



Department for Levelling Up,
Housing & Communities

Evaluation of the Changing Futures programme: feasibility study

Final Report

April 2023





© Crown copyright, 2023

Copyright in the typographical arrangement rests with the Crown.

You may re-use this information (not including logos) free of charge in any format or medium, under the terms of the Open Government Licence. To view this licence visit <http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/>

This document/publication is also available on our website at www.gov.uk/dluhc

If you have any enquiries regarding this document/publication, complete the form at <http://forms.communities.gov.uk/> or write to us at:

Department for Levelling Up, Housing and Communities
Fry Building
2 Marsham Street
London
SW1P 4DF
Telephone: 030 3444 0000

For all our latest news and updates follow us on Twitter: <https://twitter.com/luhc>

April 2023

Contents

| | |
|--|-----------|
| Acknowledgements | i |
| Executive summary | ii |
| 1 Introduction | 1 |
| 1.1 Background to this report | 1 |
| 1.2 Key research questions | 1 |
| 1.3 Approach to the feasibility study | 2 |
| 1.4 Approach to reporting | 5 |
| 1.5 Report structure | 5 |
| 2 Individual-level outcomes | 6 |
| 1.1 Framing outcomes and links across outcome levels | 6 |
| 2.1 Summary of suggested priority outcomes | 6 |
| 2.2 Introduction | 8 |
| 2.3 Category 1: Changes in service use | 9 |
| 2.4 Category 2: Changes in experiences of support | 12 |
| 2.5 Category 3: Changes in wellbeing, health, behaviours and circumstances | 12 |
| 3 Service-level outcomes | 19 |
| 3.1 Summary of suggested priority outcomes | 19 |
| 3.2 Introduction | 21 |
| 3.3 Category 1: Changes in how support is delivered | 21 |
| 3.4 Category 2: Changes in access and engagement | 22 |
| 3.5 Category 3: Changes within the workforce delivering support | 22 |
| 3.6 Measurement approaches | 23 |
| 4 System-level outcomes | 24 |
| 4.1 Summary of suggested priority outcomes | 24 |
| 4.2 Introduction | 27 |
| 4.3 Changes to be prioritised | 29 |
| | 3 |

| | | |
|----------|---|-----------|
| 4.4 | Longer-term changes of lower priority for the evaluation | 30 |
| 4.5 | Measurement approaches | 30 |
| 5 | Measuring impact | 32 |
| 5.1 | Summary of potential approaches to impact evaluation | 32 |
| 5.2 | Introduction | 37 |
| 5.3 | Role for theory-based, mixed methods and qualitative approaches | 39 |
| 5.4 | Level A: Intervention-level impact for individuals receiving support directly from interventions | 41 |
| 5.5 | Level B: Local population-level impact of the whole local programme of work (not just specific interventions) | 45 |
| 5.6 | Level C: System-level impact of the funding/programme model | 56 |
| 5.7 | Economic impact | 60 |
| 6 | Process evaluation | 66 |
| 6.1 | Summary of findings | 66 |
| 6.2 | Role of the process evaluation | 66 |
| 6.3 | The benefits of formative or iterative approaches | 67 |
| 6.4 | Scope for closer focus on specific themes or local areas | 67 |
| 6.5 | Key process evaluation questions | 68 |
| 7 | Engagement of stakeholders | 70 |
| 7.1 | Summary of findings | 70 |
| 7.2 | Introduction | 70 |
| 7.3 | Co-production | 70 |
| 7.4 | Engaging people using services in the evaluation | 74 |
| 7.5 | Engaging local programme partners | 77 |
| 7.6 | Engaging evaluation audiences | 79 |
| 8 | Conclusions | 81 |
| | Appendix A: Summary of outcome measurement tools identified during feasibility study | 82 |
| | Appendix B: Summary of findings on the approaches to conceptualising recovery | 90 |

| | |
|---|-----------|
| Appendix C: Examples of evidence sources on “what works” in supporting people experiencing multiple disadvantage | 93 |
| Appendix D: Evidence review protocol and bibliography | 95 |

Acknowledgements

Thank you to all of the stakeholders who shared their views and expertise as part of the consultation element for this feasibility study. Thank you also to the NECG members who co-facilitated the focus groups with people with lived experience of multiple disadvantage.

