.

It is expected that these interviews with staff, employers and partners will occur during working hours and last approximately 60 minutes, although employer interviews may be shorter (45 minutes), depending on their availability. Interviews will be arranged to be flexible and responsive to their needs. Participants in these interviews will not receive any financial incentive or 'thank you' for taking part, as it is expected that they will perceive and understand the potential benefits of their involvement, in terms of the ongoing improvement of the trials and, in the longer term, furthering the evidence base. Interviews will be confidential, and such confidentiality will only be breached if the interviewee is considered to be at risk of personal harm. The appropriate course of action for safeguarding will be decided by the project manager on an individual basis.

Translation and inclusion

We will state explicitly during recruitment that should they wish, interviewees may choose for someone else to be present during the interview or to answer any questions on their behalf. We will ask participants if they require any other special arrangements, for example having an interpreter present and/or taking breaks during the interview. Interviews will be kept to a reasonable length, so as to not overburden respondents. Our approach aims to ensure the research is accessible to all participants, and to allow us to provide services such as translation, braille and signing we have set aside a contingency fund which can be drawn on as required.

2.3 Service user surveys

Sample

Across the two sites, it is anticipated that 14,100 adults will be randomised into treatment and control groups between Spring 2018 and the end of Summer/Autumn 2019. We expect these caseloads will form a flow sample for the surveys so that everyone within each site is interviewed at a similar point in relative to when they were randomised. This section sets out how the volume of participants might be distributed over time.

WMCA anticipate that 6,600 adults across four areas with a health condition or disability who have been out of work for at least four weeks will be randomised evenly into treatment control groups over the 18 months. The site has mapped an increase in caseload over the first year as teams build their staff numbers and capacity.

SCR anticipate that 7,500 adults across five areas who are either unemployed and seeking work or who are in work but struggling/off-sick will be randomised evenly into treatment and control groups. At this stage, it is anticipated that the number of new referrals will be distributed evenly across months. On average, this will mean 1,250 referrals per quarter (625 treatment and 625 control). It is not yet known whether there will be an even split between the in-work and out-of-work groups.

We anticipate that the sample for the interim and final surveys will be formed of all those for whom baseline measures are collected and who are subsequently randomised into the treatment and control groups. The randomisation tool will record full contact details that will be transferred securely to NatCen as electronic data on a monthly basis. The interim and final surveys will then take place on a rolling basis in accordance with the numbers randomised.

The Table 2.1 shows the distribution of cases randomised per quarter according to the latest estimates by sites.

Table 2.1: Estimated numbers randomised per quarter in each site

| | Q1 Oct-Dec 17 | Q2 Jan-Mar 18 | Q3 Apr-Jun 18 | Q4 Jul-Sep 18 | Q5 Oct-Dec 18 | Q6 Jan-Mar 19 | Total |
|-------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------|
| WMCA total | 100 | 875 | 1,413 | 1,412 | 1,400 | 1,400 | 6,600 |
| <i>Treatment</i> | 50 | 428 | 707 | 706 | 700 | 700 | 3,300 |
| <i>Control</i> | 50 | 437 | 706 | 706 | 700 | 700 | 3,300 |
| SCR total | 1,250 | 1,250 | 1,250 | 1,250 | 1,250 | 1,250 | 7,500 |
| <i>Treatment</i> | 625 | 625 | 625 | 625 | 625 | 625 | 3,750 |
| <i>Control</i> | 625 | 625 | 625 | 625 | 625 | 625 | 3,750 |
| Total | 1,350 | 2,125 | 2,663 | 2,662 | 2,650 | 2,650 | 14,100 |
| Cumulative total | 1,350 | 3,475 | 6,138 | 8,800 | 11,450 | 14,100 | |

Source: Evaluation consortium based on information from local sites 2017

The **Error! Reference source not found.** illustrates how this would translate into the sample size issued per month for the interim and final surveys. The numbers do not take account of any attrition (cases that withdraw their agreement to be part of the trial). All trial participants will be invited to take part in the final survey, even if they did not respond to the interim survey.

Table 2.2: Estimated issued sample per month in each site

| | Randomised | | Interim survey | | Final survey | |
|--------|------------|-----|----------------|-----|--------------|-----|
| | WMCA | SCR | WMCA | SCR | WMCA | SCR |
| Feb-18 | 33 | 417 | | | | |
| Mar-18 | 33 | 417 | | | | |
| Apr-18 | 34 | 417 | | | | |
| May-18 | 292 | 417 | 33 | 417 | | |
| Jun-18 | 292 | 417 | 33 | 417 | | |
| Jul-18 | 291 | 417 | 34 | 417 | | |
| Aug-18 | 471 | 417 | 292 | 417 | | |
| Sep-18 | 471 | 417 | 292 | 417 | | |
| Oct-18 | 471 | 417 | 291 | 417 | 33 | |
| Nov-18 | 471 | 417 | 471 | 417 | 33 | |
| Dec-18 | 471 | 417 | 471 | 417 | 34 | |
| Jan-19 | 470 | 417 | 471 | 417 | 292 | 417 |
| Feb-19 | 467 | 417 | 471 | 417 | 292 | 417 |
| Mar-19 | 467 | 417 | 471 | 417 | 291 | 417 |
| Apr-19 | 466 | 417 | 470 | 417 | 471 | 417 |
| May-19 | 467 | 417 | 467 | 417 | 471 | 417 |
| Jun-19 | 467 | 417 | 467 | 417 | 471 | 417 |
| Jul-19 | 466 | 417 | 466 | 417 | 471 | 417 |
| Aug-19 | | | 467 | 417 | 471 | 417 |
| Sep-19 | | | 467 | 417 | 470 | 417 |
| Oct-19 | | | 466 | 417 | 467 | 417 |
| Nov-19 | | | | | 467 | 417 |
| Dec-19 | | | | | 466 | 417 |

| | Randomised | | Interim survey | | Final survey | |
|--------------|-------------|-----|----------------|-------------|--------------|-------------|
| | WMCA | SCR | WMCA | SCR | WMCA | SCR |
| Jan-20 | | | | | 467 | 417 |
| Feb-20 | | | | | 467 | 417 |
| Mar-20 | | | | | 466 | 417 |
| Apr-20 | | | | | | 417 |
| May-20 | | | | | | 417 |
| Jun-20 | | | | | | 417 |
| Total | 6600 | | 6600 | 7506 | 6600 | 7506 |

Source: Evaluation consortium based on information from local sites 2017

Contact strategy

The contact strategy has been designed to maximise inclusion, engagement and response at each wave while maintaining ethical rigour around informed agreement to the research activities.

The contact strategy begins with the referral and agreement processes that ensure participants are fully informed about the trial and what participation involves, give agreement freely and understand how to withdraw agreement at a later date should they wish to do so (which will be via service providers).

Trial participants will provide a full range of contact details at baseline to enable contact via post, email, text and telephone.

The contact strategy for the interim and final surveys will involve the following steps:

- Advance letter sent by post to the trial participant inviting them to complete the survey online and providing the URL to access the survey. The letter will be motivational, accessible and tailored to the treatment/control status.
- For the participants who have indicated at baseline they need to complete the interview face-to-face or by telephone, the advance letter will be adapted to refer to the relevant mode.
- A series of email and text reminders will follow with embedded links to the online survey tool. Using different communication channels increases the chance of people responding.
- After two weeks, the fieldwork will move to mixed online and telephone for a further four weeks.
- Participants who have not yet completed (or partially completed) the survey will be contacted by telephone interviewers with the option of completing the survey by telephone. Phone calls will be made on different days of the week and at different times up to a maximum of eight attempts.

Between the interim and final survey, the sample will be contacted to retain engagement in the study. Change of address postcards will be distributed to encourage participants to notify NatCen of any changes. While there is a risk of losing contact with participants who move out of area, our strategies at recruitment and to maintain engagement will minimise this.

As with all studies involving fieldwork, NatCen will provide information about the study on its 'Taking Part' webpage (<http://natcen.ac.uk/taking-part/>). This will provide details about the study in line with the initial participant information sheet, including how data will be handled, the surveys

and contact details of how to decline to participate. Participants often check for web presence to be reassured about legitimacy.

For costing purposes, we have anticipated a minimum response rate of 60% at each wave although each survey will operate as a census with options to enable participation above this level. All trial participants will be invited to take part in the final survey, even if they did not respond to the interim survey.

We will monitor response closely in the early stages and adapt the contact strategy accordingly within the constraints of the HRA approval.

Mode

The interim and final surveys will be carried via a mixed-mode approach of telephone, web and face-to-face. It is anticipated that of the completed surveys, the majority will be completed by telephone, with smaller proportions online and around 5% via face-to-face survey. The mixed-mode approach provides flexibility to encourage engagement with the survey and adaptation to the needs of trial participants. The online survey can be completed in multiple sittings and in combination with telephone completion, enabling partially completed cases to become fully complete.

The approach will maximise efficiency in two main ways:

- Offering face-to-face surveys only for the participants for whom telephone or online completion would be a barrier. We will identify these cases at baseline.
- Leading with the online approach to avoid wasting resource on completing cases by telephone where they could have been completed online.

The questionnaire will be suitable for the different modes (so, for example, not relying on visual aids such as show cards) and aim to minimise mode effects.

The telephone and field interviewers will be fully briefed on the content of the questionnaire and the health needs of the participants but they will not be directly informed of the treatment status of participants when interviewing them to ensure blinding in data collection.

Questionnaire content

We expect that the final questionnaire will cover the following topics:

- Current employment status
- Job quality, satisfaction, match
- General self-efficacy
- Work search self-efficacy
- Work search activity
- General health
- Impact of health conditions
- Mental health, anxiety and depression
- Impact of health conditions on work productivity
- Wellbeing

- Use of services

We envisage that the final survey will be approximately 30 minutes when administered by telephone.

At this stage, we expect interim survey will comprise a subset of questions from the final survey, focusing less on employment-related questions and more on intermediate outcomes (such as self-efficacy and motivation) and of experiences and perceptions of services. The interim survey will include mental health and wellbeing measures to identify participants' status early on in the intervention and any risk of harm associated with the intervention. The interim survey is likely to be of approximately 20 minutes' duration when administered by telephone.

Translation and inclusion

The population in the two sites has considerable diversity in ethnicity but a very small minority (less than 1%) are likely not to be proficient in English. It is therefore unlikely to be proportionate and cost effective to translate survey materials, the randomisation tool, and materials in respect of agreement to participate in the trial into multiple languages.

Instead, we will address needs through interpreters, using NatCen field interviewers where possible and friends/family members as a last resort.

Cognitive testing

A selection of questions from the baseline assessment, and interim and final surveys were cognitively tested with adults with a mental health condition or MSK between 2nd and 15th January 2018. The purpose of the cognitive testing was to ensure that some of the key questions for measuring impact are easily understood, consistently interpreted and likely to result in reliable data from respondents. Standard proprietary questions were not in scope for the cognitive testing as these could not be altered.

An external agency recruited the 15 adults according to pre-specified quotas for in-work/out-of-work status and mental health condition/MSK while also achieving a range across age, gender and education level. All participants were recruited from London. An incentive was provided in the form of high street vouchers.

Interviews lasted around an hour, and included thorough probing as well as techniques such as 'think aloud' and observation, to ensure the questions were fit for purpose. These three cognitive techniques allow the researchers to identify questions that are at risk of being misunderstood or misinterpreted.

The researchers followed a protocol to ensure a consistent approach to testing the questions. Interviews were recorded with permission and the charted into a framework for synthesis and analysis.

Key findings and recommendations were collated:

- Suggestions were made as to wording changes to questions and response categories.
- A question capturing one of the primary outcomes, continuous employment, was tested in different formats and a simpler form adopted.

It was evident that a substantial minority of participants showed signs of frustration or distress in response to questions relating to employment status and barriers to employment. Although this was not surprising given the challenging circumstances of participants (and may be less of an issue for those who have actively opted in to the trial), it highlighted the need for measures to support participants and enable research participation to be a positive experience. On balance

and considering experience on other similar projects, we consider it reasonable to proceed with the following mitigation strategies:

- Thorough briefing of telephone and face to face survey interviewers (and training for employment advisers completing the baseline assessment) to detect signs of distress and manage appropriately, including pausing, rescheduling or terminating the interview and signposting to support.
- Explaining the purpose of the baseline assessment and surveys being to improve support for people with health conditions.
- Adding an introductory statement to the baseline assessment explaining the purpose and rationale for the questions being asked and providing a normalising statement to explain that people with health conditions often experience challenges in finding and staying in work, and explaining that the assessment aims to understand the individuals' experience.
- Explaining why questions on employment are being asked and normalising questions on employment and barriers with introductory text such as 'some people find...' to avoid people feeling singled out.
- Reading out response categories for some of the questions so that people are responding yes/no rather than explaining their own situation.
- Considering the feasibility of piloting the interim and final survey instruments in full (see next section).
- Launching the telephone and online modes simultaneously to ensure the majority of participants are engaged by an interviewer who can detect signs of distress.

The baseline assessment, and interim and final surveys are included in the document set for this resubmission (January 2018).

Survey Pilot

Purpose

In response to the findings of the cognitive testing, the NatCen team recommended carrying out a full pilot of the interim and final surveys. The main purpose of the pilot would be to check:

- Survey length: participants in the cognitive interviews took longer to answer some questions than we anticipated. A pilot would enable us to check the length of the interim and final surveys in full with participants who have chosen to take part in the trial.
- Response to the questions/levels of distress: between a half to a third of the cognitive testing participants responded negatively to questions about employment and barriers. A pilot would enable us to check the reaction to the surveys when they are carried out with introductions and explanations and in the context where people have opted to take part.

The pilot would also enable us to test the processes of accessing the baseline data and preparing the sample for the interim survey.

Scope

We propose to pilot the interim and final surveys with early recruits to the trial. We have costed on the basis that we invite the first 50 cases randomised from each site to take part in one combined pilot of the interim survey approximately two months after the baseline assessment and

randomisation. The intention is for participants to be evenly split between SCR and WM sites for two reasons: (1) so that we can identify any specific site-related difficulties and (2) so that sites are affected similarly by the removal of pilot cases from the final analysis. A key risk to this will be if one site launches the trial later than the other or if either site achieves fewer than 50 referrals in the first month. The constraint is that the pilot needs to be completed in time to allow for the agreement of post-pilot changes, HRA approval and programming changes. We will monitor early flows and consider an alternative strategy with WHU if necessary.

An indicative timetable is included below. Assuming that HRA approval for any post-pilot changes is received within a few weeks, it should be possible to launch the interim survey 17 weeks later (4 months post-randomisation). This would ensure that any additional cases from the first month (over and above the 100 included in the pilot) could be surveyed at the agreed point. There is also a buffer of some weeks in case it takes longer than 5 weeks to secure 100 cases.

| Date | Activity |
|-------------|---|
| Week 1 | Go live. Trials launched |
| Week 5 | Sample transferred from ONS to NatCen Programme signed off for pilot |
| Week 6 | Sample prepared for pilot Telephone interviewer briefing |
| Week 7 | Advance letters sent out |
| Weeks 8-9 | Pilot fieldwork |
| Week 10 | Pilot debrief and recommendations for programme changes agreed Revisions sent to HRA |
| Weeks 11-15 | HRA approval Programme changes implemented and tested |

We would invite the same 100 adults (regardless of whether they took part in the first pilot, but excluding any who have opted out) to take part in the pilot of the final survey eight or nine months after the baseline assessment. This would allow time to agree and implement any changes before launching the final survey for the first cases 12 months post randomisation. Our proposal is to aim for 50 completed interviews by a combination of telephone and online modes across a two week fieldwork period on each occasion. If the sample appears insufficient to achieve 50 interviews in the pilot of the final survey, we can supplement the sample with new cases.

We would monitor the length of survey and reaction to the questions and if necessary, recommend changes to the survey instrument to the consortium, WHU and HRA.

2.4 Information from national databases

The study intends to use information on intervention and control groups from national administrative databases to build up a fuller picture of their circumstances prior to starting on the trial, and the outcomes that they experience. These data will be used in the impact and economic analyses. Access to these data sources will be in compliance with the data access protocols required by national data owners.

Recruitment into the trial will close in Spring 2019. Someone randomised in March 2019 would potentially receive support from the trial for a further nine or 12 months after this date (depending on which trial they are part of). The length of time that is then required for national databases to be updated varies depending on the source. Data requests can be made at any time after March

2019, once baseline management information is available for all those randomised. However, it is necessary to wait until complete national administrative data records are available for all participants before the data extracts are drawn. We are considering drawing data at two points: to link with the completion of the interim survey in order that interim outcomes can be reported alongside the interim survey findings. Second, we will draw national data at an agreed point following trial closure. For example, to observe outcomes three months after the end of support from the trials for all participants (i.e. for the period up to June 2020) data extracts would need to be drawn at the following dates:

- **From NHS Digital.** Datasets should be updated around two months or less after June 2020 and so data extracts could be drawn sometime after August 2020.
- **From DWP.** Data on benefits receipt should be updated around four months after June 2020 and so data extracts could be drawn after October 2020.
- **From HMRC.** If real time information is available, this is likely to be subject to a two-month lag, and so data extracts could be drawn sometime after August 2020. The data on employment and earnings that HMRC have historically supplied to DWP has a six-month lag, and so if it was necessary to use this, data extracts would have to be drawn sometime after December 2020. There is a nine-month lag in the supply of data on self-employment to HMRC, so to obtain this it would be necessary to obtain data extracts sometime after March 2021.

2.5 Phase 3: analyses and reporting

2.5.1 Analysis of impacts on employment, health and wellbeing using survey data and data from national database

Full details of the statistical and data analysis plan are provided in section 3.

In line with the intention to treat principle (Shadish, Cook and Campbell, 2002), impact estimates will be based on the full sample of all those randomly assigned. Estimation will be carried out using linear regression of the outcome of interest on a variable indicating whether the individual was assigned to the treatment group rather than the control group¹⁸. Baseline (pre-randomisation) variables will also be included in the regression specification. The same approach will be used for continuous and discrete outcomes, with robust standard errors used to take account of heteroskedasticity.

Impacts will be based on both administrative data and survey data. For outcomes observed in administrative data, impacts will be based on the full sample of trial participants. For outcomes observed in survey data, impacts will be based on the sample of survey respondents. We will carry out a detailed analysis of non-response in order to establish whether there appears to be any difference across treatment arms that could influence impact estimates. Assuming there is no such difference, non-response weights will be calculated using baseline data and impact estimates for survey outcomes will be based on this weighted data. The degree to which impacts vary will be explored by carrying out additional analysis on subgroups of trial participants. These subgroups will be defined on the basis of characteristics at baseline (i.e. pre-randomisation).

2.5.2 Analysis of interview findings

The large volume of qualitative data collected through the process evaluation will be managed and analysed using the Framework approach. This involves the identification of key themes from the

¹⁸ Linear regression has been selected since results are more straightforward for non-statisticians to interpret

data to develop a thematic framework. This framework is then used to classify and organise the data from each respondent. The coded data for each theme is then reviewed in detail, drawing out the range of experiences and views across respondents. Emergent patterns are interrogated in relation to key characteristics of interest. In this study, the key hypothesis is that the IPS approach can lead to the client and system level outcomes described in the Theories of Change. We will therefore ensure that the analysis framework captures each element of the Theories of Change to explore and test the experiences and activities that lead to the key outcomes.

Analysis will take place on an iterative basis after each wave of fieldwork to allow us to meet interim and final reporting requirements, whilst also feeding into subsequent stages of the process study and refinement of the Theories of Change. On a quarterly basis, we will hold full-consortium analysis sessions, where key themes, patterns and issues emerging from the interviews are discussed and triangulated with other findings. This will allow regular review of the findings from the process study and present opportunity for the other evaluation teams to feed in to the interpretation of findings and shaping of subsequent waves of fieldwork. Representatives from the WHU will also be invited to these meetings.

Preliminary findings will also be shared with the trials through the Theory of Change refinement workshops mentioned earlier. This would facilitate the input of wider expertise into the interpretation of findings and help to fine tune the Theories of Change, supporting us to draw out key issues of interest and relevance.

2.5.3 Economic analysis

Data requirements

Analysis of the management information, survey data and linked administrative data will be used to assess the impact of the IPS service and determine the likely benefits. These sources will also be used to identify costs arising from the service. Impact estimates will be used to estimate the value of the benefits that the service produces and so it will only be possible to produce the economic evaluation once the impact analysis is complete. Where it is not possible to derive estimates of the direct costs or benefits of the service, the process study will be used to collect information on indicative costs. Therefore, the economic analysis will make use of the same data sources used in the other strands of the study, in pseudo-anonymised form, supplemented with aggregate data from publically available sources.

To date, the following costs and benefits have been identified, arising from the costs of administering IPS, as well as its potential impact on employment:

- Costs of delivering service
- Costs/savings from changes to use of benefits/tax credits¹⁹
- Costs/savings from changes in tax receipts
- Changes in output
- Costs/savings from changes in use of healthcare services
- Increased use of transport

¹⁹ If IPS type services were to be rolled out nationally for individuals who are not working due to health conditions, this would be in the context of Universal Credit (rather than the legacy benefits e.g. Jobseeker's Allowance and Employment Support Allowance). As such, we will look to take account of the change to Universal Credit within the economic analysis.

- Increased demand for childcare
- Changes in eligibility for free school meals
- Changes in the use of statutory sick pay

The following section outlines the process of calculating the return on investment.

Methods

Identifying direct and indirect costs and benefits

The approach to calculating the return on investment from the IPS service will be based on that set out in the HMT Green Book. We anticipate that the estimates will be made for the return on investment (ROI) of the service, rather than of the trial.

The first step will be to identify all the likely costs and benefits to the exchequer and society of the service, even if it is likely to be difficult to attach a value to all components. Both direct and indirect costs and benefits will be considered. There is a possibility that other costs and benefits are identified during the course of the evaluation (perhaps as a result of the process study). It is not realistic to foresee all potential costs and benefits of the service before the trials commence and so it will be necessary to keep the expected costs and benefits under review and add to the list as the trials develop.

In assessing the impact of the trials it is important to consider how they might differ for different individuals. The impact analysis will include a sub-group analysis which will be used to identify how the costs and benefits of the trials vary for different groups of participants. If sample sizes are sufficient and differences in impact are identified, the economic analysis will consider how the characteristics of those affected shape the likely return on investment.

Valuing costs and benefits

Wherever possible, values will be based on market prices, but the HMT Green Book sets out a number of approaches that can be used where prices cannot be easily observed. These will form a starting point in exploring alternative options, if necessary.

The process study will also be used to collect indicative information on the market value of costs or benefits from the service. Another potential source is information from past studies. In these cases, the approach will be to produce both a central valuation estimate, as well as minimum and maximum estimates to explore how sensitive the return on investment is to different assumptions about how costs and benefits are valued.

As the costs and benefits of the trials will accrue over a matter of years, it is necessary to take into account both inflation over this period, and the present value of costs and benefits that may be incurred at some point in the future. To reflect a known preference to receive benefits now, rather than in the future, and to defer costs until later, the analysis will take into account 'discounting'. The Green Book recommends that a discount rate of 3.5 per cent is used, so something that was worth £100 in 2016 will be worth £103.5 at the end of 2017. All costs and benefits will be adjusted by the discount rate to calculate net present value i.e. the difference between costs and benefits, valued in today's prices.

Values will be considered in real terms, taking into account price inflation over time.²⁰ The possibility that certain components of costs and benefits experience atypical inflation over time will

²⁰ Where possible, the analysis will adjust prices by the most appropriate measure of inflation for that particular component, for example, allowing for differences in wage and price inflation.

also be assessed. For example, it is possible that Britain's forthcoming exit from the EU reduces the supply of labour and generates increased wage inflation.

Estimating the return on investment

Having calculated the costs and benefits per trial participant and the number of participants, an estimate of the overall return on investment from the IPS service will be produced. The costs associated with running the randomised control trial will not be included in this estimate as the aim is to determine the likely return on investment from the service itself. The evaluation team will work with WHU and trial sites to seek to separate out the costs of running the trials from the costs of running the service. Where it is impossible to find a defensible way of attaching monetary values to particular costs or benefits, this will be documented and the report will indicate the likely impact of these on the overall conclusions from the economic impact of the service.

The ability to observe outcomes from the service and actual costs makes it possible to produce a more accurate estimate of the return on investment than is possible before implementation. However, it is likely that there will still be a degree of uncertainty over some aspects of the model. Therefore, the analysis will show how the return on investment might vary with adjustments to the assumptions underlying the model.

Comparing trial participants with the eligible population

The economic analysis will then move on to explore differences between the characteristics of sites participating in the trials and areas more generally. This will involve comparing the characteristics of trial participants with those of the population of individuals who would be eligible for the service if it was extended nationwide. There will be greater uncertainty in this part of the analysis due to the fact that it will be based on projections about what the likely impact of the trials would be if they were extended more widely.

As the trials are only running in two areas, there may be substantial differences in characteristics between the trial sites and areas which might potentially be subject to the intervention in the future. These differences in characteristics might result in differences between the estimated return on investment calculated in the study and that which would be expected if the trials were carried out in a different set of areas.

Risks to the achievement of the expected return on investment across a wider range of areas will be considered and the potential impact of these risks assessed. For example, pilot schemes sometimes receive a higher level of resourcing than is available to other areas when roll-out of a policy is extended. This can mean that the impact of pilots can be greater than subsequent impacts when a policy is expanded to other areas. A sensitivity analysis will be used to assess the likely impact of varying the assumptions which underlie the analysis. This will consider by how much any benefits produced by the service in the trial sites would need to fall before the return on investment disappeared, or likewise, the increase in costs which would cancel out any return. This makes it possible to assess the likelihood that, in practice, the expected return on investment would not materialise. The sensitivity testing is likely to include a Monte Carlo analysis to assess the likely impact of uncertainty over multiple factors. For each factor, there may be an expected cost, as well as likely maximum and minimum values. A Monte Carlo analysis can be used to estimate the most likely total cost and the probability that costs are above or below particular values.

Timing

The ratio of costs to benefits from the service will vary depending on the period of time considered in the analysis²¹. Ideally, costs and benefits should be considered over the full period that they are

²¹ This will be linked to the agreements on the retention of national and other data from the evaluation – currently three years and indefinitely for anonymised data.

likely to accrue, but constraints on the amount of time available for the evaluation mean that it is only possible to consider the return on investment over a more limited period.

It is likely that few benefits will be realised in the period immediately after the start of the trials. Also, the fixed costs associated with commencing delivery (such as staff training) may mean that costs in the early stages of the service may be relatively high. The benefits of the service are likely to develop over a longer time-period, whilst total costs may fall as the intervention becomes established. For example, if the service is effective in assisting someone to move from benefits to employment, this may not result in a net benefit to the exchequer initially if the impact is to shift the individual from benefits to tax credits. However, over a longer period of time, the individual may progress in work and attain higher earnings, reducing entitlement to tax credits and increasing savings for the exchequer. Thus the return on investment is likely to vary, depending on the time-period considered. Whilst a longer-term study would be required to address this issue fully, the economic analysis will seek to differentiate between fixed and variable costs to give some indication of how at least some components of the return on investment might evolve over time. It will not be possible to estimate longer-term benefits from the service over the lifetime of the current evaluation however.

2.5.4 Reporting and dissemination

The reporting of interim and final findings will need to balance the need to allow sufficient time for data to be collected and analysed with the pressure to monitor early flows, drop-outs etc. and the need to align with wider policy imperatives (such as internal WHU reporting requirements and the comprehensive spending review (CSR) timetable. The evaluation plan has a reporting timetable which takes these factors into account, but this will need to be kept under review, in consultation with WHU, to ensure it can be adapted should circumstances change.

The main reporting stages²² are set out below:

- **Early findings on referral flows, allocations and the data dashboard:** we anticipate that, in the early weeks of the trial, weekly reports will be required to help monitor referrals, agreement to participate, randomisation and assignment to arms of the trial. This will allow an early assessment to be made of the mechanics of the trial and for corrective action to be taken as quickly as possible. In addition, we anticipate that the Unit will require a weekly update on referrals as well as a monthly summary of early MI data (contained in the data dashboard reporting mechanisms);
- **Early findings from the qualitative research:** we recognise that some form of interim reporting on the trial may be needed to inform any policy submissions ahead of interim quantitative findings being available. This will be based on a 'snap shot' of available data in Q2 of 2018;
- **Interim survey findings:** a report of interim survey findings (and accompanying presentation) will be available in Q3 of 2018. This will include an analysis of intakes, the balance & drop-out rates of participants in each arm of the trial and some early process evaluation findings from each site);
- **Follow-up survey findings:** a report (and presentation) of survey findings to Q3 in 2019 (including the full interim survey sample and the final sample, and administrative data to this date) will be produced in Q1 of 2020;
- **Trial Reports:** a report (and presentation) for each local area will be produced in Q4 of 2020;

²² Not all these reports are expected to be published and many are intended for internal use only

- **Synthesis:** a synthesis report, including results from the trial and including data on impact, process and economic evaluation will be produced, together with a presentation, in Q1 of 2021. This will unify findings in respect of the key research questions although segment them based on the differing target groups, and operational designs, in each area. It will conclude with the identification of implications and issues for any future roll-out. This report will include in the appendices an update to the literature review conducted during the design phase, which will include findings from the RISE study into IPS along with other more recent evidence on the IPS intervention. These data will be used to contextualise the findings from the evaluation.

Other opportunities for dissemination of information about, and findings from, the trial are likely to include:

- publication of trial protocols;
- articles reporting on quantitative and qualitative findings from the trial in peer-reviewed journals;
- papers presented to academic, policy or practitioner conferences and seminars;
- delivering local practitioner workshops for healthcare professionals, employers and JCP employees.

The evaluation team will liaise with the WHU over any concrete plans to progress these or other dissemination options.

3 Statistics and data analysis

3.1 Sample size

West Midlands Combined Authority (WMCA)

The sample size was chosen to be the largest possible, subject to the practical constraints around local delivery capacity and budget (IPS support has a high average cost). A large sample size was desirable in order to maximise the statistical power of the trial, which is particularly important when considering multiple outcomes and subgroups and the unknown magnitude of effect the intervention might have (if any).

Power calculations for the predicted sample size are presented below for the case of a binary employment outcome. The likely size of effect (difference between treatment and control groups) is uncertain. We assume a 'business as usual' (BAU) employment outcome of 15% for the control group. This means that in the control group, we expect around 15% of individuals to achieve employment. This assumption is based on the latest Work Programme Statistics²³ (with data up to December 2016). These data show that the proportion of those starting a benefit claim on the grounds of sickness or disability who, within a year of referral to the programme, had entered work and remained employed for three months averaged 8.4% over the 2015 intake (nationally) for those with 12-month prognosis complaints and 16.1% for those without a 12-month prognosis complaint. There is also the question of what the average outcome of the treated group is likely to be. Perhaps the best clue here is the Working Well pilot in Greater Manchester²⁴ which is on course to achieve its target of 20% of participants starting work.²⁵ Combined, a reasonable and prudent approach might be to assume for design purposes that IPS will increase labour market outcomes among those out of work from 15% (the BAU case) to 20%, a difference of five percentage points.

Estimates of statistical power reflecting these assumptions – where power is the probability of detecting an effect where one exists - are summarised in **Error! Reference source not found..** These are based on simulations performed using the statistical package R. Impacts were estimated using linear regression of the outcome on a treatment group indicator, assuming an R-squared of 0.1 (this assumption for R-squared is based on results from a previous trial²⁶). The table

²³ The Work Programme, since June 2011, has been the Government's main welfare-to-work scheme. Referrals ended in April 2017.

²⁴ The Working Well pilot was introduced in March 2014 with the aim of supporting 5,000 clients, who have completed the Work Programme but not found work. The focus is on those claiming sickness and disability benefits. The pilot addresses specific barriers to work and offers up to one year of in-work support.

²⁵ http://www.greatermanchester-ca.gov.uk/download/downloads/id/215/working_well_annual_report_2016.pdf

²⁶

<http://webarchive.nationalarchives.gov.uk/20130314010347/http://research.dwp.gov.uk/asd/asd5/rports2011-2012/rrep765.pdf>

presents estimates of power (1 minus the probability of type 2 error) and statistical significance (type 1 error)²⁷:

- Column (3) relates to the expected total number of trial participants. It indicates that estimates based on the full sample have very high statistical power (1.00, to two decimal places) to detect the assumed impact of a 5 percentage point difference. This is relevant to the case of an outcome that is recorded in administrative data and therefore available for all trial participants.
- Column (2) relates to the expected total number of trial participants who respond to the survey. It indicates that estimates based on the sub-sample of survey respondents (assumed to be 65% of the full sample) have very high statistical power (0.99) to detect the assumed impact of 5 percentage points.
- Column (1) shows the smallest sample size that can deliver adequately-powered estimates of the assumed impact of 5 percentage points. A sample size of 1,350 delivers power of 0.79, just short of the conventional 0.80 target. This suggests that the trial is adequately-powered for subgroups that accounts for fewer about one quarter of the full sample or two-fifths of the survey sample.

Table 3.1: Power calculations WMCA

| | (1) | (2) | (3) |
|--------------|-------|-------|-------|
| Power | 0.79 | 0.99 | 1.00 |
| type 1 error | 0.05 | 0.05 | 0.05 |
| N | 1,350 | 3,500 | 5,300 |

The results in this table are based on an assumed impact of +5 percentage points, relative to the 15% BAU base. Impacts were estimated using linear regression of the outcome on a treatment group indicator, assuming R-squared of 0.1. The results are based on 1,000 simulations.

Source: Evaluation consortium based on information from local sites 2017

Overall, the trial will be well powered to detect the anticipated impact for the full sample or survey sub-sample. There are four reasons why it is particularly helpful for the trial to be highly-powered to detect overall impacts. First, there is strong policy interest in how impacts vary across the population; the results above indicate the size of subgroup that can be considered (i.e. $n=1,500$). Second, while the power considerations relate to the case of a single outcome, the trial will examine multiple outcomes and this will need to be taken into account during estimation, reducing the power levels for individual outcomes²⁸. Third, existing evidence relates to those with more severe conditions so we cannot be sure how strong the impacts for those with mild/moderate conditions will be. Fourth, public policy trials often suffer with low participation numbers; aiming for a large sample offers some protection against this so that even with lower than expected numbers there may still be sufficient power to achieve a robust impact estimate.

Sheffield City Region (SCR)

The sample size was chosen to be the largest possible, subject to the practical constraints around local delivery capacity and budget (the IPS support has a high average cost). A large sample size was desirable in order to maximise the statistical power of the trial, particularly important when

²⁷ A type 1 error is the probability of wrongly rejecting a true null hypothesis; type 2 error is the probability of wrongly not rejecting a false null hypothesis

²⁸ In short, each additional outcome tested means adjusting the threshold for statistical significance.

considering multiple outcomes and subgroups. Furthermore, the SCR sample involves participants who are in work at baseline as well as participants who are out of work at baseline and there is interest in understanding how the intervention operates for both groups. The calculations in this section are based on 30% of the trial participants being in work at baseline. As mentioned earlier, this proportion will be controlled by adjusting the randomisation ratios.

Power calculations for the predicted sample size are presented below for the case of a binary employment outcome. The likely size of effect is uncertain. We assume a BAU employment outcome of 15%. This is based on the latest Work Programme Statistics (with data up to December 2016) which show that the proportion of new ESA claimants who, within a year of referral, had entered work and remained employed for three months averaged 8.4% over the 2015 intake (nationally) for those with 12-month prognosis complaints and 16.1% for those without a 12-month prognosis complaint. There is also the question of what the average outcome of the treated group is likely to be. Perhaps the best clue here is the Working Well pilot in Greater Manchester which is on course to achieve its target of 20%.²⁹ Combined, a reasonable and prudent approach might be to assume for design purposes that IPS will increase labour market outcomes among those out of work at baseline from 15% (the BAU case) to 20%. For the in-work group, there is less evidence available regarding expected BAU outcome levels. Conservatively, we assume a BAU outcome of 50% and that the assumed 5 percentage point impact of IPS for the out of work groups will hold also for the in-work group.

Estimates of power reflecting these assumptions are summarised in **Error! Reference source not found..** These are based on simulations performed using the statistical package R. Impacts were estimated using linear regression of the outcome on a treatment group indicator, assuming an R-squared of 0.1 (this assumption for R-squared is based on results from a previous trial). **Error! Reference source not found.** presents' estimates of power (1 minus the probability of type 2 error) and statistical significance (type 1 error):

- Column (3) relates to the expected total number of trial participants and so is relevant to the case of an outcome that is recorded in administrative data. It indicates that estimates based on the full sample of participants who are out of work at baseline have very high statistical power (1.00, to two decimal places) to detect the assumed impact of 5 percentage points (top panel). For those in work at baseline (middle panel), power is estimated to be quite close, at 73%, to the conventional target of 80%. The pooled sample (bottom panel) again shows very high power.
- Column (2) relates to the expected total number of trial participants who respond to the survey. It indicates that estimates based on the sub-sample of survey respondents (assumed to be 65% of the full sample) who are out of work at baseline have very high statistical power (0.99) to detect the assumed impact of 5 percentage points (top panel). Estimates based on those in work at baseline appear under-powered at 57% (middle panel). For the pooled sample or respondents, power is again high (100%, rounded - bottom panel).
- Column (1) shows the smallest sample size that can deliver adequately-powered estimates of the assumed impact of 5 percentage points. For the OOW group, the smallest adequately-powered subgroup is of size 1,300 (one quarter of all OOW participants). Subgroup analysis for the IW group is not adequately powered. For the pooled sample, the smallest viable subgroup is of size 1,750 (again about a quarter of all participants).

²⁹ http://www.greatermanchester-ca.gov.uk/download/downloads/id/215/working_well_annual_report_2016.pdf

Table 3.2: Power calculations

| | (1) | (2) | (3) |
|--|-------|-------|-------|
| Out of work at baseline: BAU = 0.15 | | | |
| power | 0.81 | 0.99 | 1.00 |
| type 1 error (1 outcome) | 0.05 | 0.05 | 0.05 |
| N | 1,300 | 3,420 | 5,250 |
| In work at baseline: BAU = 0.50 | | | |
| power | 0.80 | 0.57 | 0.73 |
| type 1 error (1 outcome) | 0.05 | 0.04 | 0.04 |
| N | 2,700 | 1,460 | 2,250 |
| Pooled sample: 70% out of work (BAU = 0.15) 30% in work (BAU=0.50) at baseline | | | |
| power | 0.81 | 1.00 | 1.00 |
| type 1 error (1 outcome) | 0.05 | 0.05 | 0.05 |
| N | 1,750 | 4,800 | 7,500 |

The results in this table are based on an assumed impact of +5 percentage points, relative to the BAU bases set out in the three panels of the table. Impacts were estimated using linear regression of the outcome on a treatment group indicator and an indicator of whether IW or OOW. The results in the table were based on 1,000 simulations.

Source: Evaluation consortium based on information from local sites 2017

Overall, the trial will be well powered to detect the anticipated impact for the out of work sample or the pooled sample. Estimates based on the in-work group as a whole come close to the target level of power but are under-powered when based on survey respondents only (and will be similarly limited for sub-groups). There are four reasons why it is particularly helpful for the trial to be highly-powered to detect overall impacts. . First, there is strong policy interest in how impacts vary across the population; the results above indicate the size of subgroup that can be considered. Second, while the power considerations relate to the case of a single outcome, the trial will examine multiple outcomes and this will need to be taken into account during estimation, reducing the power levels for individual outcomes. Third, existing evidence relates to those with more severe conditions so we cannot be sure how strong the impacts for those with mild/moderate conditions will be. Fourth, public policy trials often suffer with low participation numbers; aiming for a large sample offers some protection against this so that even with lower than expected numbers there may still be sufficient power to achieve a robust impact estimate.

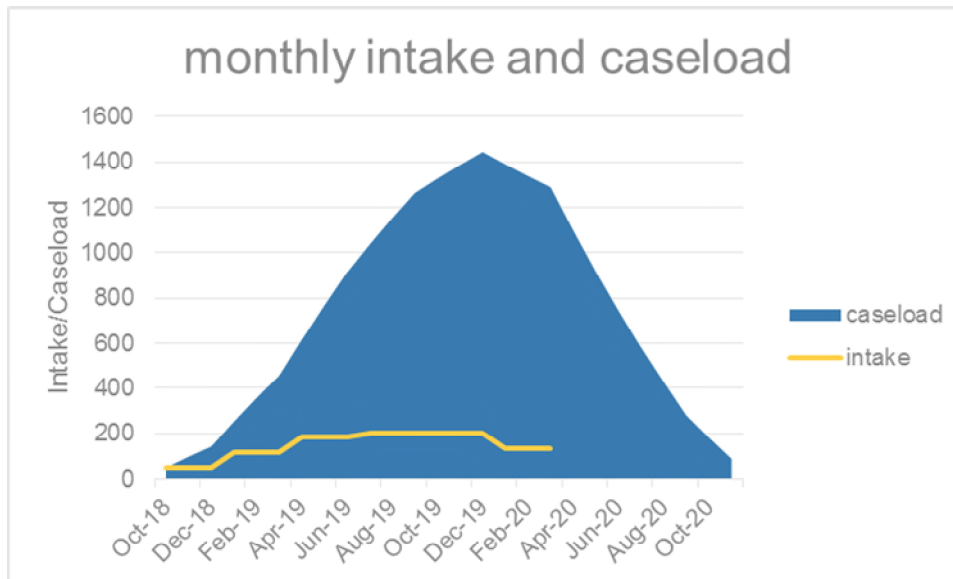
3.2 Planned recruitment rate

West Midlands Combined Authority (WMCA)

Figure 3.1 shows how trial numbers are projected to vary over the course of the delivery period in WMCA. These are estimates; it is a challenge to be precise since the numbers potentially eligible are not known and the voluntary nature of trial participation introduces a further uncertainty. Drawing on benefit records, the caseload is expected to steadily build to peak at the end of 2019 at

about 1,450. Intake stops after March 2019 and the last individual from the trial will complete treatment in November 2020.