Executive summary

Introduction

This report presents the findings of a feasibility study to inform the upcoming evaluation of the [Changing Futures programme](#), which aims to improve outcomes for adults experiencing multiple disadvantage¹ by testing a more joined up, person-centred approach in local areas and across government. The feasibility study was conducted by Cordis Bright, CFE Research, Revolving Doors Agency and the National Expert Citizens Group (NECG) and the final report was prepared in March 2021. The study was commissioned by the Department for Levelling up, Housing and Communities (DLUHC).²

Priority outcomes

There was a relatively strong consensus during the feasibility study consultation process about the types of outcomes which should form the focus of the Changing Futures programme and – by extension – its evaluation. These were identified by both people with lived experience of multiple disadvantage and those with experience of programmes to support them. They also resonate with outcomes identified via the evidence review. (Please see Figure 2 for an overview of suggested priority outcomes and broad measurement approaches).

It will not be feasible to focus on all of the potential programme outcomes within the resources for the evaluation and for programme delivery. Consultation with local areas will be required during the set-up stages to determine which outcomes are most meaningful in the context of their local programmes of work, and the scale and nature of outcomes measurement which will be feasible. At this stage, our recommendations are to focus on:

- **Individual-level outcomes**³ which focus on early outcomes of importance to the concept of recovery, experiences of support, engagement with support and (if economic impact evaluation is a priority) changes in service use.
- **Service-level outcomes** which focus on how support is delivered, changes in access and engagement and changes within the workforce delivering support.

¹ The Changing Futures programme considers that people are experiencing multiple disadvantage if they experience three or more of the following issues: homelessness, substance misuse, mental health issues, domestic abuse and contact with the criminal justice system. The explicit inclusion of domestic abuse represents a change from previous definitions of multiple disadvantage (such as those used by the Fulfilling Lives programme and Making Every Adult Matter coalition), which have focused on the other four issues included in the Changing Futures definition.

² When this report was commissioned, the Department was known as the Ministry of Housing, Communities and Local Government (MHCLG). References to MHCLG have been retained throughout this report.

³ The inclusion of specific individual-level outcomes is based on the assumption that meaningful outcomes frameworks of individually-determined outcomes (discussed in section 2.5.1) will not be desirable to MHCLG and partners.

- **System-level outcomes** which focus on changes to multi-agency partnerships and working and changes to system learning and development processes.

Outcomes at these levels might be achieved in one or more of the areas participating in the Changing Futures programme. Alongside this, a key purpose of the programme is to understand whether and how central funding and resource can be used to catalyse changes at a local and national level, and to generate learning about how to do so. Therefore the evaluation should also focus on generating and capturing learning about the programme's success at creating change, and exploring key mechanisms for change.

Challenges for the evaluation

This important evaluation is likely to be challenging to deliver, particularly from an impact evaluation perspective. There is a range of reasons for this, but the chief of them are:

- The multi-site nature of the programme, with flexibility for local areas to design and implement local programmes of work that they believe fit their local contexts. This means that interventions and programmes of work will differ in each local area.
- The desire for robust evaluation of impact at both individual- and system-level, and the possible tensions between the approaches required for these two different strands of impact evaluation.
- The personal nature of priority outcomes for individuals receiving support, which runs counter to the idea of an evaluation focusing on pre-determined individual-level outcomes.
- The multiplicity of individual-level outcomes which might be affected and the absence of a clear lead variable for understanding impact via experimental or quasi-experimental methods.
- The absence of an obvious counter-factual control group for experimental and quasi-experimental methods.
- The complex and “noisy” systems into which the programme is being introduced, which will all differ in structure, relationships and context. This poses challenges for understanding and attributing impact via traditional evaluation approaches.
- The innovative nature of interventions and related activities, where an emphasis is placed on learning by doing, meaning that interventions and activities are unlikely to remain fixed for the duration of the programme. This also means that there are limited examples of previous similar evaluations.
- The relatively short timescales for the programme and intervention, making it important to determine which outcomes are most likely to be achievable and demonstrable in the timescales and to be realistic about evaluation approaches that can be set up and delivered in the two-year window.

It may not be feasible to resolve all of these challenges within the resource available to the evaluation. As a result, MHCLG and partners will need to decide on the highest priorities for the evaluation. This will help to identify the most relevant approaches to pursue so that

evaluation resource can then be focused on solutions to challenges for these specific approaches and the methods they require.

Approaches to impact evaluation

There are three key levels at which the impact of the Changing Futures programme might usefully be explored by the evaluation, as follows:

- Level A: Intervention-level impact.
- Level B: Local population-level impact.
- Level C: Local system-level impact.

These levels are outlined in Figure 1, alongside the feasibility study findings of the most feasible and robust approach to impact evaluation for each level.