Figure 3.1: Illustration of monthly caseload using WMCA estimates of onflow

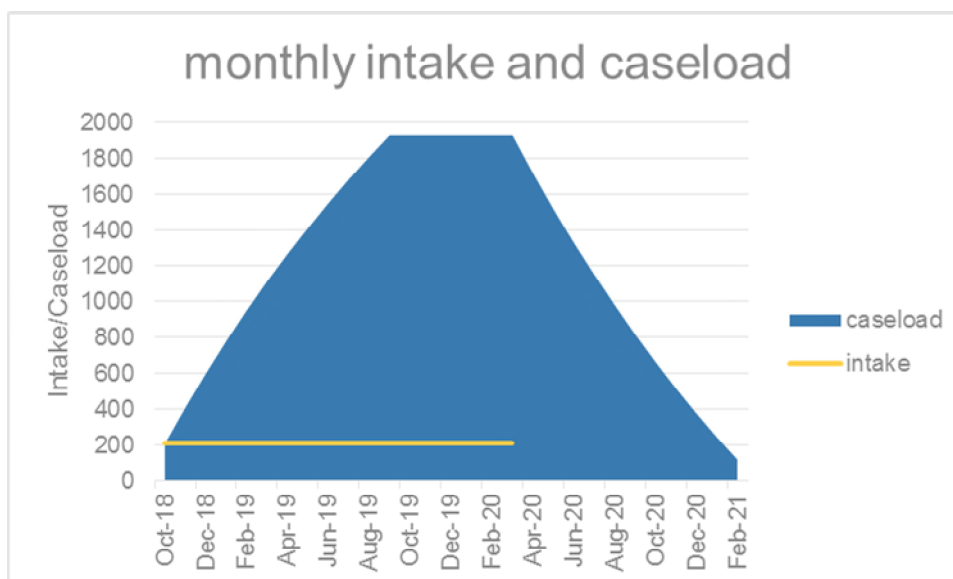


Source: Evaluation consortium based on information from local sites 2017

Sheffield City Region (SCR)

Figure 3.2 shows how trial numbers are projected to vary over the course of the delivery period on SCR. These are estimates; it is a challenge to be precise since the numbers potentially eligible are not known and the voluntary nature of trial participation introduces a further uncertainty. Monthly onflow is estimated to be 210 and a monthly attrition rate of 5% is assumed. Caseload builds steadily, plateaus at 1,930 and then declines.

Figure 3.2: Illustration of monthly caseload using SCR estimates of onflow



Source: Evaluation consortium based on information from local sites 2017

3.3 Statistical analysis plan

3.3.1 Summary of baseline data and flow of patients

During the course of the trial, baseline characteristics will be compared across treatment and control. For continuous variables, means, medians and standard deviations will be reported by treatment arm; for categorical variables, distributions will be compared; for binary variables proportions will be compared. In line with CONSORT guidelines³⁰, we will not carry out significance tests of treatment-control differences since such differences will be random by virtue of the design of the trial³¹. Instead we will examine standardised differences for any suggestion that there may be a possible problem with the randomisation algorithm. The response to this would depend on the nature of the difference found but would prompt an examination of the randomisation computer code.

Monitoring reports presenting this information will be produced weekly in the first six weeks of the trial, then fortnightly and then, three months after trial start, monthly. A CONSORT flow diagram will be produced that tracks dropout and attrition.

3.3.2 Primary outcome analysis

For the primary outcome, treatment and control group means and measures of spread (e.g. variance) will be reported. Impacts will be estimated using linear regression of the outcome on a variable indicating whether the individual was assigned to the treatment group or the control group. Baseline (pre-randomisation) variables will also be included in the regression specification.

We will adhere to the intention to treat (ITT) principle. In line with this, impact estimates will be based on the full sample of all those randomly assigned. Impacts will be based on both administrative data and survey data. For outcomes observed in administrative data, impacts will be based on the full sample of trial participants. For outcomes observed in survey data, impacts will be based on the sample of survey respondents. We will carry out a detailed analysis of non-response in order to establish whether there appears to be any difference across treatment arms that could influence impact estimates. Assuming there is no such difference, non-response weights will be calculated using baseline data and impact estimates for survey outcomes will be based on this weighted data.

In view of the nature of the intervention, there are two primary outcomes of equal interest, one covering the employment domain and the other covering the health domain. This raises a multiple testing issue, whereby the chances of spuriously finding a significant effect increases. We will adjust critical p-values to take account of the multiple testing issue (e.g. using the Benjamini-Hochberg approach).

Non-compliance: The ITT impact estimate is based on assignment outcome rather than realised participation. Hence, individuals not adhering to their randomisation outcome, for instance, those assigned to IPS not showing up – does not alter the approach taken to ITT estimation. However, we will examine and report on the degree of non-compliance. In view of the BAU nature of the control condition, only one-sided compliance is likely (control group members will have access restricted to the IPS support available under this trial, meaning that it is non-compliance with IPS for those allocated to it that we will be more concerned with). The proportion of individuals not receiving the IPS treatment will be reported. We will estimate the Complier Average Causal Effect

³⁰ <https://www.consort-statement.org/>

³¹ For example, with 20 variables assessed, we might expect on (5%) to be significantly different by chance alone. With such large sample sizes, even substantively small differences could be 'significant' hence focusing on standardised differences rather than significance testing.

(CACE) using instrumental variable regression, where the randomisation outcome is the instrument.

3.3.3 Secondary outcome analysis

Impacts on secondary outcomes will be estimated in the same way as the primary outcomes. Again, critical p-values will reflect the multiple-testing problem. However, the greater number of secondary outcomes may reduce result in a threshold adjustment that reduces the statistical power of individual tests to the extent that potentially important effects are missed³². In short, with more outcomes to test, we have to make it more difficult to find statistically significant results, because with more testing, we increase the probability of finding chance differences. Consequently, the analysis of secondary outcomes is more exploratory; perhaps generating hypotheses rather than providing firm conclusions.

3.3.4 Subgroup analyses

The degree to which impacts vary will be explored by carrying out additional subgroup analysis. These subgroups will be defined on the basis of baseline characteristics. We anticipate the subgroups listed below are likely to be of interest in respect of policy and practice, although any decisions have yet to be finalised. In addition, decisions on the final analysis strategy may be empirically driven.

- Referral route: eg GP referral, self-referral, other
- Personal characteristics: e.g. age, gender, ethnicity
- Skills and labour market: e.g. highest qualification, prior labour market experience, duration of unemployment/inactivity
- Benefits receipt: e.g. Jobseeker's Allowance, Employment and Support Allowance, Universal Credit
- Health condition category: e.g. musculoskeletal, mental health, others
- In- or out-of-work at commencement of the trial (SCR only)

The analysis will be carried out by repeating the estimation approach used for primary outcomes but now interacting with a variable identifying subgroup membership with the treatment variable. As with the primary outcomes, we will adjust p-values to take account of multiple outcome testing.

3.3.5 Adjusted analysis

The main impact estimates will be based on the results of linear regression analysis, controlling for baseline characteristics. These baseline characteristics are intended to increase the precision of estimates. It is expected that this will include baseline employment and health measures, summary measures of employment and health histories and personal characteristics (such as age, sex, ethnic group, partnership status, dependent children, education). As a sensitivity check, we will compare these main results against unadjusted treatment-control comparisons.

³² Using a simple approach - see Dunn (1961) on the application of Bonferroni's work to means and confidence intervals - with two outcomes tested, the threshold for 'significance' would be divided by two, meaning the 'critical threshold' was 0.025 rather than 0.05. For two tailed tests, this would mean 0.0125 at either end of a normal distribution.

3.3.6 Interim analysis and criteria for the premature termination of the trial

An interim analysis will be carried out based on survey data collected four months post-randomisation. In addition, it is intended that the analysis includes outcomes recorded in national administrative data-sets. This will explore the early effects of the treatment for a subset of the survey outcomes measured in the final survey. The statistical approach will be the same as that used for the analysis in the final report. In addition, it will report on the services used by both the treatment and control group members.

3.3.7 Procedure(s) to account for missing data

In principle, the analysis based on administrative data should be unaffected by missing data. In practice, it may be that some trial participants cannot be matched to their administrative records, perhaps because the identifying variables required for linking (NHS number and national insurance number) were wrongly recorded in the randomisation tool. In such a scenario, we would expect there to be no systematic differences across treatment arms and therefore that estimating results on the subsample of individuals who can be matched will not introduce any bias.

The problem of missing data is greater when considering survey data, mainly due to unit non-response. Non-response that is independent of treatment status does not invalidate the trial; treatment control comparisons still capture the impact of the treatment. However, these impacts may relate only to the respondent population rather than the trial population. In order to address this, non-response weights will be calculated using baseline data and impact estimates for survey outcomes will be based on this weighted data. A more difficult problem arises when there is a difference across treatment and control groups in the probability of survey response. It may still be the case that treatment-control comparisons provide valid estimates of causal impacts, but this relies on the difference in non-response being uncorrelated with outcomes. Should that not be the case, the ability of the trial to provide unbiased estimates is reduced. Dealing with such a “Missing Not At Random” (MNAR) problem relies on using non-experimental techniques which themselves are not guaranteed to be effective. Consequently, should MNAR be a concern, we will report the survey-based estimates with an appropriate caveat about their robustness. Our main strategy for judging this will be to use outcomes reported in administrative data in order to compare full sample and respondent sample impact estimates.

With item non-response, missing values will be imputed. However, we will not impute values for outcome measures where they are missing.

4 Trial setting

The locations listed in the table below include the NHS Trusts and CCGs in which Employment Specialists will administer the agreement to participate process and meet with trial participants. In including this information, our main aim is to identify the settings in which the research will take place. This is noted in the bullet points beneath the table.

Table 4.1: Locations for trial delivery

| | WMCA | SCR |
|---|--|---|
| Areas of the Regions | Sandwell & West Birmingham Birmingham & South Central Dudley & Wolverhampton. | Barnsley, Bassetlaw, Doncaster, Rotherham and Sheffield |
| Primary and Community care settings | <ul style="list-style-type: none"> • The Dudley Group NHS Trust • Dudley and Walsall Mental Health Trust • Black Country Partnership NHS Foundation Trust • Heart of England NHS Foundation Trust • University Hospitals Birmingham NHS Foundation Trust • Birmingham and Solihull Mental Health NHS Foundation Trust • Birmingham Community Healthcare NHS Foundation Trust • The Royal Wolverhampton NHS Trust (Birmingham Healthy Minds) • Wolverhampton CCG • Dudley CCG • Sandwell & West Birmingham CCG • Birmingham South Central CCG | <ul style="list-style-type: none"> • Sheffield CCG • Doncaster CCG • Rotherham CCG • Barnsley CCG • Bassetlaw CCG • Sheffield Health and Social Care NHS Foundation Trust • Rotherham Doncaster and South Humber NHS Foundation Trust • South West Yorkshire Partnership NHS Foundation Trust • Nottinghamshire Healthcare NHS Foundation Trust • Sheffield Teaching Hospitals NHS Foundation Trust • Rotherham Hospital NHS Foundation Trust • Barnsley Hospital NHS Foundation Trust • Bassetlaw Hospital NHS Foundation Trust |

Source: Drawn from local sites' intervention protocols, 2017

44 - Evaluation of the health-led employment trials

- **Survey administration and service user interviews** – this will usually be by phone. In a few instances the survey will be administered face-to-face in service user's homes, or exceptionally, in a quiet, private room in one of the Primary or Community Care settings listed above.
- **Interviews with practitioners** – these will usually be by phone or in a private room in one of the Primary or Community Care settings listed above.

5 Agreement to participate in the trial

The trials will operate on the basis of written, informed and voluntary agreement of service users to being part of the trial (and thus to the possibility of being either in the control or intervention group). This process will be led by local sites and further detail is contained in their trial protocols.

As part of this process, agreement will be sought of service users to having their data shared with the Consortium and the organisations owning data sets to be used in the evaluation for research purposes.

Once they have agreed to take part in the trial and the evaluation, participants can withdraw their agreement at any point, and they will then not be contacted again.

It is vital that agreement to participate is informed, given freely and can be withdrawn (in respect of primary research activities such as interviews and surveys) if the service user changes their mind.

Process to seek agreement for trial participation – led by local sites

When an individual is referred to either the service in the WMCA or SCR, they will have an initial face-to-face meeting with an Employment Specialist. The Specialist will determine if the individual meets the inclusion criteria, and if so, will invite them to be part of the trial and undertake the process of asking for informed agreement to participate. The administration of the agreement process will take place in a quiet and private space.

It is vital that agreement is informed, given freely and that potential participants are aware that their informed agreement can be withdrawn at any point without consequence (including a continued ability to receive the service) if the service user changes their mind (and that this is recorded on the service database). Such a withdrawal will mean they receive no further contact from the research team and/or employment specialist depending on if they wish to withdraw from receipt of the IPS support.

It will be made clear to service users that if they do not agree to be part of the trial, they will be eligible for the range of local support already available for their health needs and to find employment and the Employment Specialist will provide information about these services (details of these services form part of the trial protocols in each local area). The Employment Specialist will clearly explain that if they do agree to be in the trial, it is possible that they will either receive the research trial service or services as usual, and that, in relation to the research trial service, it is not known whether or not it is more effective, less effective or the same in a primary care setting for the target group, as the existing services.

Time for consideration

WMCA

In WMCA Service users will have a 5-minute window to reflect on their inclusion into the trial and they will be explicitly asked if they would like to ring anyone with whom they would like to discuss the trial.

SCR

In SCR service users will be asked to make a decision during the initial meeting once the trial and the agreement to participate has been fully explained. They may also have received a leaflet about the service from the referral source. It is therefore considered appropriate to allow the individual to take the decision at this point whether they wish to continue to voluntarily participate in the trial, whilst noting that individuals are not required to make a decision at this point and can go away and take more time if they wish, and reminding that they can withdraw from the trial support and research activities (but not processing of data already held) without consequence at any point.

Signing agreement materials

Where agreement is obtained, the service user will be asked to sign the agreement form and this will be logged as completed using a code in the management information database used by the service. Furthermore, all agreement forms will be stored with the service user's employment case records.

It will be possible for a representative to sign the form of the service user is not physically able to do so.

Withdrawing agreement

Once agreement has been given, service users will be able to opt out of the trial support and contact with the research activities such as interviews and surveys at any time by telling the Employment Specialist or by using the opt-out form provided to them at the point at which they give their agreement.

Training and guidance for Employment Specialists

The consortium will provide clear written guidance will be provided to Employment Specialists about how to administer the agreement process.

5.1 Agreement materials

In recognition of the importance that the trial and the service is equally accessible for a wide range of service users, special consideration has been given to the design and language of the Participant Information Sheets (PIS) and agreement forms. Two versions of each have been developed; accessible (easy-to-read) and plain English formats. The accessible versions have been worded to ensure the comprehension and comfort of those with literacy difficulties. These have been updated in the latest submission to accommodate the role of the Office for National Statistics (ONS) which is providing the safe haven for the research data. Local site trial branding has been added to these documents.

The development of these materials has sought guidance and input from service users in a user testing workshop in June 2017. They have also been shared with DWP, NHS Digital and HMRC – the owners of the national datasets which the evaluation hopes to access. They have also been designed in collaboration with the WMCA and SCR.

In both trial sites careful consideration has been given as to whether agreement materials need to be translated into languages other than English.

- The WMCA are seeking feedback as part of their planned of GP and primary care engagement sessions on whether specific languages need to be considered beyond English in specific regions. If this is the case then they will ensure that relevant leaflets and agreement materials are translated.

- In the SCR the starting point is that it is important that all residents are able to participate in the trial, regardless of their language. Analyses conducted by SCR suggest minimal translation needs, but has identified the main non-English languages. Employment Specialists will ask at referral if they need translation, and if the needed, SCR will translate of trial participation/service related materials.

5.2 Agreement to survey and interview

The informed agreement obtained during the first meeting with an Employment Specialist, at the point of sign-up to the trial (see Trial Protocol), will cover agreement for the respondent to be contacted again to be invited to participate in the surveys/ interviews.

- **Surveys:** an advance letter, which will be sent to respondents ahead of the surveys start date, will explain that the survey is voluntary. Finally, verbal agreement will be obtained by NatCen interviewers prior to conducting the telephone or face-to-face survey.
- **Interviews:** an advance letter will also be sent to all clients sampled for the in-depth interviews 2 weeks prior to any interview recruitment taking place. The letter will explain that the interview is voluntary, confirm the purpose of the research and length of the interview and enable the participant to opt out if they prefer. Those that do not opt out of the interviews will be contacted by a researcher for recruitment and asked again if they are willing to take part. Finally, verbal agreement will be also obtained by researchers prior to conducting phone interviews and written agreement for any face-to-face interviews.

6 Disclosure protocol

This section provides the protocol that will be followed by all consortium members should a participant or other interviewee make a disclosure of concern when in contact (by phone or in person) with a member of the research team.

Agreed disclosure policy (the Consortium)

Should members of the evaluation consortium see, hear or experience anything during their contact with staff or participants that gives them cause for concern as to harm or illegal behaviour, they will follow the agreed disclosure policy.

The process for disclosure will be administered at an organisational level, but the requirements and timescales will apply across the consortium. The process has three stages:

- i. The members of the evaluation consortium will raise their concerns with the project lead at their organisation.³³ This should happen within 24 hours of the receipt of the information giving rise to a concern.
- ii. The organisation lead³⁴ will decide whether to convene a Disclosure Panel (comprised of at least 2 senior staff members). The Panel will make an assessment of the concern and make a decision about whether a disclosure should be made, to whom it should be made, and what should be disclosed. This discussion should take place within 24 hours of the Panel being convened.
- iii. The organisation lead will record the actions and outcomes of the discussion with the Disclosure Panel. The Panel will make any disclosure as agreed within 48 hours, where possible, of the concern being raised. Disclosures will be shared with IES as consortium lead in an anonymised form.

The only exception to this process is where there is a clear and immediate need to call the emergency services (police, ambulance, fire service) because of an immediate danger to life.

Timings set out for the process, above, provide time limits; however, the decision-making process will be expedited wherever possible.

In reaching a decision whether or not to disclose, the Panel will consider:

- The seriousness of the alleged harm or illegal behaviour.
- The strength of evidence for the evaluation consortium member's concerns.
- The ability of the individual involved to seek help for him or herself.

³³ Using existing organisational incident reporting processes.

³⁴ Where the disclosure issue has been reported by a NatCen fieldworker, the decision about whether to take the issue to the NatCen Disclosure Panel will be made by the NatCen's Director of Field.

- Whether the situation is already known to support services (e.g. IPS worker, GP, health visitor, social worker) or others capable of intervening (e.g. family members).

The Panel will decide exactly what and how much information is to be passed on, by whom and to whom. Wherever possible disclosure will be made directly to a representative public body or authority or known agency already working with the individual. The Panel will always consider whether the disclosure should be made to the IPS worker.

Notification of a disclosure

Where a disclosure takes place, the consortium lead will inform the WHU of the nature of the disclosure without sharing personal information. It is expected that in most cases it will be appropriate to inform the IPS service that a disclosure has been made (even if the details of the information disclosed is not divulged to the IPS service).

If the disclosure relates to an 'urgent safety measure' or 'serious adverse event' the REC will be notified in line with the procedures for reporting such events.

All consortium members will receive clear instructions and support on the disclosure process and regular reminders about how to escalate any concerns they have in their contact with participants and staff during the course of the evaluation.

7 Risk management and mitigation

7.1 Risks from agreement process and how they will be mitigated

The collection of service users' personal data and the matching and linkage with other datasets is not expected to cause any significant emotional or time burden for participants. As described, in the trial protocols, informed agreement will take place during the first meeting with an Employment Specialist. This will be a face-to-face encounter during which research participants will be asked to agree to the collection, sharing and storage of their personal data.

7.2 Risks related to data protection, confidentiality and how they will be mitigated

The evaluation will involve the collection of sensitive, personal data about service users – their health usage, employment and benefit details, as well as details of their interactions with the IPS service. Collecting these data is essential to address the research questions and to allow participants records in national administrative data sets to be identified for use in the study. The reasons for using measures derived from multiple administrative data sources rather than survey data or management information alone are:

- To reduce respondent burden and the costs associated with primary data collection;
- To be able to assess the impact of the trials on a wider range of outcomes than would be possible with survey data, or from a single source of administrative data alone;
- To construct a detailed picture of historical information on participants and assess the comparability of intervention and control groups;
- To avoid reliance on recall and thus to improve the accuracy of impact estimates;
- To reduce the risk to reliability if survey response rates are low;
- To provide flexibility over the time-periods that can be considered in the analysis, compared with a survey which takes place at a fixed point in time.