Impact evaluation at levels A and B would involve a greater focus on quantitative measurement of harder outcomes experienced by individuals and a greater emphasis on the direct attribution of these outcomes to aspects of the Changing Futures programme. Impact evaluation at level C would draw more heavily on complexity theory and on qualitative and participatory approaches to understanding systems change and its relationship to the Changing Futures programme.

Exploration of service-level outcomes could be incorporated at any of the three levels.

Impact evaluation at any of these levels could be delivered across all areas involved in the Changing Futures programme, but could also focus on a smaller number of areas of interventions of particularly high interest. This might enable more resource-intensive approaches to be delivered in selected areas.

It is our recommendation that the evaluation focuses as a minimum on level C, the impact at local system-level. This is because one of the key distinguishing characteristics of Changing Futures is the central focus on enabling systems change via an injection of funding and support. Focusing on an impact evaluation at this level may be the most effective way to produce a useful evaluation at national level, given the bespoke nature of local programmes and implementation.

If resource allows, we then recommend pursuing an evaluation of level A, intervention-level impact, as a second priority. If well designed, this could add value in improving the evidence base for the impact of specific interventions in different systems and contexts. However, many of the challenges outlined in the previous section are particularly applicable to developing and delivering an experimental or quasi-experimental evaluation design to understand individual-level impact, which also builds in approaches to unpick the nature of and context for interventions in more detail in order to say more about which elements of an intervention are important in determining impact.

Figure 1: Summary of different levels and approaches to impact evaluation

| | Level A: Intervention-level impact | Level B: Local population-level impact | Level C: Local system-level impact |
|---|--|--|---|
| Main focus of impact evaluation | The impact for individuals receiving support directly from interventions funded or catalysed by the Changing Futures programme | The impact of the whole local programme of work relating to the Changing Futures programme (not just specific interventions) for all individuals experiencing multiple disadvantage in the local area (not just those receiving support from interventions) | The impact of the Changing Futures programme model on systems change activity and systems in the local areas included in the programme |
| Most robust of the feasible approach(es) | <p>Interrupted time series / change over time data study with no comparison group.</p> <p>Data for beneficiaries at two or more time points is compared in order to understand change in key indicators of outcomes.</p> <p>For change over time studies, data for a baseline time point is compared to data at one or more later time point (typically before/at the start of an intervention and after an intervention).</p> <p>For interrupted time series, data at multiple time points before an</p> | <p>The feasibility or otherwise of these approaches would need to be established by further exploration of the likelihood of any changes for the local multiply disadvantaged population being visible in the selected lead variables.</p> <ul style="list-style-type: none"> • Difference in difference with either one or two lead variables or synthetic controls. • Qualitative comparative analysis. <p>This could be delivered as part of a mixed methods approach, which could include qualitative approaches and additional quantitative methods.</p> | <p>Qualitative and participatory approaches, to be determined by selected evaluators. These could possibly be combined with quantitative approaches using frameworks for assessing system maturity or specific changes and outputs as proxies for some outcomes.</p> |

| | | | |
|---------------------------|---|--|--|
| | <p>intervention and multiple points after an intervention is required⁴.</p> <p>The feasibility of these approaches should be confirmed once more when the specific outcomes and indicators of interest to the evaluation have been finalised.</p> <p>The data study element could be delivered as part of a mixed methods approach, which could include qualitative approaches and additional quantitative methods.</p> | | |
| Our recommendation | We recommend pursuing an evaluation at this level as the second priority. | We do not recommend pursuing an evaluation at this level, unless further programme and evaluation planning establishes a clear lead variable at population level, in which change is likely to be visible given the scale of the programme activity and the evaluation timescales. | We recommend pursuing an evaluation at this level as the first priority. |

⁴ As a result, an interrupted time series design would likely require the use of administrative data as this is the only type of data likely to be available for multiple time points prior to an intervention.

It would be possible to conduct impact evaluation at one or more of these levels, depending on the resource available for the evaluation. The decision on which level(s) to include depends on whether MHCLG and partners wish to place greater emphasis on evaluating outcomes at individual-level (attributable directly to interventions or otherwise) or at system-level (contributed to by the programme funding and model). There were divergent views amongst consulted stakeholders about which of the levels was most important and useful, depending on stakeholders' own methodological expertise and on their views on the priority focus for the evaluation.

Neither the consultation nor the evidence review provided definitive findings on the impact evaluation methods which would be most suitable at any of the three levels. Final approaches and methods would need to be determined with the selected evaluators, in consultation with local areas.