Set against this is the fact that administrative data is not collected specifically for the purpose of the evaluation and so measures may not be ideally suited to capturing the outcomes required to assess the impact of the programme. Administrative data may also be incomplete or have limited coverage of the population of interest, which may affect the interpretation of findings. There is also the risk that linking together multiple sources of information increases the probability that individuals can be identified, which could cause stress, embarrassment and a breach of rights to privacy. The impact of a data leak could also be worse where the information held on individuals is more wide-ranging.

To reduce this risk, a considerable amount of time has been spent by the Evaluation Consortium, WHU, WMCA and SCR to identify and map how data will flow: how it will be collected, stored, transferred, used and destroyed (work package 1.4). A detailed data flow document is appended to the research protocol, showing in detail all the different ways that data will be collected.

A full account of how data will be managed and confidentially ensured is set out in Section 7.

7.3 Risks and benefits from survey and how they will be mitigated

Providing service users with an opportunity to answer questions about IPS services offers them a chance to contribute and feed in their views to the evaluation consortium and to shape future services. Participants will be told that although their contribution to the trial may not benefit them directly, that there may be significant benefits for others in the future. As well as the value this contribution places on the views of participants, collecting survey data from all those taking part in the trial (both the treatment and control groups) provides invaluable evaluation data, allowing the research team to develop robust impact findings and to assess the value of rolling the evaluation out to a wider population, potentially benefitting large numbers of future IPS services users.

Alongside the benefits are a number of risks which must be considered and mitigation strategies developed.

One risk is that agreement to take part in the survey is not free and informed. This will be mitigated by sending an advanced letter to all participants (including information about the survey and making it clear they do not have to participate), and seeking explicit informed agreement at the point of each interview, during the introduction to the survey.

A second risk is that topics in the survey around health conditions, wellbeing and the impact of health conditions on respondents' lives may be considered sensitive. There are a number of ways in which the risk of adverse reactions can be mitigated:

- Questions will have been scrutinised and tested to ensure that there are none in the survey that are considered too potentially distressing using i) user testing during survey development, and ii) full cognitive testing immediately prior to the survey being finalised and piloted.
- NatCen interviewers are trained and experienced in administering surveys dealing with sensitive topics. Interviewers will be able to refer respondents to sources of support, either online (accessible to those completing the survey online) or as a paper leaflet.

A third risk from surveys is that they place undue burden on respondents. To ensure this does not happen, the interim survey is expected to be approximately 20 minutes, and the outcome survey approximately 30 minutes when administered by telephone, which our experience tells us is appropriate to all delivery modes being proposed.

Alongside risks to participants are a number of factors that have the potential to undermine the value of the trial and its findings. These risks include low survey response, which could result from a poor understanding of the importance of the evaluation, lack of engagement from referrers/ those involved in delivering the trial and a poor communication strategy. These risks will be addressed by ensuring all materials are carefully designed and accompanied by a well-considered communication strategy, and by ensuring all those involved in delivery are given opportunities to engage and develop buy-in to the trial.

In addition to a lack of engagement from trial participants, survey response may also be affected by the quality of sampling information provided by trial sites. If the information collected is incomplete or of poor quality, this will affect NatCen's ability to make contact with trial participants, and therefore of achieving high response rates and numbers needed for robust impact estimates. Efforts will be taken to ensure that trial sites understand the importance of collecting

comprehensive and accurate contact information. This will include details for a 'stable contact' - someone who can help put interviewers in touch with the respondent in the event that they change their contact information.

7.4 Risks and benefits from qualitative interviews with service users and how risks will be mitigated

Providing an opportunity for service users to share their experiences and views of the trials is an important element of the interviews for the process study and survey. This gives users the chance to contribute to the study, which we hope will be a positive experience that makes them feel valued. Gaining the perspectives of individuals who participated in the trial will also add unique insights, which through triangulation with other data sources, will generate rich and comprehensive findings.

Conducting interviews with service users also presents risks that must be mitigated. This includes that service-users are unclear on the aims and objectives of the research or are unaware that the research is voluntary and therefore their agreement is not fully informed. There is also a risk that service users are worried about their responses being shared amongst delivery staff or other authorities (such as DWP), and therefore do not share honest opinions or actual experiences – skewing the findings.

Interviewers will ensure that participants are informed of the aims and objectives of the research and what their participation will entail at the point of recruitment and prior to commencement of the fieldwork. Agreement for interviews will be conducted in line with Social Research Association (SRA) ethical guidelines, to ensure that participants:

- are informed from the first contact with the research team that they may decline to answer any questions put to them or to participate in the research process;
- know that they can withdraw from the interview at any point;
- do not feel compelled to participate and their involvement is voluntary and based on informed agreement.

When recruiting people with multiple and complex barriers, it is particularly important that everyone is able to access and understand the purpose of the research and what participation involves. All potential interviewees will be provided with information about the interviews in formats that consider the access needs of those being contacted, including easy-read versions. In screening calls, potential participants will be assured that participation is completely voluntary, whilst also being provided with the opportunity to opt-out of the interview, protecting them from any undue intrusion or distress.

Although there is a risk that the topics raised in the interview cover sensitive or personal issues that could result in undue intrusion, distress, personal embarrassment or psychological or other harm, consortium researchers are trained to be sensitive to cultural, religious, gender, health and other issues in the research population when undertaking their work. To address such risks when working with vulnerable and sensitive groups, researchers will:

- Take special care to protect the interests of those with mental health issues, those with disabilities and learning difficulties, the elderly and other vulnerable groups.
- Make special arrangements for participants with diminished capacity or when a person's understanding is limited due to age or learning difficulties

Research materials are developed with the barriers faced by participants in mind, ensuring that the design, presentation and the language of materials reflects the needs and learning abilities of participants. For example, topic guides have been developed to ensure that questioning is not unnecessarily intrusive. In addition, researchers have been provided with in-house disability

awareness training and will have information on local sources of support relating to the nature of the issues discussed.

Interviewers from the Evaluation Consortium will ensure that research participants are protected from undue intrusion, distress, indignity, physical discomfort, personal embarrassment or psychological or other harm when completing an interview. Consortium researchers are experienced in asking questions about health, wellbeing and employment or looking for work sensitively, using clear and appropriate language to ensure that participants feel comfortable discussing their situation and views. Through emphasising the voluntary nature of the study, staff will ensure that interviewees are aware that they do not have to answer any questions that they do not want to. Researchers will also use the topic guide flexibility, only asking the relevant questions for that participant at the time of the interview.

Interviewers will be able to refer respondents to sources of support, either online (accessible to those completing the interview by phone) or as a paper leaflet to be designed by the trials.

Consortium researchers understand that there may, under exceptional circumstances, be a need to breach confidentiality. From our experience of working on many other similar projects, we do not expect to encounter such a situation often. These breaches of confidentiality may occur only if there appears sufficient evidence to raise serious concern about:

- the safety of service users
- the safety of other persons who may be endangered by the service user's behaviour
- the health, welfare or safety of children or vulnerable adults

Consortium researchers will breach confidentiality if a service-user is considered to require immediate protection, particularly, if it is discovered that they:

- are being subject to abuse or neglect
- are self-harming or threatening self-harm

The Evaluation Consortium has agreed that these grounds for disclosure will apply across all data collection methods (interviews and surveys). Any questions about breaching confidentiality would be first raised with the Project Manager to agree an appropriate course of action by which to raise concerns.

7.5 Risks and benefits from qualitative interviews with practitioners

As with service users, providing an opportunity for practitioners to share their experiences and views of the programme is an important element of the evaluation. It gives them the chance to contribute their views, which we hope will be a positive experience in itself, and will also be seen as making a positive contribution to the improvement of services and furthering the evidence base. Gaining the perspectives of frontline staff delivering the pilot will also be essential for the process study, adding unique insights on the operation of the trial and the achievement of outcomes. Through triangulation with other data, this will generate rich and comprehensive findings.

The evaluation involves inviting professionals to take part in interviews. Doing so is voluntary and there is no obvious risk to their safety. Interviews will be confidential, and such confidentiality will only be breached for safeguarding purposes if the interviewee is considered to be at risk of personal harm. The appropriate course of action will be decided by the Project Manager. The main burdens on practitioners will be of time and inconvenience. Hence, we will be flexible when organising interviews and complete these by telephone wherever possible at a time that is suitable for the interviewee, which removes additional time burdens of travel time and room booking. We will also limit the interviews to 60 minutes to and conduct individual rather than group interviews (unless otherwise desired) to ensure that participation in the research is not overly time-consuming, which will be important as interviews will be conducted during working hours. With

respondent permission, interviews will be recorded with encrypted Dictaphones, and fully transcribed.

Clear information on the purpose of the interviews and the topics to be covered will be provided in advance, giving practitioners an opportunity to ask any questions or raise any concerns that they may have and to prepare for the interview, if they feel that this is necessary.

One other risk is that staff and professionals may worry that their response are used to monitor their performance or are reported back to managers or colleagues. Therefore, interviewers will emphasise that there are no right or wrong answers to the questions being asked and ensure that the interviews occur in a safe space where participants feel comfortable discussing the topics at hand. Respondents will be assured of the confidentiality of personal responses and that all reporting will be non-attributable, with quotes anonymised and any identifying information removed (for example, identifiable demographic characteristics, specific experiences or references, and any idiosyncratic phrases used or content).

7.5 Potential risks for researchers

Although we foresee few risks to researchers from this research, researchers have received training prior to conducting fieldwork, which briefs staff on the particular challenges that they may face when conducting research with certain groups, and guidance on the fieldwork process including gaining informed agreement and interviewing techniques. In addition L&W have recently completed training on Managing Challenging Interviews which will be cascaded to all those undertaking interviews through research briefings.

During any fieldwork that is conducted individually outside of the office, a buddy system will be utilised, whereby a member of staff monitors the safety of the individual completing fieldwork until their appointments are completed for each day. The person monitoring safety has a fieldwork diary and will know the times and venues of the scheduled research appointments. The interviewee will let the buddy know when they enter and exit the interview to inform them of their safety. If the buddy cannot get hold of the interviewee when their interview was due to finish, there is an escalation procedure in place.

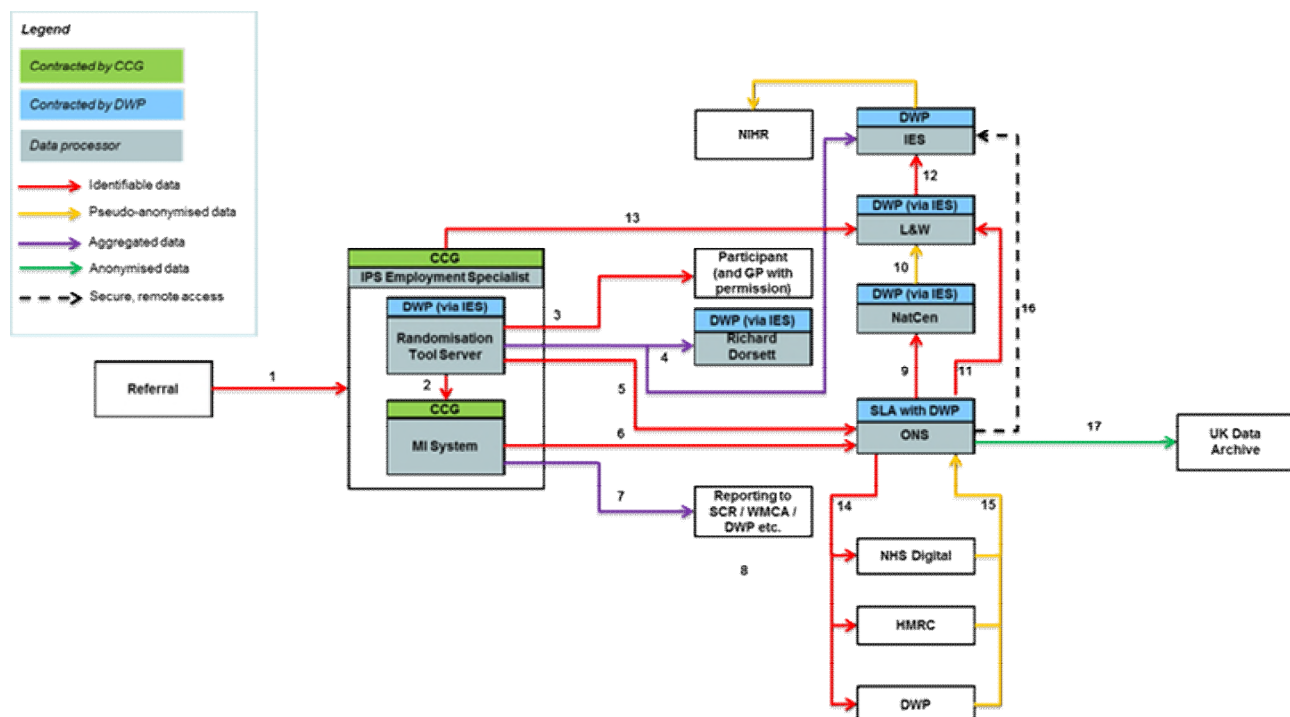
8 Data management

8.1 Introduction to the data collected and used in the trial and evaluation

The trial presents a complex and interacting data collection architecture between the local sites (WMCA and SCR), the evaluation consortium (largely via the safe haven provided by ONS) and latterly the ADRN and the UK Data Archive. This is summarised in the map below.

This section of the research protocol focuses on data collection, storage, transfer and deletion within the evaluation consortium, with the appendix providing full information on the intended flow of data between all parties.

It must be noted that this is the current proposal for data flows, but there may need to be minor adjustments during the lifetime of the evaluation.



8.2 Data management by the evaluation consortium

The consortium will collect and store all data in compliance with the Data Protection Principles under the Data Protection Act 1998 (DPA) until May 2018 and thereafter in compliance with the legislative provisions enacted so as to provide for the terms of the General Data Protection Regulations 2016 (GDPR).

All consortium organisations and operators within them conform to the Generic Security Accreditation Document (GSAD) standards operated by the Department for Work and Pensions.

All have the highest standards of data storage and security. As lead organisation, IES has oversight to ensure all standards are adhered to.

- **Secure physical storage:** Paper-based information will be stored in locked filing cabinets on the secured access premises of consortium members. Data will be pseudo-anonymised and personal identifiers replaced with codes.
- **Secure electronic storage:** All data will be held on the on-shore secure, encrypted servers of the consortium. No data will be transferred to personal computers. The movement of data throughout the evaluation and between the evaluation team, trial sites and national data owners, has been logged in the data flow diagram.
- **Secure destruction:** The data held by the evaluation will be securely destroyed and deleted at an agreed date (normally 12-24 months), following sign-off and publication of the final, synthesis report.
- **Access control measures:** Data will be held on the consortium's servers within folders that restrict access to the research team.
- **Secure data transmission:** all personal, sensitive, confidential data will be transferred securely and in encrypted format.

Detailed arrangements are set out in Table 8.1. We are awaiting full information from the safe haven (Office for National Statistics) although note, this organisation operates the highest standard of data security and protection.

Table 8.1: Data flows and storage

| Description of data | Transfer | Storage | Destruction |
|--|---|---|---|
| Pseudo-anonymised data owned by NHS Digital | From NHS Digital to ONS (safe haven) | Added to the Office of National Statistics (ONS) Data Integrator. Stored securely on ONS servers with highly restricted access. | To be confirmed |
| Pseudo-anonymised data owned by HMRC | From HMRC to ONS (safe haven) | Added to the Office of National Statistics (ONS) Data Integrator. Stored securely on ONS servers with highly restricted access. | To be confirmed |
| Pseudo-anonymised data owned by NHS Digital, DWP and HMRC, plus pseudo-anonymised management information | From to ONS (safe haven) to evaluator | Pseudoanonymised linked data set once transferred data will be held securely on IES secure servers with restricted access. | Five years after the end of the trial i.e. by March 2024. |
| Management information, including personal data to enable matching | From service providers to ONS (safe haven) | Added to the Office of National Statistics (ONS) Data Integrator. Stored securely on ONS servers with highly restricted access. | To be confirmed. |
| Management information, including personal data to enable matching | From ONS safe haven to NHS digital / DWP / HMRC | Added to the Office of National Statistics (ONS) Data Integrator. Stored securely on ONS servers with highly restricted access. | To be confirmed. |

| Description of data | Transfer | Storage | Destruction |
|--|--|--|---|
| Names, addresses, phone numbers, information on language proficiency, treatment status and health/disability of individuals as well as email addresses of all in the treatment and control groups, needed to send out survey invitations | From the service providers to the relevant part of the evaluation consortium (NatCen) | Once transferred data will be held securely on NatCen systems with restricted access. | Data will be securely deleted in line with DWP requirements at the end of the project. |
| Completed surveys – online, telephone and face-to-face | From the evaluation consortium (NatCen) to the data haven, and between members of the evaluation consortium (NatCen to IES). | Data will be held securely on NatCen systems with restricted access. Personal identifiers will be removed from the analysis dataset prior to transfer. Sample information will be held separately and contact details updated as new information is received. | Data will be securely deleted in line with DWP requirements at the end of the project. |
| Contact details and demographic and other information to support purposive sampling (for example gender, age, ethnicity, employment and health status) for a sample of the treatment group, needed to invite them to be interviewed | From the intervention providers to the relevant part of the evaluation consortium (L&W) by secure, encrypted email transfer. | <p>If it is necessary to store personal data for the project in hard copy format, we will do so in lockable cabinets and ensure this is labelled anonymously. We will dispose of such data when it is no longer required, by shredding it and giving this to waste collection services.</p> <p>Access to the systems or service that will handle this data are controlled through the implementation of the Password / Permissions Protocol. This documented control ensures that permissions to project-specific file structures and usernames / passwords to the secure server are controlled, monitored and audited periodically. One Way Hash (SHA1, SHA2, or MD5) is employed at domain login, using Kerberos Security Encryption built into Windows 2008 Server (only the server can decrypt the information).</p> | |
| Interview audio recordings – on digital recorders | Digital recorders that will be used for the qualitative fieldwork use a numerical passcode to protect recording playback, and an encryption password to upload the recordings to a computer. | | Recordings are automatically deleted from the Dictaphones as soon as they have been transferred to the secure server. |

| Description of data | Transfer | Storage | Destruction |
|--|--|--|-----------------|
| Interview audio recordings, transcripts – on servers | Will not be transferred between consortium organisations | Each organisation will store the data for this project on our secure server, which is accessed via a secure 128bit encryption and 2 phase authorisation. Auditing software ensures that all actions and commands are logged and reviewed periodically. | |
| Randomisation / baseline data | Anonymised to the consortium for monitoring of the RCT | Management information data set will be held securely on IES secure servers with restricted access. | To be confirmed |

Source: Evaluation consortium 2018

8.3 How confidentiality of data will be ensured by consortium's working practices

The consortium has carefully planned data flows to ensure the collection and storage of named data is minimised, and that all data collected – named, pseudo-anonymised and/or anonymised – is held in line with our obligations under the Data Protection Act and GDPR. This entails, for example, ensuring that response data is pseudo-anonymised and stored separately from named data. For example, RCT data, and data deriving from national data sources will be pseudo-anonymised by safe havens, transferred to the IES on-shore, encrypted server to a folder only accessible to the quantitative impact assessment team³⁵ where they will be held for processing.

The named data to enable survey sampling will be securely transferred³⁶ direct to NatCen and stored on their on-shore, encrypted server. Identifier data will be held securely until deletion in a separate, restricted access folder from survey responses.

The named data to enable service user interviews to be recruited will be first, securely transferred to L&W³⁷. IES and RAND are expected to undertake some of these interviews hence L&W will transfer sample to each, using PGP encrypted emails. In each case, named data will be held in restricted access folders on each organisation's on-shore, encrypted servers. Once recruited, a participant code will be created which will be used on all recordings and transcripts and in analysis. Participant identifiers will be held separately, in secured folders, from their response data, until deletion. The team conducting qualitative research at IES will not be involved in the quantitative data analysis, and vice versa. Neither team will be able to access the secure folders operated by the other.

For interim and final impact analysis, it is expected that survey data will be combined with RCT and outcome data stemming from national databases. Survey responses will be pseudo-anonymised to allow matching but to ensure no personal identifiers are transferred. Data transfer will take place between NatCen and IES using PGP encrypted email.

8.4 Plans for data storage, destruction, archiving after the end of the trial

As noted, all electronic data associated with the evaluation will be held on the on-shore, encrypted servers of the organisations involved in its delivery. The consortium will agree with the Unit, a deletion date for named data. Typically, this is once the final report is published. Where named data is in hard copy, this will be held in locked cabinets within consortium organisations' restricted access offices. These data will be shredded, and taken away by a waste collection company that offers secure destruction of confidential data.

The trial agreement documents set out that the consortium may keep pseudo-anonymised and anonymised data for a period of three years following the publication of the report. Again, a final deletion date will be agreed with the Unit as part of the evaluation contract.

The Unit expects that data from the evaluation will be stored in the Essex Archive, and the agreements allow for this. Such data will be securely transferred by the Consortium to the archive. Once successfully transferred, these data will then be deleted from the servers of consortium members.

³⁵ Led by Richard Dorsett and Helen Gray

³⁶ Transfer protocols are yet to be established, however will use appropriate levels of encryption

³⁷ Appropriate transfer protocol will be agreed, using an appropriate level of encryption

9 Research management, governance and peer review

9.1 Programme and project management

The evaluation is being conducted by a consortium of independent research organisations led by the Institute for Employment Studies. The consortium members are: RAND Europe, National Centre for Social Research, Learning and Work Institute and Richard Dorsett. IES, a not-for-profit organisation, is an independent, apolitical, international centre of research and consultancy in public employment policy and HR management. It is a focus of knowledge and practical experience in employment and training policy, the operation of labour markets, and HR planning and development.

Stephen Bevan, IES Head of HR Research Development, has oversight and overall responsibility for the evaluation as Chief Investigator. He is also the Programme Director for the evaluation. Becci Newton (Associate Director) and Rosie Gloster (Senior Research Fellow) comprise the IES project management team. They will coordinate the inputs from the evaluation experts across the consortium who lead on various components of the research and who meet weekly, as follows:

- Helen Gray, IES Principal Economist: administrative data, outcome measures and economic assessment; lead on information governance;
- Emily Tanner, NatCen Head of Children, Families & Work: user surveys, evaluation team advice on trial agreement materials and processes;
- Tony Wilson, LWI Policy and Research Director: theory of change, process evaluation, management information analysis;
- Richard Dorsett, RDR Ltd Director, randomisation process, monitoring and analysis;
- Emily Disley, RAND Associate Group Director for Home Affairs: ethics.

IES has an established and robust approach to project management involving: inception meetings with clients, at which formal and informal communication protocols are agreed and the risk register is developed further. The nominated Project Director is the key point of contact with the client but additional points of contact are two Project Managers and the Administrator for the project. Clients are updated weekly on the progress of the research, notified of any difficulties encountered and advised of potential solutions, informed of emerging findings, and consulted in the development of research materials. Clear lines of responsibility and project monitoring combine to ensure projects are completed to client expectation, timetable and budget.

Regular contact by email and telephone, between the consortium and the WHU, its NHS England account managers as well as local sites, has characterised communications during the design phase. This will continue to ensure there is a close understanding of progress and issues as the trials enter their delivery phase.

9.2 Peer review

The evaluation design and approach has been subject to review by the Director of IES, Nigel Meager, and Ton Ling Senior Research Leader at RAND Europe, neither of whom will be involved in the delivery of the evaluation, as well as by a senior member within the consortium:

- Nigel Meager, Director of IES, peer reviewed the full protocol including the statistical approach
- Professor (emeritus) Tom Ling, Head of Evaluation, RAND Europe RAND Europe
- Dr Alex Sutherland, Research Leader at RAND Europe, reviewed the protocol in full focusing on the statistical approach.

Their review and commentary is documented in Appendix B.

Outputs from the evaluation consortium will be subject to quality assurance processes organised by the Department for Work and Pensions.

The evaluation consortium plans to also to submit results for publication in peer reviewed journals.

9.2.1 Biographies for the peer review team

Nigel Meager, BA, MPhil, FAcSS, FRSA, Director of the Institute for Employment Studies

Nigel has a long and varied research track record covering the functioning of national, regional and local labour markets, unemployment, skill shortages, labour market flexibility, changing patterns of work and equal opportunity policies and practices. He has, since the late 1980s, had a particular interest in the role of self-employment and small businesses in the labour market, and has published widely on this topic. A major strand of his work has focused on the evaluation of public training and employment programmes and active labour market measures, with a particular focus on the participation of disabled people and other disadvantaged groups in the labour market. He recently led the official evaluation of the UK government's Work Programme, for the Department of Work and Pensions. Much of his work has an international emphasis and he is especially interested in the comparison of labour market policies between European countries, and in the identification and transfer of good practice in policy development.

He has been a specialist adviser to various select committees of the British House of Commons: the Education and Employment committee (in 1996-97, and 1998-99), the Trade and Industry committee (2004-05) and the Work and Pensions committee (2008-09). He has been the UK representative on the European Commission's Expert Group on the Employment of Disabled People and a member of the Employment and Training Committee of the Royal National Institute of Blind People. He has been a member of the Advisory Group on the Impact of Employment Regulation of the UK Department for Business, Innovation and Skills, and the expert advisory panel of the UK Commission for Employment and Skills (UKCES). He was also a Visiting Fellow of the UKCES (2011-12). From 2008-2012 he was chair of the Executive Committee of the UK Association of Research Centres in the Social Sciences (ARCISS). He is a Trustee of the Social Research Association and in October 2016 he was conferred as a Fellow of the Academy of Social Sciences in recognition of his contribution to social science.

Professor Tom Ling, Head of Evaluation, RAND Europe

Tom Ling is Head of Evaluation at RAND Europe leading evaluations and impact assessments focused on the key challenges facing organisations delivering public benefits and health and wellbeing impacts in particular. This work includes leading research and evaluation on the delivery of better integrated health and social care (the Integrated Care Pilots for UK Department of Health

and Social Care), quality improvement (Engaging with Quality for the Health Foundation, embedded evaluation of Project Q to improve quality and safety), the evaluation of Innovation Health and Wealth for the DH, the evaluation of research and health care (CLAHRCs), the changing role of Community Hospitals (NIHR) and projects on commissioning health and social care for NHS England and for OECD. He was recently invited to contribute to a policy dialogue on healthcare commissioning in the Republic of Ireland by the Irish Government. His current health research focuses on innovation and improvement and their relationship to driving quality and productivity in the UK health system. He has over twenty years of experience in researching on and leading research projects and he has published widely on evaluation, accountability, implementing health reforms, and related topics. His roles outside RAND Europe include a professorship (Emeritus) at Anglia Ruskin University, a membership of the College of Assessors at the Health Foundation, and he is a PI the Behaviour and Health Research Unit, University of Cambridge.

Dr Alex Sutherland, Research Leader, Communities, Safety & Justice, RAND Europe

Alex is a senior researcher with responsibility for proposing and leading projects covering a range of topics across the remit of RAND but particularly in the areas of evaluation, criminal justice, education and community cohesion. He currently leads five randomised controlled trials in education and criminal justice and has a background in quantitative evaluation.

Since 2012/2013 Alex has been an Associate Member, Department of Social Policy and Intervention, University of Oxford; has been acting as viva examiner; teaching evaluation methods; advisor to the Sinovuyo Caring Families Programme an RCT to reduce child abuse in Kenya and a member of the Violence Research Centre, Institute of Criminology, University of Cambridge.

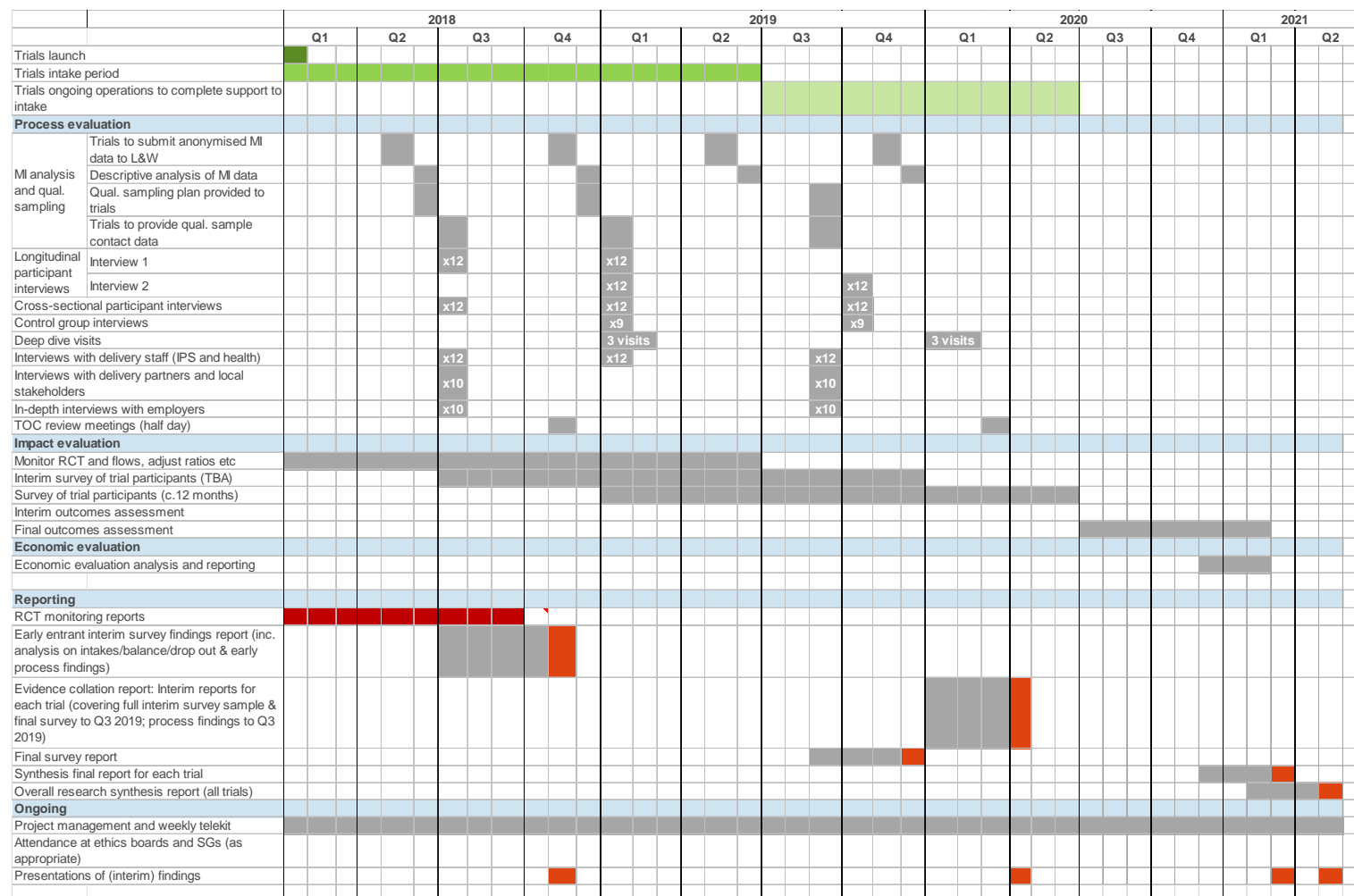
10 Research timetable

The design phase for evaluation commenced in February 2017 and is operating throughout the period until the trial's intended commencement in Spring 2018.

The evaluation will run in parallel with trial delivery from Spring 2018 through until Spring 2019. It will continue to track outcomes beyond trial closure for a further 12 months, with potential to extend this to assess longer term outcomes using administrative data. Evaluation reporting will conclude no later than Summer 2021.

The evaluation is comprised of multiple strands of activity, provided in detail in the methods section. Both RCT monitoring and the process evaluation will commence from the trial's start. The RCT will be monitored throughout the trials' delivery, with this monitoring being intensive over the first quarter, to provide rapid feedback on any adjustments required, but intensity of reporting will reduce over time. Management information will be regularly reviewed throughout the delivery phase and process evaluation qualitative research will also be active throughout delivery.

A draft timetable for the evaluation is supplied overleaf. This sets out the current expectations and understanding of the research activity required. This may be adjusted to be responsive to the delivery of the trials.

Figure 10.1: Schedule for the evaluation components

Please note, the timetable is indicative

Source: Consortium 2017

Appendix A: Evaluation design and approach

This appendix describes how the evaluation of the trials was designed. The evaluation has three phases:

- **A design phase** ending in Spring 2018 – in which the new services are being designed and procured, the evaluation design worked out, and relevant applications made for approval of the evaluation
- **An implementation and evaluation phase** between Spring 2018 and Summer 2020 – in which service users will be recruited into the trial, and data collection will be undertaken.
- **An analysis and reporting phase** between Summer 2018 and Spring 2021, including interim survey reports and presentations, trial reports and a final synthesis report. Early monitoring data collected will be analysed and reported to allow early referral and intake problems to be identified and managed in a timely manner.

Phase 1: Design

Literature review and scoping

A review of existing evaluations of IPS models was conducted to inform the evaluation design. At the same time, the consortium made contact with the WMCA and the SCR teams to familiarise themselves with the planned IPS services. The following research questions guided the review:

- Drawing on past IPS logic models, what are the key elements of the trialled interventions and what are the possible pathways which will lead to the intended outcomes?
- How the trialled interventions will contribute to the existing evidence base?

In order to respond to these questions the consortium conducted a targeted literature review and collaborated with another research team conducting systematic reviews on IPS approaches - the RISE project³⁸ - to limit duplication of work. In total, nine articles were supplied by the RISE researchers. These provided a variety of IPS approaches which draw on somewhat different intervention logics.

The consortium also ran a search for evidence on IPS for the 'in work' (IW) group on Google Scholar but found no results. The RISE team has identified very limited evidence on IPS used for the IW group but recommended one study (Coole et al., 2012). We checked references used in Coole et al. and searched for other articles citing it but did not find more relevant articles.

³⁸ The RISE project is conducting systematic reviews that will report on: a) moderators and mediators in IPS, and b) specific components in the IPS-approach such as interventions that are associated with improved outcomes. The review includes a full methodological assessment allowed the consortium to extract information on how the trials were designed and implemented, who did randomisation, at which point, what system was used, etc from relevant papers.

However, some trials in the shortlist of documents included IW group in their populations (Fergusson et al., 2012; Li-Tsang et al., 2008; Magura et al., 2007; Reme et al., 2014).

These articles provided some information relevant to the intervention design proposed by the Sheffield City Region (SCR), which comprises IPS support for the traditional target group (OOW), as well as for the IW group (supporting people in maintaining their current employment or in moving from one job to another).

Finally, the shortlist of papers was supplemented with additional purposefully chosen articles which tested the IPS model with different subgroups (e.g. long vs. short-term conditions, different age groups, etc.) and implementation settings (multiple vs. single sites, large vs. small scale). All 20 articles included in the review (with a brief explanation for reasons for their inclusion) are listed in table at the back of this appendix.

Developing a Theory of Change for the trials

Theory of Change is an approach used to map the connections between activities and outcomes within an intervention, to generate hypotheses about how the intervention will achieve the desired change. Using this approach within evaluation facilitates a more systematic focus on explaining *how* and *why* an intervention works (or does not work).

For these trials, we have developed three interlinked theories of change for each site. These comprise:

- An intervention level Theory of Change, which shows how client change should be achieved;
- A health systems Theory of Change, which shows how change in health systems should be achieved (in development);
- An employer Theory of Change which shows how change in employer behaviour should be achieved.

These are supplied as supporting documents to the research protocol (document set 7).

The three Theories of Change are linked, in that activities in the systems level Theories of Change support client outcomes and vice versa. The individual and systems level Theories of Change for each site will also be drawn together in an overarching programme-level Theory of Change for the trial as a whole.

To develop the Theories of Change, the consortium first reviewed site documentation and conducted initial scoping interviews/group discussions with site-level stakeholders and WHU staff, focusing on the aims and intended outcomes for the trials. Along with insights from the literature review about how IPS works in other contexts, this fed into the production of draft Theories of Change for the two trials. The intervention (client-focused) Theories of Change were then tested with stakeholders at each site in half-day workshops during May 2017, followed by a further workshop in each site to finalise the revised Theories of Change and to consider the outcome measures needed to test the Theories of Change. Subsequently, system-level Theories of Change were further refined through a telephone discussion with key stakeholders at each site.

The Theories of Change set out, in a systematic way, the activities that are intended to be delivered as part of the trial, and how they will result in the desired outcomes. This has been used to guide the evaluation design, by identifying the key outcomes to be measured and the processes to be explored in qualitative research. The Theories of Change are 'live' documents and will be further refined during the remainder of the design phase and subsequently tested against practice during delivery.

Finalisation of outcomes

The Theory of Change workshops with sites were used to identify likely intermediate and main outcomes from the trials. A cost-benefit modeller, developed by the WHU, was also reviewed as a starting point in mapping the full range of expected costs and benefits from the trials. In addition, the review of existing evidence was used to identify outcome measures used in other studies which assessed Individual Placement Support (IPS) interventions. This made it possible to identify measures which had been previously tested and which met our main requirements for the study, namely that they:

- Were likely to be sufficiently sensitive to be able to capture any impacts from the trials.
- Offered high levels of internal consistency.
- Minimised the burden on respondents.
- Were appropriate to carry out the economic analysis.

Having identified the full range of outcomes that the trials seek to effect, the most appropriate source of information, and the measures most likely to meet the requirements set out above, a proposed list of outcome measures was collated. This list was then revised in response to comments received from the Work and Health unit and discussions with representatives of NHS Digital, HMRC and DWP regarding other sources of administrative data and the likely reliability of particular data items.

The rationale for the decisions reached on the selection of measures can be found towards the end of this appendix in the section entitled 'Further details on the rationale for the selection of measures within the evaluation'.

Finalising evaluation design and data collection tools

Having selected the outcomes to be measured in the trial, the next set was to decide how to measure these in survey instruments and interviews to be conducted with service users.

Interview topic guides

Separate topic guides for the process evaluation have been designed for service users, delivery staff, employers, key stakeholders and trial partners. These have been included as an annex to this document. The topic guides are currently indicative drafts to show the anticipated coverage of the interviews, but may be adapted flexibly during the delivery of the trials in order to respond to emerging themes and changes. In addition, the topic guides are designed to allow flexibility within the interviews themselves, so interviewers can tailor questions to be responsive to the participant and their experiences within clearly defined areas for discussion. This will ensure that the requisite amount of detail is captured, while minimising the burden on respondents, and eliciting a positive interview experience.

The topic guides explore and test the underlying assumptions in the Theories of Change. This will include (as appropriate to the respondent) their experiences of activities undertaken, intermediate outcomes, and longer term outcomes. Interviews will explore the extent to which these occurred as anticipated in the Theories of Change, perceived linkages between support activities and outcomes experienced, and the critical success factors and barriers to achieving outcomes. The topic guides also allow the research team to explore whether trial design has been implemented as intended and the processes by which system level change (in health professionals' and employer behaviour) may occur.

Survey design

One advantage of an RCT is that it facilitates collection of baseline data. This is described in more detail in the main body of the protocol but in this section, the principles guiding information collection are discussed. In addition to baseline data, some outcomes will be observable from administrative data sources. However, to observe those outcomes that most closely reflect the theory of change described above requires conducting post-randomisation surveys of trial participants.

The purpose of including a survey of trial participants is to collect the outcome measures that cannot be obtained through administrative or management information data sources and to investigate perceptions and experiences for the purpose of understanding 'process' elements such as the barriers/facilitators to impact and the levers of change.

The survey design developed alongside the Theory of Change workshops and the specification of outcome measures (described above).

Questionnaire content

The questionnaire design was based around the primary and secondary outcomes. The questionnaire also aimed to capture (1) the intermediate outcomes identified in the theory of change that might indicate progress towards the primary and secondary outcomes and (2) experiences of services and perceived impacts.

A wide range of measures was considered for the outcomes and the selection was made in collaboration with multiple stakeholders including the Work and Health Unit, officials from the two sites, individuals from across the evaluation consortium and service users.

Questions to include were selected based on how well they met the following primary criteria:

- **Validity.** The key measures should have high levels of internal consistency and should have other validity data collected from field tests and/or application in similar trials.
- **Burden.** The administration of tools should minimise the demands made of participants in order to mitigate the risk of drop-out and to maximise the opportunity of gathering meaningful data at baseline and follow-up.
- **Sensitivity.** The tools should be sensitive enough to capture changes in health and wellbeing over the course of the trial.

And the following secondary criteria:

- **Specificity.** Some tools measure general health or health-related quality of life and others are more specific to particular diagnoses.
- **Applicability.** The health-led trials will include participants who are out of work at the point at which they are recruited and an additional (though smaller) group who are in work on recruitment. Key measures needed to be applicable to both groups.
- **Economic Analysis.** Some measures are routinely used in economic evaluation to look at the costs and benefits of interventions
- **Comparability.** Some tools needed to be considered because they are routinely used in other trials funded by government and the comparability of findings, effects and impact is an important secondary outcome of the health-led trials.