If impact evaluation is required at multiple levels, there is an argument for commissioning these separately as they will require evaluators with different specialist skillsets. In this case, careful thought will be required as to how to integrate the findings from both evaluations to avoid findings which fail to explore the important connections between the impact of funded interventions and the maturity and features of systems in which they are delivered.

Economic evaluation

The decision on which level to focus on may also be determined by the extent to which economic evaluation is a priority. Within the programme and evaluation timescales, it is most feasible to examine economic impact at the level of impact for individuals receiving support directly from interventions (level A). At a cohort level, longer-term fiscal impact may be seen over a longer period of time.

In all of the quantitative approaches to measuring intervention-level impact listed in section 5.4, a unit cost can be attached to inputs, outputs and outcomes to gain an understanding of economic impact. For example, if the lead variable is hospital admissions a tariff cost can be used. Social outcomes, such as change in wellbeing, can also be monetised, as can health outcomes (using QALYs).

Once unit costs are applied, economic impact can be assessed through a range of approaches. However, the same challenges to measuring impact as set out earlier in this chapter apply to the assessment of economic impact. For this reason, approaches to economic impact assessment that rely on a counterfactual were not widely recommended by stakeholders. A more proportionate and realistic approach may be to assess costs and benefits based on the beneficiaries' service use before and after service use, and for this assessment to focus on changes in a limited set of outcome areas. However, this approach is still likely to be relatively resource-intensive and challenges in generating and collating high-quality service use data might affect the robustness of the analysis and findings. quality service use data might affect the robustness of the analysis and findings. Further investigation would be required by the appointed evaluators to determine whether likely levels of validity justify the use of evaluation resource.

Importance of theory-based methods

Within any of these levels of impact evaluation and irrespective of the specific approaches taken within them, we recommend a theory-based, mixed methods overall approach to understanding impact, including qualitative methods. These are particularly important in exploring changes taking place in complex systems. A central theory of change has already been drafted and this will be a helpful starting point for theory-based methods, which must also remain open to identifying other changes which might emerge as important, and particularly to unplanned or unintended changes. Any theory of change will need to be localised too, in order to enable evaluators to work with local areas to explore systems change activities and their results.

Role of process evaluation

A strong process evaluation will also be crucial to understanding how different aspects of the programme have been delivered and why different approaches and interventions have been successful or less successful. It can complement theory-based impact evaluation approaches and provide a more in-depth understanding of the detail of implementation and how this links to impact, providing supporting information to explore the mechanisms and context for change at individual, service and system level. Formative and iterative approaches may be useful, especially at the systems level. There is also scope to focus in on specific themes or areas of interest, in order to unpick these in further detail.

Co-production and stakeholder engagement

The quality of the evaluation is likely to be improved if it is co-produced at all stages with people with lived experience of multiple disadvantage. This is a resource-intensive process which requires adequate investment if it is to be delivered meaningfully and safely.

Whilst this lived experience input is valuable throughout the evaluation process, peer researcher models may be particularly valuable in improving the engagement of people using services in the evaluation. In addition to involving peer researchers, other key mechanisms for promoting the engagement of people using services include transparency about the purpose and nature of involvement, offering a range of ways to engage, using qualitative consultation methods and taking a tailored, strengths-based approach.

Effectively engaging local programme partners in the evaluation is crucial to its success. It is therefore important to generate buy-in by ensuring the evaluation is relevant to local partners and by being realistic about the extent to which they can support the evaluation. Building early relationships centred on the development of local theories of change may be effective, and will also support the theory-based approaches which are important to evaluations in complex systems.

Ensuring dedicated local resource is available for evaluation activities is crucial, particularly in relation to the collation and sharing of individual-level data by someone with experience and expertise in doing so. This element of the evaluation is likely to be challenging and may also benefit from dedicated central resource from MHCLG. This might include: central collation of administrative data for indicators where data is available centrally; direct support to local areas to establish effective consent and data collection processes and negotiate access to data which is held locally.

It is also important to find ways to ensure that the volume of data is manageable. The two main approaches to this are reducing the number of variables included in the dataset (by focusing in on a smaller number of priority outcomes and indicators) or reducing the number of individuals included in the dataset (by selecting only a sample of clients from each local area to be included in this element of the evaluation). Finally, the quality of data collection is likely to be higher if the data is also of use to the local projects responsible for collating it. For example, the Fulfilling Lives quarterly data dashboards have been helpful in providing an accessible tool for projects to explore their data and compare it with others.