Once measures were selected, key stakeholders and experts from the Government's Health and Work Unit and the consortium fed into the questionnaire design. The input was also sought from

user groups in both the WMCA and SCR comprising adults with lived experience of mental health problems (mild to severe), substance misuse and physical disabilities drawn from established patient stakeholder groups. Drawing on their experience of living with health conditions, these individuals gave feedback on issues such as the sensitivity and clarity of questions being asked, which was incorporated into the final survey design. Whilst it was not possible to change the wording of proprietary measures, user views were used to inform decisions about the use of different scales, and inclusion of particular items.

User testing

Service users were engaged in the survey design process through user testing workshops carried out in each site. The aim was to obtain feedback on the broad approach of the surveys, the agreement materials³⁹ and key question topics from the questionnaire. Service users were engaged as a group rather than individually and were asked not to provide details about their own circumstances.

The user groups from each site comprised adults with lived experience of mental health problems (mild to severe), substance misuse and physical disabilities drawn from established patient stakeholder groups. Drawing on their experience of living with health conditions, these individuals gave feedback on issues such as the sensitivity and clarity of questions being asked, which was incorporated into the final survey design.

Cognitive testing

The survey were cognitively tested and will be piloted to ensure understanding, acceptability and suitability. Post-pilot adjustments will be made prior to the start of the trial. Ideally, the cognitive testing took place in December 2017, to allow amendments to the outcome measures to be approved by the HRA and implemented within the randomisation tool prior to the start of recruitment.

Survey timings

The evaluation consortium worked in collaboration with sites to determine the content and approach to collecting baseline data prior to randomisation. It was decided to carry out an interim survey four months after randomisation to collect 'in-service' early outcomes and experiences of engaging with services. The decision on the timing of the final survey was informed by the duration of the IPS interventions in each site and the constraints of the delivery timetable. At this stage, it is anticipated that the final survey will be timed for 12+ months post-randomisation. The difference in duration of the interventions poses challenges for pooled analysis across sites – the timing of the final survey may be reviewed.

Randomisation

Another important task during the design of the evaluation was to agree on the approach to randomisation. Randomisation is central to the objective of achieving a reliable estimate of the impact of adding IPS to the range of employment and health support usually available. By allocating to treatment or control conditions on a purely random basis we hope to achieve groups that, in the absence of the treatment, would be expected to experience similar outcomes on average. The differences in average outcomes that we do observe can then be confidently attributed to IPS.

³⁹ Due to changed understanding of requirements in light of GDPR, consent materials are now redrafted as agreement materials and have not, in this form, been subject to user testing

Randomisation will be carried out using an online software platform developed by the Behavioural Insights Team (BIT) and used in the Islington IPS trial. It conforms to the ISO27001 standard. This is described further in chapter 3 in the main body of the protocol.

Economic evaluation design

The economic evaluation aims to estimate the social return on investment from the trials, both for the exchequer and for society as a whole. As well as seeking to estimate the return on investment resulting from the trials, the analysis will estimate likely returns if the trials were extended nationwide.

In designing this part of the evaluation it has been necessary to consider the range of information required to estimate all costs and benefits likely to arise for both the exchequer and society, taking into account unit costs/benefits and the numbers of individuals participating in the trials. The design phase has also considered the type of information which will be required to estimate how the ratio of costs to benefits might vary if the trials were extended to areas with different characteristics.

The Theory of Change workshops with sites were used to identify likely intermediate and final outcomes from the trials. A Cost-benefit modeller, developed by the WHU, was also reviewed as a starting point in mapping the full range of expected costs and benefits from the trials. In addition, the review of existing evidence was used to identify outcome measures used in other studies which assessed Individual Placement Support (IPS) interventions.

Whilst the WHU's cost-benefit modeller provides an initial estimate of the expected return on investment, this is based on assumptions about the likely impact of the trials and the costs that will be incurred. Information collected during the course of the trial will be used to improve on this initial estimate. This involves producing quantitative estimates of the impact of the trials on different outcome measures and information on the costs of implementation, as well as the financial value of any costs and benefits. As a result, the final estimate of the return on investment from the trials will benefit from more comprehensive information on costs and benefits than that available at the time the cost-benefit modeller was produced, making it be more robust and defensible.

Having identified the full range of outcomes that the trials seek to effect, the most appropriate source of information, and the measures most likely to meet the requirements set out above, a proposed list of outcome measures was collated. Survey data will be used to assess the benefits of the trials, both by comparing outcomes for intervention and control groups and observing changes compared to baseline information contained in the management information. Due to limits on the amount of information that it is possible to collect on the survey, analysis of linked administrative data will also be used to estimate the impact of the trials and to value any benefits or costs. This will make it possible to take into account a wider range of potential costs and benefits from the programme and to lessen reliance on recall to improve the accuracy of the analysis. The list of potential outcome measures was revised in response to comments received from the Work and Health unit and discussions with representatives of NHS Digital, HMRC and DWP regarding other sources of administrative data and the likely reliability of particular data items.

To estimate the return on investment from the trials it is necessary to collect data on the costs of administering the trials and the value of the benefits that result. Much of the required data on costs will be collected from providers, through the supply of management information. Some information used to value costs and benefits will come from publically available sources, such as information on the rate of statutory sick pay, or estimates of the costs to the public purse of attending Accident and Emergency etc.

Further details on the rationale for the selection of measures within the evaluation

This section sets out the rationale for the choice of measurement tools by describing a set of criteria for the selection of these tools:

1. **Validity.** The measures should have high levels of internal consistency (as measured, for example, by Cronbach's Alpha coefficient) and should have other validity data (e.g. content validity, predictive validity) collected from field tests and/or application in similar trials.
2. **Burden.** The administration of tools should minimise the demands made of participants in order to mitigate the risk of drop-out and to maximise the opportunity of gathering meaningful data at baseline and follow-up. Thus, tools should have few items, take very little time to complete and should have realistic recall periods (i.e. more recent or immediate the time period for self-reports of symptoms the better).
3. **Specificity.** Some tools measure general health or health-related quality of life and others are more specific to particular diagnoses. If specific measures, for example, of mental and physical (such as musculoskeletal) health at baseline and follow-up are required then well-established and condition-specific measures may be preferable, bearing in mind the limitations imposed by other criteria in this list. Some more general physical health measures also collect data on mental health, which can be helpful if comorbidity is likely to be prevalent among sub-groups of participants.
4. **Applicability.** The health-led trials will include participants who are out of work at the point at which they are recruited and an additional (though smaller) group who are in work on recruitment. Most mental health and musculoskeletal tools are suited to both groups. Some measures are also designed to collect data on the extent to which a given health condition affects attendance, hours worked and self-reported productivity at work. Some of these measures are also frequently used to quantify the costs of lost productivity attributable to ill-health at work, which may support wider economic evaluation activity.
5. **Economic Analysis.** Some measures are routinely used in economic evaluation to look at the costs and benefits of interventions or to calculate standardised indicators such as Quality Adjusted Life Years (QALYs). As economic evaluation is a key strand of the national evaluation, the utility of the tools used to collect baseline and follow-up data for this purpose is considered along with the other criteria.
6. **Sensitivity.** The tools should be sensitive enough to capture changes in health and wellbeing over the course of the trial. Related to this, some way of assessing the magnitude of this change would also be helpful. For example, if similar measures are used in other studies, it would be possible to consider the relative impact of the current trial.
7. **Comparability.** Some tools may need to be considered because they are routinely used in other trials funded by government and the comparability of findings, effects and impact is an important secondary outcome of the health-led trials.

Arguably, the four most important criteria (in rank order) are:

1. Burden – experience from previous trials and from early work carried out at site level indicates that 'light-touch' data collection is essential to reduce the costs of the trials and to minimise the risk of participant attrition;
2. Sensitivity – to use outcome measures which are likely to be sufficiently sensitive to detect any changes;
3. Validity – to maximise the credibility of the trials and to optimise the replicability, transferability and scalability of the interventions using well-validated tools is an important consideration;
4. Economic Analysis – the trials need to be able to identify, track and quantify both the costs and benefits of the interventions. This means that credible baseline and follow-up data needs to be collected which will support this objective.

These criteria have been the focus in identifying the most suitable mode for collecting each of the outcome measures required for the evaluation.

Table 10.1 summarises the reasons for drawing each of the potential outcome measures from each mode, taking into account the criteria of validity, burden and the requirement to support the economic evaluation.⁴⁰ It also details any restrictions on the reuse of existing questions, where these are known. As the measures drawn from the management information will be aligned with those used on the survey, these two sources are grouped together.

Table 10.1 Reasons for choice of outcome measures from each mode

| Mode | Measure | Include? (Yes/No) | Reasons for inclusion or rejection (considering burden, validity, whether required for economic analysis and restrictions on reuse of existing questions) |
|--|---|-------------------|--|
| Survey (including related Management information questions) | EuroQol-5D-5L (EQ-5D-5L) 5-item. Comprises 5 dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression | Yes | More concise than alternative measures. Widely used and validated – used across WHU trials. Track record in economic evaluations. Also includes a Visual Analogue Scale, but there is the issue of modifying the VAS for different modes. Necessary to register study in order to use the scale and there may be a licencing fee. The size of the fee depends on the size of the study, type and funding source. |
| | 12 or 36-item Short Form Health Survey (SF-36 or SF-12) – generic quality of life measures | No | Greater respondent burden than EQ-5D-5L. |
| | Musculoskeletal Health Questionnaire (MSK-HQ). 2-item. On joint, back, neck, bone and muscle pain and stiffness, including how much the respondent has been bothered by this in the two weeks prior to interview. | Yes | Short Includes a psychosocial element (to pick up comorbid mental illness) Can be used by those both in and out-of-work. Need permission from Oxford University Innovations but freely available to publically funded health-care providers, non-commercially funded academic researchers and other non-commercial users. |
| | General Anxiety Disorder-7 (GAD-7), 7-item, and Patient Health Questionnaire-8 (PHQ-8), 8-item on anxiety and depression. | Yes | Measures are commonly used and well-established, they are also sensitive to small changes in condition. No restrictions on use and also being used on Group Work trial. The 8 item variant of the PHQ9; this excludes the item 'Thoughts that you would be better |

⁴⁰ Although the other criteria mentioned in section **Error! Reference source not found.** are not specifically referenced here, they were also considered in selecting appropriate outcome measures. Material on sensitivity will be added once the HRA application has been submitted and will be included in the draft technical report.

| Mode | Measure | Include? (Yes/No) | Reasons for inclusion or rejection (considering burden, validity, whether required for economic analysis and restrictions on reuse of existing questions) |
|------|--|----------------------|--|
| | | | off dead or of hurting yourself in some way', which is not recommended for use in self-administered surveys (i.e. web completion). |
| | PATIENT HEALTH QUESTIONNAIRE-9 (PHQ 8) (i.e. 9 minus suicide)) 8-item on anxiety and depression. | Yes | Measures are commonly used and well-established, they are also sensitive to small changes in condition. No restrictions on use |
| | Hospital Anxiety and Depression Scale (HADS) 14-item. Features seven questions for anxiety and seven for depression giving a separate score for each. | No | Measures anxiety and depression simultaneously, whilst giving a separate score for each Used in the study with participants with common mental health problems (Reme et al. 2014) and 'in work group' (Coole et al 2012). Overlap with GAD-7 and PHQ-8. |
| | Clinical Global Impressions (CGI) Scale, 3-item | No | Short, measures illness severity (CGIS), global improvement or change (CGIC) and therapeutic response and can be used by IPS specialists and participants (i.e. could be used for triangulation) |
| | World Health Organisation Disability Assessment Schedule (WHODAS-II). 36 or 12-item. Measures the impact of physical and mental health conditions on daily activities. Does not focus on confidence and skills to manage health condition - see PAM in alternative measures below. (NB we do not yet have access to this measure to assess it fully). | No | Concise Validated Covers 6 different functioning domains: cognition (2 items), mobility (2 items), self-care (2 items), getting along (2 items), life activities (2 items), and participation (2 items). Overlap with other health questions and limited added value. |
| | Work and Social Adjustment Scale (WSAS), 5-item | No | Measures the effect of a particular disorder on the respondent's work and leisure activity. Considered as an alternative to WHODAS-II, but overlap with EQ-5D, so limited added value given need to minimise respondent burden. |
| | Patient Activation Measure (PAM) – Measure of the knowledge, skills and confidence a person has in managing their own health and care | No | Possible alternative to WHODAS-II but we do not have access to the measure itself to assess yet. |
| | Short Warwick-Edinburgh Mental Well-being Scale (SWEMWBS) 7-item. Short form of WEMWBS, items for which are all worded positively and cover both feeling and functioning aspects of mental wellbeing. | Yes | Widely used and validated Seen as more simple than PHQ-9 Other measures such as GAD-7 and PHQ-9 are deficit-based. Need to register use of scale, but no fees for use. |

| Mode | Measure | Include? (Yes/No) | Reasons for inclusion or rejection (considering burden, validity, whether required for economic analysis and restrictions on reuse of existing questions) |
|------|---|----------------------|--|
| | Office for National Statistics Personal Well-being Questions (ONS-4), 4-item | No | Duplication with SWEMWBS and GAD-7/PHQ-9, but strong preference of WHU to include to aid comparability with other trials, so we propose using the first item only to reduce overlap. |
| | General Self-Efficacy Scale (GSE Scale) 10-item. Designed to assess perceived self-efficacy and ability to cope with daily hassles as well as adaptation after experiencing stressful life events | Yes | Widely used, including group work who trial. Can be used in same instrument as JSSE. No restrictions on use of scale provided source acknowledged in written reports. |
| | Understanding Society questions on GP visits (Wave 9 main survey) | Yes | Restricted to GP visits, rather than other questions as high degree of overlap with information available from NHS Digital, which does not rely on recall. Also preferable to use admin data sources to reduce respondent burden. However, no information on GP visits currently available, so a survey question will be used to address the needs of the economic evaluation. |
| | Number of prescriptions | No | Likely to be difficult for respondents to recall accurately. Differences in the costs of drugs mean imprecise data likely to be of limited value to the economic evaluation. ⁴¹ |
| | Employed ⁴² , self-employed, unemployed, inactive | Yes | To be cross-referenced with admin data/ contingency measure if problems accessing admin data. |
| | Length of time in employment/out of work | Yes | To be cross-referenced with admin data/ contingency measure if problems accessing admin data. |
| | Benefit status | Yes | To be cross-referenced with admin data/ contingency measure if problems accessing admin data. |
| | Employment status since randomisation | Yes | To be cross-referenced with admin data/ contingency measure if problems accessing admin data. |
| | Costs of using secondary services (as a result of IPS referral) | No | Required for economic analysis, but possible to estimate from |

⁴¹ Note also that individual-level data on prescriptions is not currently available from NHS Digital. The data on prescriptions that NHS Digital do hold are aggregated at practice-level and so are of limited value for the evaluation.

⁴² Whilst in theory it might also be possible to collect information on occupation and industrial sector, this would require backcoding. The evidence review also suggested it was not generally considered in other IPS studies and is likely to be of limited value in terms of assessing outcomes in the current trials.

| Mode | Measure | Include? (Yes/No) | Reasons for inclusion or rejection (considering burden, validity, whether required for economic analysis and restrictions on reuse of existing questions) |
|------|--|----------------------|---|
| | | | published data sources. |
| | Costs of travelling to work | No | Required for economic analysis, but possible to estimate from published data sources and likely to be of limited additional value relative to increased respondent burden. |
| | Costs of childcare | No | Required for economic analysis, but possible to estimate from published data sources and likely to be of limited additional value relative to increased respondent burden. |
| | Work Productivity and Activity Impairment Scale – General Health (WPAI-GH). 6-item. Measures effect of health problems on individual's ability to work and perform regular activities. | Yes | Captures impact on productivity and time off sick resulting from improvements in managing health condition at work Concise Widely used and validated Can be used in economic models. No restrictions on using scale, but asked to provide notification of publications. |
| | Workplace Employment Relations Study (WERS), Qs A7-A9. 19 sub-questions but scope to condense Captures: Job Control, Job Satisfaction (including security and pay) and Work Stress | Yes | Questions from robust national study Captures wider range of the factors influencing job satisfaction and job quality than the other alternative measures considered |
| | Utrecht Work Engagement Scale (UWES/UWES-9), 17 or 9-item | No | Captures job satisfaction and engagement with work. More focus on how job makes individual feel (e.g. inspiration, immersion, challenged) than WERS, and all statements are worded positively |
| | Psychosocial quality of work measure (12-item) taken from Household, Income and Labour Dynamics in Australia (HILDA) Survey. | No | Captures some similar areas to WERS (job demands, control, security and fairness), but not work stress. |
| | Usual earnings from employment (weekly/ monthly/ annual) | Yes | To be cross-referenced with admin data/ contingency measure if problems accessing admin data. |
| | Usual weekly hours worked | Yes | To be cross-referenced with admin data/ contingency measure if problems accessing admin data. |
| | Attitudes to employment with health condition (from NatCen survey of ESA claimants in the Work-related Activity Group) | Yes | To explore intermediate outcomes for survey respondents. |
| | Job Search Self-Efficacy Scale (JSSE). | Yes | Widely used in similar studies, including group work trial. Multiple variants used in other |

| Mode | Measure | Include? (Yes/No) | Reasons for inclusion or rejection (considering burden, validity, whether required for economic analysis and restrictions on reuse of existing questions) |
|----------------------------|---|----------------------|---|
| | 8-item. Self-efficacy relating to finding employment | | studies, NatCen currently seeking advice on whether there are any restrictions on its use. |
| | Whether any job applications in recent 4-week period | Yes | To explore intermediate outcomes. |
| | Whether any job offers in recent 4-week period | Yes | To explore intermediate outcomes. |
| | Whether any jobs accepted in recent 4-week period | Yes | To explore intermediate outcomes. |
| | Training/education | Yes | To explore Intermediate outcomes and support economic evaluation. |
| | Engagement and satisfaction with service at end of programme/Perceived effect of service on motivation and confidence to find work (from NatCen survey of ESA claimants in the Work-related Activity Group) | Yes | To explore intermediate outcomes for survey respondents. |
| | 12-item latent and manifest benefits of employment scale | No | Uncertain whether inclusion would be justified, given additional respondent burden. |
| | Demographic information | Yes | To explore interplay between individual characteristics and the impact of the trials. |
| Administrative data | Length of spell on out-of-work benefits | Yes | To reduce respondent burden and avoid reliance on recall for start and end dates of benefit spells. Required for economic analysis. |
| | Proportion claiming out-of-work benefits at monthly intervals | Yes | To reduce respondent burden and avoid reliance on recall for start and end dates of benefit spells. Required for economic analysis. |
| | Amount of benefits paid | Yes | Would require additional data matching by DWP for anything other than the most recent amount for the current benefit claim, but would provide an indication of costs for economic analysis. |
| | Receipt of free school meals | No | Would require additional data linking (to National Pupil Database) for limited additional benefit. Some potential to estimate based on claims for other benefits. |
| | Gross pay from employment | Yes | To reduce respondent burden and increase accuracy of data. Required for economic analysis to estimate benefits and progress in work. |

| Mode | Measure | Include? (Yes/No) | Reasons for inclusion or rejection (considering burden, validity, whether required for economic analysis and restrictions on reuse of existing questions) |
|------|---|----------------------|---|
| | Gross income from self-employment (business income minus business expenses) | Yes | To reduce respondent burden and increase accuracy of data. Required for economic analysis to provide information from self-employment. |
| | Tax paid during employment spell | Yes | To reduce respondent burden and increase accuracy of data. Required for economic analysis to estimate costs/benefits to the exchequer |
| | Total tax due on income earned from self-employment | Yes | To reduce respondent burden and increase accuracy of data. Required for economic analysis to estimate costs/benefits to the exchequer from self-employment |
| | Length of employment spell | Yes | To reduce respondent burden and avoid reliance on recall for start and end dates of employment spells. Required for economic analysis to observe whether employment sustained e.g. whether in employment lasting three months or more six months after randomisation etc. |
| | Proportion employed at monthly intervals | Yes | To reduce respondent burden and avoid reliance on recall for start and end dates of employment spells. Required for economic analysis to observe whether employment sustained. |
| | Amount of tax credits received | Yes | To reduce respondent burden and increase accuracy of data. Required for economic analysis to estimate welfare costs/savings. |
| | Student loan repayments | Yes | To reduce respondent burden and increase accuracy of data. Required for economic analysis to estimate costs/benefits to exchequer |
| | RTI earnings | Yes | To reduce respondent burden and avoid reliance on recall for earnings. Expected to be more complete than WPLS data. Required for economic analysis. |
| | RTI total tax paid | Yes | To reduce respondent burden and increase accuracy of data. Required for economic analysis to estimate costs/benefits to exchequer |
| | RTI amount of student loan repayments | Yes | To reduce respondent burden and increase accuracy of data. Required for economic analysis to estimate costs/benefits to exchequer |

| Mode | Measure | Include? (Yes/No) | Reasons for inclusion or rejection (considering burden, validity, whether required for economic analysis and restrictions on reuse of existing questions) |
|-------------|---|------------------------------|--|
| | RTI employer national insurance contributions | Yes | To reduce respondent burden and increase accuracy of data. Required for economic analysis to estimate costs/benefits to exchequer |
| | RTI employee national insurance contributions | Yes | To reduce respondent burden and increase accuracy of data. Required for economic analysis to estimate costs/benefits to exchequer |
| | RTI length of time with employer | Yes | To reduce respondent burden and avoid reliance on recall of start and end dates for employment spells and to update information available from the WPLS. Required for economic analysis. |
| | RTI proportion of those randomised in causal employment | Yes | To reduce respondent burden and to provide information on the nature of employment. |
| | RTI hours worked at monthly intervals | Yes | To reduce respondent burden, reduce reliance on respondent recall and to provide information on progression in employment over time. Required for economic analysis. |
| | Days in hospital | Yes | To reduce respondent burden and avoid reliance on recall. Required to estimate health care costs for economic analysis. |
| | Proportion in hospital at monthly intervals | Yes | To reduce respondent burden and avoid reliance on recall. Required to estimate health care costs for economic analysis. |
| | Proportion of hospital spells non-elective | Yes | To reduce respondent burden and avoid reliance on recall. To capture emergency use of health services. |
| | Total number of procedures | Yes | To reduce respondent burden and avoid reliance on recall. Required to estimate demand on health services for economic analysis. |
| | Number of outpatients appointments attended | Yes | To reduce respondent burden and avoid reliance on recall. Required to estimate demand on health services for economic analysis. |
| | Number of operations | Yes | To reduce respondent burden and avoid reliance on recall. Required to estimate demand on health services for economic analysis. |
| | Total days of inpatient care | Yes | To reduce respondent burden and avoid reliance on recall. Required to for economic analysis. |

| Mode | Measure | Include? (Yes/No) | Reasons for inclusion or rejection (considering burden, validity, whether required for economic analysis and restrictions on reuse of existing questions) |
|-------------|--|------------------------------|---|
| | Total number of outpatients procedures | Yes | To reduce respondent burden and avoid reliance on recall. Required to estimate demand on health services for economic analysis. |
| | Total number of visits to A & E | Yes | To reduce respondent burden and avoid reliance on recall. To capture emergency use of health services for economic analysis. |
| | Total amount of time in A & E | Yes | To reduce respondent burden and avoid reliance on recall. To capture emergency use of health services for economic analysis. Having access to information on time demands for services gives some flexibility over how costs/savings can be valued. |
| | Total amount of diagnosis time in A & E | Yes | To reduce respondent burden and avoid reliance on recall. To capture emergency use of health services for economic analysis. Having access to information on time demands for services gives some flexibility over how costs/savings can be valued. |
| | Total amount of treatment time in A & E | Yes | To reduce respondent burden and avoid reliance on recall. To capture emergency use of health services for economic analysis. Having access to information on time demands for services gives some flexibility over how costs/savings can be valued. |
| | Number of days from referral to mental health service to discharge | Yes | To reduce respondent burden and avoid reliance on recall. Required for economic analysis. |
| | Total number of mental health appointments attended | Yes | To reduce respondent burden and avoid reliance on recall. Required for economic analysis. |
| | Length of time from referral to discharge from IAPT | Yes | To reduce respondent burden and avoid reliance on recall. Required for economic analysis. |
| | Length of time from opt-in to IAPT to discharge | Yes | To reduce respondent burden and avoid reliance on recall. Required for economic analysis. |
| | Number of IAPT appointments | Yes | To reduce respondent burden and avoid reliance on recall. Required for economic analysis. |
| | Number of IAPT appointments at low or high intensity | Yes | To reduce respondent burden and avoid reliance on recall. Required for economic analysis. |
| | Proportion of appointments attended | Yes | To reduce respondent burden and avoid reliance on recall. Required for economic analysis. |

| Mode | Measure | Include? (Yes/No) | Reasons for inclusion or rejection (considering burden, validity, whether required for economic analysis and restrictions on reuse of existing questions) |
|-------------|---|------------------------------|---|
| | Total amount of clinical contact required | Yes | To reduce respondent burden and avoid reliance on recall. Required for economic analysis. |
| | Proportion requiring psychotropic medication | Yes | To reduce respondent burden and avoid reliance on recall. Required for economic analysis. |
| | Improvements in assessment scores over time (PHQ9, GAD7, Work and social adjustment scale scores for work; home management; social leisure activities; private leisure activities; and close relationships) | Yes | To reduce respondent burden and avoid reliance on recall. Required for economic analysis. |
| | Length of time from referral to community services healthcare provider to discharge | Yes | To reduce respondent burden and avoid reliance on recall. Required to estimate healthcare savings for economic analysis. |
| | Length of time from referral to community services healthcare provider to completion of treatment | Yes | To reduce respondent burden and avoid reliance on recall. Required to estimate healthcare savings for economic analysis. |
| | Number of community services appointments attended. | Yes | To reduce respondent burden and avoid reliance on recall. Required to estimate healthcare savings for economic analysis. |
| | Total amount of time receiving community services care | Yes | To reduce respondent burden and avoid reliance on recall. Required to estimate healthcare savings for economic analysis. Having access to information on time demands for services gives some flexibility over how costs/savings can be valued. |
| | Activity-limiting health problem | Yes | To reduce respondent burden and avoid reliance on recall. Required for economic analysis. |
| | Demographic information | Yes | To explore interplay between individual characteristics and the impact of the trials. |

Source: Evaluation consortium 2017

References included in the literature review

| # | Short reference | Full references | Reason for inclusion |
|-----|----------------------------|---|--|
| 1) | * Allott et al., 2013 | Allott, K.A., Cotton, S.M., Chinnery, G.L., Baksheev, G.N., Massey, J., Sun, P., Collins, Z., Barlow, E., Broussard, C., Wahid, T. and Proffitt, T.M., 2013. The relative contribution of neurocognition and social cognition to 6-month vocational outcomes following Individual Placement and Support in first-episode psychosis. <i>Schizophrenia research</i> , 150(1), pp.136-143. | IPS (augmented with another intervention) for people with severe mental health illness |
| 2) | * Areberg & Bejerholm 2013 | Areberg, C. and Bejerholm, U., 2013. The effect of IPS on participants' engagement, quality of life, empowerment, and motivation: a randomised controlled trial. <i>Scandinavian journal of occupational therapy</i> , 20(6), pp.420-428. | IPS for people with severe mental health illness |
| 3) | * Au et al., 2015 | Au, D.W., Tsang, H.W., So, W.W., Bell, M.D., Cheung, V., Yiu, M.G., Tam, K.L. and Lee, G.T.H., 2015. Effects of integrated supported employment plus cognitive remediation training for people with schizophrenia and schizoaffective disorders. <i>Schizophrenia research</i> , 166(1), pp.297-303. | IPS (augmented with another intervention) for people with severe mental health illness |
| 4) | * Bell et al., 2014 | Bell, M.D., Choi, K.H., Dyer, C. and Wexler, B.E., 2014. Benefits of cognitive remediation and supported employment for schizophrenia patients with poor community functioning. <i>Psychiatric Services</i> , 65(4), pp.469-475. | IPS (augmented with another intervention) for people with severe mental health illness |
| 5) | ** Bond et al., 2008 | Bond, G.R., Drake, R.E. and Becker, D.R., 2008. An update on randomised controlled trials of evidence-based supported employment. <i>Psychiatric rehabilitation journal</i> , 31(4), p.280. | IPS for people with severe mental illness |
| 6) | Bond et al., 2016 | Bond, G.R., Drake, R.E. and Campbell, K., 2016. Effectiveness of individual placement and support supported employment for young adults. <i>Early intervention in psychiatry</i> , 10(4), pp.300-307. | IPS for young adults with severe mental illness |
| 7) | ** Burns et al., 2015 | Burns, T., Yeeles, K., Langford, O., Montes, M.V., Burgess, J. and Anderson, C., 2015. A randomised controlled trial of time-limited individual placement and support: IPS-LITE trial. <i>The British Journal of Psychiatry</i> , 207(4), pp.351-356. | IPS for people with severe mental illness |
| 8) | * Burns & Catty 2008 | Burns, T. and Catty, J., 2008. IPS in Europe: The EQOLISE trial. <i>Psychiatric rehabilitation journal</i> , 31(4), p.313. | IPS (adapted) for people with affective/ bipolar disorders |
| 9) | Christensen et al., 2015 | Christensen, T.N., Nielsen, I.G., Stenager, E., Morthorst, B.R., Lindschou, J., Nordentoft, M. and Epløv, L.F., 2015. Individual Placement and Support supplemented with cognitive remediation and work-related social skills training in Denmark: study protocol for a randomised controlled trial. <i>Trials</i> , 16(1), p.280. | IPS (augmented with another intervention) for people with severe mental health illness |
| 10) | * Coole et al., 2013 | Coole, C., Drummond, A. and Watson, P.J., 2013. Individual work support for employed patients with low back pain: a randomised controlled pilot trial. <i>Clinical rehabilitation</i> , 27(1), pp.40-50. | IPS (adapted) for an 'in work' population of people with chronic pain |
| 11) | Ferguson et al., 2012 | Ferguson, K.M., Xie, B. and Glynn, S., 2012, June. Adapting the individual placement and support model with homeless young adults. In <i>Child & youth care forum</i> (Vol. 41, No. 3, pp. 277-294). Springer US. | IPS (adapted) for homeless young adults |
| 12) | Howard et al., 2010 | Howard, L.M., Heslin, M., Leese, M., McCrone, P., Rice, C., Jarrett, M., Spokes, T., Huxley, P. and Thornicroft, G., 2010. Supported employment: randomised controlled trial. <i>The British Journal of Psychiatry</i> , 196(5), pp.404-411. | IPS implementation issues |

| | | |
|-------------------------------|--|---|
| 13) * Li-Tsang et al., 2008 | Li-Tsang, C.W.P., Li, E.J.Q., Lam, C.S., Hui, K.Y.L. and Chan, C.C.H., 2008. The effect of a job placement and support program for workers with musculoskeletal injuries: a randomised control trial (RCT) study. <i>Journal of Occupational Rehabilitation</i> , 18(3), pp.299-306. | IPS (adapted) for people with MSK injuries |
| 14) * Magura et al., 2007 | Magura, S., Blankertz, L., Madison, E.M., Friedman, E. and Gomez, A., 2007. An innovative job placement model for unemployed methadone patients: a randomised clinical trial. <i>Substance use & misuse</i> , 42(5), pp.811-828. | IPS (adapted) for people with substance abuse |
| 15) Ottomanelli et al., 2014 | Ottomanelli, L., Barnett, S.D. and Toscano, R., 2014. Individual placement and support (IPS) in physical rehabilitation and medicine: The VA spinal cord injury experience. <i>Psychiatric rehabilitation journal</i> , 37(2), p.110. | IPS (adapted) for people with spinal cord injuries |
| 16) * Reme et al., 2015 | Reme, S.E., Grasdahl, A.L., Løvvik, C., Lie, S.A. and Øverland, S., 2015. Work-focused cognitive-behavioural therapy and individual job support to increase work participation in common mental disorders: a randomised controlled multicentre trial. <i>Occupational and environmental medicine</i> , 72(10), pp.745-752. | IPS (augmented with another intervention) for people with CMH disorders in work & out of work |
| 17) Rinaldi et al., 2011 | Rinaldi, M., Montibeller, T. and Perkins, R., 2011. Increasing the employment rate for people with longer-term mental health problems. <i>The Psychiatrist</i> , 35(9), pp.339-343. | IPS for people with severe mental health illness |
| 18) Sveinsdottir et al., 2014 | Sveinsdottir, V., Løvvik, C., Fyhn, T., Monstad, K., Ludvigsen, K., Øverland, S. and Reme, S.E., 2014. Protocol for the effect evaluation of Individual Placement and Support (IPS): a randomised controlled multicenter trial of IPS versus treatment as usual for patients with moderate to severe mental illness in Norway. <i>BMC psychiatry</i> , 14(1), p.307. | IPS for people with moderate to severe mental health illness |
| 19) van Veggel et al., 2015 | van Veggel, R., Waghorn, G. and Dias, S., 2015. Implementing evidence-based supported employment in Sussex for people with severe mental illness. <i>British Journal of Occupational Therapy</i> , 78(5), pp.286-294. | IPS implementation issues |
| 20) Waghorn et al., 2014 | Waghorn, G., Dias, S., Gladman, B., Harris, M. and Saha, S., 2014. A multi-site randomised controlled trial of evidence-based supported employment for adults with severe and persistent mental illness. <i>Australian occupational therapy journal</i> , 61(6), pp.424-436. | IPS implementation issues |

NOTE: Articles shared by RISE are marked with (*); articles suggested by the WHU are marked with (**)

Source: Evaluation consortium 2017

Appendix B: Peer review commentary

Institute for Employment Studies 1

Peer review of the Innovation Fund Health-led Employment Trials Research Protocol

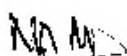
| Review Criteria | Reviewer opinion |
|---|--|
| Judgement on the quality of the proposed approach | The overall approach is thorough and comprehensive, and in its design and implementation proposes a high quality mixed-methods evaluation, taking good account of the existing knowledge and research literature (both methodological and with regard to experience of these kinds of intervention). It is underpinned by a well-resourced, expert research team across the consortium. In my judgement, therefore, the approach is of high quality. |
| Judgement on effectiveness of the proposed research to meet research aims | The approach to all three elements of the evaluation (impact, process and cost-benefit) would, if conducted according to the specification, and assuming that the trials function in practice to deliver the necessary scale and characteristics of participants over the trial period, be likely to meet the stated research aims for each element. |
| Judgement on how well the proposal balances data collection needs in light of participant burden (all stakeholder groups) | The participant burden is significant but not excessive, and seems broadly appropriate given the aims of the evaluation. Care has been taken in design of all aspects of the study to minimise unnecessary burdens on participants. |
| Judgement on the proposed statistical approach | Overall, the approach is well-considered and the statistical methods appropriate. I and other reviewers have made comments and suggestions for minor aspects which would benefit from more detailed specification or description in the protocol. |
| Overall assessment of whether the evaluation approach is fit for purpose | Yes |

Name: Nigel Meager

Role: Director, Institute for Employment Studies

Email: nigel.meager@ies.ac.uk

Signature:



Date: 27 July 2017

Peer review of the Innovation Fund Health-led Employment Trials Research Protocol

| Review Criteria | Reviewer opinion |
|---|---|
| Judgement on the quality of the proposed approach | The approach is sensible and pragmatic given the many as yet unconfirmed parameters for the trial. The overall impression of the research design and statistical strategy is one that has been produced in spite of a number of unknowns, but it is appropriately thorough. There will need to be more specifics given for the finalised version but these are largely dependent on external input. |
| Judgement on effectiveness of the proposed research to meet research aims | The planned approach for the RCT appears feasible given the logistics and complexity of the project. If executed as planned then it will provide robust evidence on the additive effect of IPS-type interventions compared to BAU. |
| Judgement on how well the proposal balances data collection needs in light of participant burden (all stakeholder groups) | n/a |
| Judgement on the proposed statistical approach | As above, there will need to be more specifics given for the finalised version but these are largely dependent on external input. The analysis approach will need to be more explicit for the final version as noted in my comments. There are appropriate suggestions in the event of randomisation failure, missing data or survey non-response. |
| Overall assessment of whether the evaluation approach is fit for purpose | In terms of the RCT design and the proposed analyses, yes, they are fit for purpose. |

Note: As requested by IES, I have only reviewed the RCT design and statistical analyses set out in version 7.0 of the protocol (which was sent on 26th July 2017).

Name: Dr. Alex Sutherland

Role: Research Leader

Email: alex_sutherland@rand.org

Signature:



Date: 26th July 2017

Peer review of the Innovation Fund Health-led Employment Trials Research Protocol

| Review Criteria | Reviewer opinion |
|---|--|
| Judgement on the quality of the proposed approach | <p>The approach has been carefully developed and described in the research protocol and overall it is strong. I think it should be made clearer at the start who the audience for this report is – it is primarily aimed at people who are reasonably well versed in research methods and vocabulary. Within this, and in the light of there being possibly a wider audience, consider clarifying or footnoting technical terms (depending on how widely you anticipate this being read).</p> <p>At the start, in the first paragraph, I think it is important to clarify that the research is not only about costs but also about well-being and health. This helps to frame and explain the research. I have also attached brief comments on the text itself.</p> |
| Judgement on effectiveness of the proposed research to meet research aims | <p>The only significant substantive issue that needs further thought is that I do not think it generates enough information about how people access local family and community based services and supports which may help them both navigate the employment world and build their personal resilience. These resources will be unevenly spread across the country and we need to understand what they are how they add to the capabilities of people engaging with the services. I think these data need to come from the interviews and baseline and will not flow easily from the qualitative work as currently described.</p> <p>Overall, the more we can structure the study from the outset to contribute to replication and scaling up (if the evidence supports this) the better.</p> |
| Judgement on how well the proposal balances data collection needs in light of participant burden (all stakeholder groups) | <p>This has been thought about carefully and a defensible balance struck.</p> |
| Judgement on the proposed statistical approach | <p>I have read through Alex Sutherland's comments on chapter three and agree with these. I have no additional comments to make.</p> |
| Overall assessment of whether the evaluation approach is fit for purpose | <p>An important governance issue is the role of PPI and this may be addressed elsewhere. If not it needs to be considered here.</p> <p>The reporting and engagement approach is fairly pedestrian and there may be more dynamic ways to build interest. In addition, consider appropriate feedback for the study participants.</p> <p>Otherwise the approach is fundamentally well founded.</p> |

Name: Professor (Emeritus) Tom Ling

Role: Head of Evaluation, RAND Europe

Email: tling@rand.org

- A secure web form is currently under development.
- An additional gateway website is also being developed as part of wider SCR work which can be used to signpost to the self-referral web form (no personal data will be entered into this gateway website).

West Midlands Combined Authority (WMCA):

Service Providers in WMCA will receive:

- E-referrals from GPs: Data will be auto-populated by EMIS / SystemOne but GPs will also have the option to add additional data to the referral in a free-text comment box.
- Referrals via the Thrive into Work website. Data will be entered into a secure web form (currently under development by Uscreates).

The data received via referral may differ slightly based on trial area and/or referral mechanism, but will include identifiable data such as:

- Name
- Address
- NHS number
- Preferred contact method
- E-mail address
- Telephone number
- Referral source
- Reasonable adjustment information (e.g. translator, hearing loop etc.)

At the Initial Reception Meeting the IPS specialist will use the Randomisation Tool⁴³ to:

- Establish eligibility to participate in the trial: The IPS specialist will use the randomisation tool to ask a series of questions to establish eligibility, based on the inclusion / exclusion criteria specific to either SCR or WMCA. If ineligible, the IPS specialist will signpost the person to other avenues of support in their area. If eligible, the IPS specialist will continue on to the next step – administering the trial participation agreement process.
- Administer the agreement to participate process: If eligible, the IPS specialist will outline the trial, and what participation entails. A key part of this process is to ensure that participants fully understand how data will be collected, shared, linked, stored, and used for the trial, if they agree to take part.

If a person wishes to take part, they will sign an agreement form, which will be kept by the Service Provider (a copy can be given to the participant). A copy of the agreement form will also be sent

⁴³ The Randomisation Tool will be hosted on a UKCloud server. This server will comply with DWP Security Policies and Standards which adopt and apply ISO 27001 Standards and Cyber Essentials.

either by secure e-mail or by secure courier to the ONS, to ensure the evidence of agreement is kept for the length of the trial [Note – logistics to be agreed].

- Collect baseline management information: The randomisation tool will include a question to confirm that agreement to participate has been recorded. Once this has been confirmed on the tool, a series of questions will be generated by the tool, which will enable the collection of baseline management information. The variables collected at baseline:



Baseline MI captured
by the Randomisation

- Administer randomisation: Once the baseline questions are completed, the IPS specialist clicks a button on the randomisation tool to randomise the participant into either the treatment or control group. The randomisation tool uses a subset of the baseline MI to randomise into one of these two groups, depending on the randomisation ratio set by Richard Dorsett. The randomisation tool will generate a unique Trial ID which will be given to the participant.

[2] Transfer from the Randomisation Tool to the Service Provider MI system

Each day, the randomisation tool will generate a random token (a unique SHA1 hash, 40 bytes) and e-mail this to download authorised users (i.e. the IPS specialist). A link will be in the e-mail which takes the user to the main page of the randomisation tool. On clicking this link, the user will need to login to the tool. The user will then see a page asking for the download token. If they have used the link provided in the e-mail, the token will be populated already. Clicking a button on this page will start the CSV download. If the user does not download the file before a new token is issued according to their schedule, then the previous token will be expired and the new csv file will be adjusted to include the items from the last potential download. The csv, and all traffic from the tool, will be transferred over HTTPS (TLS 1.2).

The IPS specialist will upload the MI baseline data included in the csv file to their MI system. This data will be held for the length of live running. At the end of live running, all electronic data will be removed from the MI system and securely destroyed (in line with Data Protection, all physical data retained by the Service Provider will also be securely destroyed).

For those randomised into the treatment group the following variables will be generated:

[TO BE INSERTED]

For those in the control group, the following variables will be generated (these are required for monitoring purposes e.g. to be aware if a control group participant returns in an attempt to be randomised into the treatment group etc.):

[TO BE INSERTED]

[3] Notification e-mail to the participant and their GP

The randomisation tool will generate an e-mail which confirms which group the participant has been allocated to. This e-mail confirmation will be sent to the participant and, with agreement from the participant, also sent to their GP to inform them of their patient's participation in the trial.

[4] Transfer of aggregate data to Richard Dorsett and IES

The randomisation tool will regularly generate aggregated reports of anonymous data for the Richard Dorsett and the Institute for Employment Studies (IES) who will use it to monitor referrals and randomisation to:

- Monitor how well the marketing of the trials is working, including what categories of people are being referred and from where, and how appropriate (good quality) these referrals are.
- To ensure that the randomisation tool successfully assigns the specified proportion of individuals to the treatment arm; and
- To explore the characteristics of the treatment and control groups in order to confirm that two similar-looking groups are identified.

This transfer will be via a secure web-link provided by the randomisation tool which will allow Richard and the IES to download a CSV file of aggregate data. In the initial stages of the trial these will be weekly reports, and the randomisation tool has the capacity to move to less frequent reporting, based on individual preferences of the receivers. The following aggregate data will be included in the download:

[TO BE INSERTED]

[5] Baseline MI transfer to the Office for National Statistics (ONS)

The baseline MI data for both treatment and control groups will be securely transferred from the randomisation tool to the ONS data coordinator. As described earlier, the randomisation tool will generate a random token and send it to the authorised user at the ONS. The ONS data coordinator can use this to securely download a CSV file of the baseline MI. The tool can generate and send these tokens daily to the ONS, but the ONS data coordinator can choose how often to initiate the downloads (the downloaded csv will be adjusted to include the items from the last download - I would advise weekly downloads). See information at [2] for more technical detail.

[6] Transfer of ongoing MI from Service Provider MI systems to the ONS

Following the Initial Reception Meeting, the IPS specialists will have regular meetings with participants that have been assigned to the treatment group. At these meetings the employment specialist will continue to collect on-going management information for service delivery, monitoring, and evaluation purposes. This MI will be securely transferred to a data coordinator within the ONS on a weekly basis [Note – details to be confirmed, but most likely a secure export of CSV file from the MI systems to a named contact / secure inbox at the ONS]:



Ongoing MI
captured by IPS speci

- SCR: South Yorkshire Housing Association (SYHA), who are delivering the service across SCR, will use the Salesforce Information Security Management System (ISMS) to store and transfer data for this trial. This platform is hosted in the UK and complies with ISO 27001 (Salesforce has also achieved ISO 27001/27018 certification for its ISMS from an independent third party).

- WMCA: To ensure consistent data quality and security across the several service providers in WMCA, all providers will use one common MI system. This MI system is being developed by CORE and will be hosted on a UKCloud server⁴⁴.

The on-going MI will be pseudo-anonymised using the Trial ID, unless the information is to update or amend participant contact details (and in such circumstances, the data would be identifiable).

This information will be regularly transferred from the MI system to the ONS data coordinator (frequency TBC). This will be a pseudo-anonymised transfer using the Trial ID, unless the information is to update or amend participant contact details, and in such circumstances the data would be identifiable.

[7] Regular reporting for monitoring, contracts management, and financial management

The MI databases will be configured to provide regular aggregated data reports to relevant bodies (i.e. the ACS, CCGs, and Work and Health Unit).

[8] Regular reporting of data on trial participation to National Institute for Health Research (NIHR)

The NIHR require individual-level information on trial participation, including referral sources. This could be extracted via the randomisation tool and uploaded by IES. The data will be pseudo-anonymised and uploaded securely via the NIHR web portal.

Data transfer to the evaluation research teams

The ONS data coordinator will provide the research teams with information so they can carry out the fieldwork necessary for the evaluation - service user surveys and qualitative interviews - as well as to allow for a descriptive analysis of the management information.

[9] Data transfer for Service User Surveys

For every participant in both the treatment and control group, ONS data coordinator will transfer the following information to NatCen:

- Trial ID
- Surname
- Forename
- Date of birth
- Gender
- Telephone number (mobile and landline numbers)
- Address and postcode at randomisation

⁴⁴ As noted earlier, the UKCloud server will comply with DWP Security Policies and Standards which adopt and apply ISO 27001 Standards and Cyber Essentials.

- E-mail address
- Site (WMCA or SCR)
- Reasonable adjustment information (main language / fluency in English / sight / hearing-loop requirements etc.)
- Whether currently in paid work, off sick, or on temporarily reduced hours
- Health condition
- Randomisation date

NatCen will receive the information on a monthly basis from the beginning of the trial to enable resource planning for the interim survey. This data will be in an Excel file and sent from ONS data coordinator to NatCen using an encrypted e-mail (PGP or similar).

NatCen will use the information they receive from the ONS data coordinator to contact the participant and to verify their identity. Once verified, NatCen will carry out the service user survey. All participants will be contacted to be invited to take part in the survey at fixed points after randomisation – once during the trial (the interim survey, around 4 months after randomisation), and again after their time on the trial has ended (the follow-up survey, around 12 months after randomisation).

If, while conducting the surveys, NatCen discover that there are changes to participant details or circumstances (i.e. change of address or contact information), NatCen will inform the ONS data coordinator using a secure e-mail (PGP encryption or similar). This will ensure that data is up to date as per the 4th principle of the Data Protection Act (data must be “accurate and where necessary kept up to date” in relation to the purposes of the processing). These updates will be provided on a monthly basis.

[10] Data transfer from NatCen to L&W

NatCen will transfer the pseudo-anonymised data to the ONS data integrator in an Excel file using a secure e-mail (PGP encryption or similar) – once after the interim surveys are completed, and once after the final surveys are completed. NatCen will also securely send pseudo-anonymised interim survey data to L&W, which will be used to construct a sample of participants for qualitative interviews.

Once NatCen have completed the final survey and transferred the pseudo-anonymised data to the ONS data integrator, they will securely destroy any data they may still hold on their systems.

The data point of contact for NatCen is Migle Aleksejunaite – Senior Data Manager:

- migle.aleksejunaite@natcen.ac.uk
- 0207 549 7105

[11] Data transfer for Qualitative interviews

There will be three waves of qualitative interviews. For Wave 1, L&W will use the pseudo-anonymised management information to construct samples of participants that will be invited to take part in an interview (purposive sampling will allow variations to be explored across trial locations and respondent characteristics). For Waves 2 and 3, L&W will use a combination of

pseudo-anonymised management information received from ONS data coordinator, and pseudo-anonymised survey data received from NatCen.

Once a sample of participants has been selected, L&W will send a list of the sample Trial IDs to the ONS data coordinator to request their identifiable data. In return, the ONS data coordinator will securely transfer the following information for the sample to L&W (both transfers will be in Excel files and using PGP encrypted e-mail):

- Trial ID
- Surname
- Forename
- Date of birth
- Telephone number
- E-mail address
- Address and postcode at randomisation (and current address and postcode if this has changed)
- Reasonable adjustment information (language / textphone / hearing loop requirements etc.)

[12] Data transfer from L&W to IES

IES will assist L&W with the qualitative interviews. L&W will securely transfer the necessary information to IES to enable them to carry out the interviews (this will be a secure file and encrypted e-mail). L&W and IES will use the information provided by the ONS data coordinator to contact the participant by post to inform them about the research and to enable them to opt out if they prefer. Those that do not opt out will be contacted by a researcher to verify their identity, explain the purpose of the interviews, and gain their agreement to carry out an interview. At the end of the interview participants in the longitudinal sample will be asked for their agreement for L&W and IES to securely store their trial ID in order to re-contact them in six months' time for a subsequent interview. To ensure that L&W and IES have the most up to date contact details for the participant, they will re-request the details after 6 months, prior to the longitudinal interview, to ensure that they are using the most up to date information. As before, this request will be made by sending the Trial IDs to the ONS in an Excel file using a secure e-mail (PGP encryption or similar).

Once L&W and IES have completed the initial interview, they will destroy the identifiable details for the rest of the sample. A record will be kept with the trial ID of anyone who opts out or declines to take part in an interview for future waves of recruitment. IES will securely destroy any identifiable data they hold before they access the final data-set that contains anonymised national admin data collected for the impact evaluation.

Pseudo-anonymised management information and survey results will be stored separately to identifiable data for further analysis, and transcripts will be stored with the pseudo-anonymised trial identifier. These transcripts will not be transferred to the ONS data integrator, but instead will be held within the Consortium (by L&W and IES) and be used to inform the evaluation of the trial.

Descriptive analysis of Management Information

L&W will also use the pseudo-anonymised management information that is sent to them by the ONS data coordinator to conduct a descriptive analysis of the participants' (treatment group) activity on the trial and intermediate outcomes. As stated previously, this pseudo-anonymised

management information will be kept separately from any identifiable data and securely destroyed once reporting for the evaluation is completed.

On current timetables [TBC], L&W will require data from the ONS data coordinator at four time-points:

- May 2018 (pseudo-anonymised MI to carry out a descriptive analysis of the MI and to draw a qualitative sample, L&W will request identifiable data of that sample in June 2018)
- December 2018 (pseudo-anonymised MI to carry out a descriptive analysis of the MI and to draw a qualitative sample, L&W will request identifiable data of that sample in January 2019)
- July 2019 (pseudo-anonymised MI to carry out a descriptive analysis of the MI and to draw a qualitative sample, L&W will request identifiable data of that sample in August 2019)
- January 2020 (pseudo-anonymised MI only – to carry out a descriptive analysis of the MI).

L&W will feed back any changes to contact information and / or circumstances to the ONS data coordinator after each wave of interviews.

[13] Data transfer from Service Providers to L&W

L&W will also conduct interviews with a range of other people involved in the trial (for example, IPS specialists). L&W will liaise directly with the sites to get the necessary information and all transfers will use secure files and PGP-encrypted e-mails.

The data point of contact for L&W is Liz Davies – Senior Researcher:

- liz.davies@learningandwork.org.uk
- 0116 285 9689 / 07917 123 089

[14] The transfer of identifiable data to NHS Digital, HMRC, and DWP

The ONS data coordinator will securely transfer the following identifiable data to NHS Digital, HMRC, and DWP.

- Trial ID
- NHS number (to NHS Digital), or National Insurance number (to HMRC and DWP)⁴⁵
- Date of birth
- Surname
- Forename

⁴⁵ Efforts will be made to ensure that the NHS number and NINO are collected by the IPS specialist. However, there will be instance where collecting it from the participant has not been possible. In these instances the National Data Owners will still be able to use the other identifiable data provided to locate and verify the correct administrative record, but with slightly less certainty than if the NHS number / NINO were provided.

- Address and postcode.

These organisations will use this data to identify and verify the relevant health, employment and earnings, and/or benefit information needed to contribute to the evaluation of the trials.

[15] The transfer of pseudo-anonymised data from NHS Digital, HMRC, and DWP to the ONS

Each organisation will pseudonymise the relevant administrative information using the Trial ID, and securely transfer it back to ONS data integrator.

Each organisation will be asked to retain an index key of the data sent to them by the ONS data coordinator for the period of time the merged datasets are retained by the safe haven. This is to allow for potential updates to be made to the datasets during this time.

Once identifiable data is no longer required (i.e. the service delivery phase of the trial has been completed (approx. April 2019), evaluation fieldwork is completed (approx. August 2019), the identifiable data has been sent to NHS Digital, HMRC, and DWP), the ONS data coordinator will pseudonymise the locally-collected information using the Trial ID and send it to the ONS data integrator.

The ONS data integrator is responsible for receiving and linking the pseudo-anonymised data, securely holding it, and for providing secure access to the final dataset. The ONS data integrator will receive:

- Locally-collected information from the ONS data coordinator.
- Survey data from NatGen.
- Health usage data from NHS Digital.
- Employment and earnings data from HMRC.
- Benefit information from DWP.

All of the information that the ONS data integrator receives will be pseudo-anonymised using the Trial ID. The ONS data integrator will not hold or have access to any information that would enable any person to be re-identified.

Using the Trial ID, ONS data integrator will link each set of information to create merged datasets for each participant. ONS data integrator will re-pseudonymise these datasets using a new unique identifier (ONS ID). This will provide additional assurance that participants will not be re-identified by anyone accessing the merged datasets who has access to the original Trial ID (i.e. evaluation research teams).

The ONS data integrator will store the pseudo-anonymised merged datasets securely on their servers for up to 3 years after the trial ends to allow for longer-term analysis and/or other health and work-related research to be carried out. [Note – the specific end date of data retention is to be agreed].

During this time the ONS data integrator will retain an index key which can be used to match the original Trial ID against the Safe Haven ID. This will allow for updated health, employment, and benefit information to be sent by NHS Digital, HMRC, and DWP.

[16] Data access for final analysis

The Evaluation Consortium will access the final merged data-set needed for analysis in a secure environment (i.e. via the Secure Research Service). A secure environment can be set up within IES's

offices in London and Brighton. The Consortium can also access the data via ONS's offices in Titchfield and London, if more convenient. The final detail of this arrangement is to be discussed with the ONS.

Any analysis the Consortium carries out on the merged dataset will be completed within these secure environments, and all analysis will be cleared by ADRN staff before releasing the results [Note - ONS to confirm process]. This is to protect against potential statistical disclosure.

[17] Transfer of anonymised data to the UK Data Archive

After 3 years, the ONS will anonymise the data and submit it to the UK Data Archive (this includes removing the Safe Haven ID and double checking the dataset to protect against statistical disclosure). The index key that was held by the ONS to allow for updates to be made to the datasets will be securely destroyed at this point.

Data Sharing Agreements

The Health-led Trials involve numerous organisations and quite complex commissioning arrangements and it is therefore crucial that sensible and straightforward Data Sharing Agreements are in place. To ensure the most appropriate arrangements are in place, advice is being sought from DWP Commercial and Legal experts. The following is the intended Data Sharing Agreements for the Health-led Trials.

1. A Data Sharing Agreement between the CCG and the DWP

This Data Sharing Agreement will outline the data transfers that will take place between the organisations commissioned by the CCG (the IPS Service Provider and the MI Database) and the organisations commissioned by the DWP (the Randomisation Tool Server and the ONS). This includes:

- The IPS Service Provider inputting data into the Randomisation Tool Server
- The Randomisation Tool Server sending data to the MI Database
- The MI Database sending data to the ONS.

The transfers of data between processors of the same commissioner (IPS Service Provider inputting data into the MI database, and the Randomisation Tool Server sending data to the ONS) will be covered separately by agreements / contractual arrangements between the commissioner and the organisations it commissions.

2. A Data Sharing Agreement between DWP-commissioned organisations

This Data Sharing Agreement will outline the data transfers that will take place between the organisations commissioned by the DWP (the Evaluation Consortium organisations, the Randomisation Tool Server, and the ONS). This includes:

- The Randomisation Tool Server sending data to the ONS.
- The transfers of data between the Evaluation Research Teams and the ONS.
- The secure access arrangements to the final merged dataset held by the ONS.

3. Data Sharing Agreements between the ONS and the National Data Owners

These Data Sharing Agreements will be between the DWP and the National Data Owner, and will outline the data transfers that will take place between the ONS (who is the data processor for the DWP) and the National Data Owners. Each Data Sharing Agreement will include:

- The ONS data coordinator sending identifiable data to the National Data Owner.
- The National Data Owner sending pseudo-anonymised data to the ONS data integrator.

On the specific arrangements with each National Data Owner:

- The Data Sharing Agreement with NHS Digital will be developed as part of NHS Digital's Data Access Request Service (DARS) process.
- The Data Sharing Agreement between DWP and HMRC is to be explored post trial go-live.
- Advice will be sought from DWP on the agreements that need to be in place to outline its own administrative data (potentially a Data Sharing Agreement between the DWP and the ONS).

[ONS legal colleagues to review and provide feedback on the Data Sharing Agreement plans]

Opt-out Process and Subject Access Request Process

[ONS to provide feedback on the arrangements / process in place for Troubled Families – ideally I would want the opt-out forms / SARs to be sent to the ONS data coordinator]

References

- Becker, D. R., Swanson, S. J., Bond, G. R., Carlson, L., Flint, L., Smith, G., et al. (2008). Supported Employment Fidelity Scale. Lebanon, NH: Dartmouth Psychiatric Research Center.
- Blakely C, Mayer J, Gottschalk R et al. (1987). The fidelity-adaptation debate: implications for the implementation of public sector social programs. *Am J Community Psychol* 1987; 15: 253–68.
- Burns, T., Catty, J., Becker, T., Drake, R. E., Fioritti, A., Knapp, M., et al. (2007). The effectiveness of supported employment for people with severe mental illness: A randomised controlled trial in six European countries. *The Lancet*.
- Centre for Mental Health (CMH) (2015). IPS FIDELITY SCALE* UK Version 14/05/2015, as of 05/06/2017: <https://www.centreformentalhealth.org.uk/the-ips-fidelity-scale>
- Crowther, R. E., Marshall, M., Bond, G. R., & Huxley, P. (2001). Helping people with severe mental illness to obtain work: systematic review. *Bmj*, 322(7280), 204-208.
- Dartmouth Psychiatric Research Center (2011). Practice Principles of IPS Supported Employment, as of 09/06/2017: <https://www.dartmouth.edu/~ips/page48/page79/files/ips-practice-principles-002880029.pdf>
- Dunn, Olive Jean (1961). "Multiple comparisons among means". *Journal of the American Statistical Association*. 56 (293): 52–64
- Heslin M, Howard L, Leese M, McCrone P, Rice C, Jarrett M, Spokes T, Huxley P, Thornicroft G. (2011). Randomised controlled trial of supported employment in England: 2 year follow-up of the Supported Work and Needs (SWAN) study. *World Psychiatry*, 10(2),132–137.
- Marmot, M. and Bell, R. (2012). Fair society, healthy lives. *Public Health*, 126, pp.S4-S10.
- Marshall, T., Goldberg, R. W., Braude, L., Dougherty, R. H., Daniels, A. S., Ghose, S. S., Preethy, G., & Delphin-Rittmon, M. E. (2014). Supported employment: assessing the evidence. *Psychiatric Services*, 65(1), 16-23.
- McManus S., Meltzer H., Brugha, T., Bebbington, P. and Jenkins, R., (2009) Adult psychiatric morbidity in England, 2007: Results of a household survey. The Health and Social Care Information Centre, Social Care Statistics.
- NHS England, Five Year Forward View for Mental Health for the NHS in England, 2016.
- Modini, M., Tan, L., Brinchmann, B., Wang, M. J., Killackey, E., Glozier, N., Mykletun, A., & Harvey, S. B. (2016). Supported employment for people with severe mental illness: systematic review and meta-analysis of the international evidence. *The British Journal of Psychiatry*, bjp-bp.

- Overall, J.E. and Gorham, D.R., 1962. The brief psychiatric rating scale. *Psychological reports*, 10(3), pp.799-812.
- Priebe S., Huxley P., Knight S., Evans S. (1999). Application and results of the Manchester Short Assessment of Quality of Life (MANSA). *Int J Soc Psychiatry*. 45 (1): 7–12.
- Rinaldi M and Perkins R (2007). Implementing evidence-based supported employment. *Psychiatric Bulletin* 31: 244–249.
- Rogers, P., Petrosino, A., Huebner, T. A., Hacsí T. A. (2000). Program Theory Evaluation: Practice, Promise, and Problems. *New Directions for Evaluation*, Number 87, Fall 2000. DOI: 10.1002/ev.1177.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Shadish, W.R., T.D. Cook, and D.T. Campbell (2002) *Experimental and quasi-experimental designs for generalized causal inference*, Boston, MA, USA: Houghton Mifflin
- Swanson, S. J., Becker, D. R., Drake, R. E., & Merrens, M. R. (2008). *Supported employment: A practical guide for practitioners and supervisors*. Kearney, NE: Morris Publishing.
- Twamley, E. W., Padin, D. S., Bayne, K. S., Narvaez, J. M., Williams, R. E., & Jeste, D. V. (2005). Work rehabilitation for middle-aged and older people with schizophrenia, *Journal of Nervous & Mental Disease*, 193, 596–601.
- van Rijn, R. M., Carlier, B. E., Schuring, M., & Burdorf, A. (2016). 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. *Occupational and environmental medicine*, oemed-2015.
- Zigmond, AS. & Snaith, RP. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica*. 67 (6): 361–370