Overview of priority outcomes and measurement approaches

Figure 2: Summary of suggested priority outcomes for the programme and evaluation

Individual-level outcomes

| Priority outcomes | Measurement approaches |
|--|--|
| <p>Category 1: Changes in service use</p> <ul style="list-style-type: none"> a. Increased engagement with specialist support services. b. Increased engagement with planned support. c. Changes in the use of emergency or crisis support and unplanned service use. | <ul style="list-style-type: none"> • Linked administrative datasets. • Area-level/local population-level indicators. • Aggregated data on engagement with services. |
| <p>Category 2: Changes in experiences of support</p> <ul style="list-style-type: none"> a. Feeling better informed about support and the options available to them. b. Having increased freedom to make choices about support and goals. c. Greater trust of services, or individual people working within them. d. Building and maintaining more open and honest relationships with a consistent worker. | <ul style="list-style-type: none"> • Qualitative consultation with people receiving support and staff members/ volunteers delivering support. |
| <p>Category 3a: Changes in wellbeing, health, behaviour and circumstances – shorter-term</p> <ul style="list-style-type: none"> a. Feeling more hopeful about the future. b. Supportive relationships with trusted workers in services. c. Feeling a greater sense of purpose. d. Having more opportunities to take part in meaningful activities. e. Increased sense of agency and control. f. More stable accommodation. g. Improved financial situation (including access to appropriate benefits). | <ul style="list-style-type: none"> • Standardised and validated tools for self-completion or completion with workers. • Qualitative consultation with people receiving support and staff members/ volunteers delivering support. |

| | |
|--|--|
| <p>h. Access to appropriate medication (for physical health, mental health and/or substance misuse treatment).</p> <p>Category 3b: Changes in wellbeing, health, behaviour and circumstances – longer-term</p> <p>a. Improved or maintained relationships with supportive friends.</p> <p>b. Improved or maintained relationships with children and/or other family members.</p> <p>c. Feeling safer.</p> <p>d. Feeling less lonely and/or isolated.</p> <p>e. Improved education and employment outcomes or employability.</p> <p>f. Improved wellbeing.</p> | |
|--|--|

Service-level outcomes

| Priority outcomes | Measurement approaches |
|--|---|
| <p>Category 1: Changes in how support is delivered</p> <p>a. Support is more trauma-informed and/or psychologically-informed.</p> <p>b. Support is more flexible, person-centred and strengths-based.</p> <p>c. Support is more holistic.</p> | <ul style="list-style-type: none"> • Fidelity tools designed to support service improvement. • Standardised and validated tools on worker-client relationships. • Bespoke surveys of staff. • Qualitative consultation with staff members/ volunteers delivering support, their managers and individuals receiving support. |
| <p>Category 2: Changes in access and engagement</p> <p>a. Relevant support is offered at the time and place where people need it.</p> <p>b. Services persevere and are creative in engaging people.</p> | <ul style="list-style-type: none"> • Change over time in output measures. |

| Priority outcomes | Measurement approaches |
|---|---|
| <ul style="list-style-type: none"> c. Fewer people are declined support without an alternative support mechanism in place. d. Fewer people drop out of support, including maintaining support during key transitions. | <ul style="list-style-type: none"> • Qualitative consultation with staff members/ volunteers delivering support, their managers and individuals receiving support. |
| <p>Category 3: Changes within the workforce delivering support</p> <ul style="list-style-type: none"> a. People delivering support have a greater and shared understanding of multiple disadvantage and its constituent parts. b. People delivering support have a better awareness of the range of services and support available and how to enable people to access them. c. There are more support work roles for people with lived experience of multiple disadvantage. d. People delivering support have a greater sense of efficacy and autonomy. e. People delivering support receive appropriate support themselves, resulting in reduced staff burn-out. | <ul style="list-style-type: none"> • Bespoke surveys of staff. • Standardised and validated tools on staff burn-out. • Qualitative consultation with staff members/ volunteers delivering support, their managers and individuals receiving support. |

System-level outcomes

| Priority outcomes | Measurement approaches |
|---|--|
| <p>Category 1: Changes to multi-agency partnerships and working</p> <ul style="list-style-type: none"> a. Increased collaboration and investment in the agenda by systems actors, including people commissioning, designing, delivering and using services. b. Greater diversity of voices in partnerships, including more effective involvement of people with lived experience of multiple disadvantage. c. More effective information sharing, at both operational and strategic levels. | <ul style="list-style-type: none"> • Qualitative and participatory approaches. • Use of existing tools to audit partnerships and system maturity. • Use of outputs as a proxy for outcomes. |

-
- d. Better coordinated support with clear pathways into services.

Category 2: Changes to system learning and development processes

- a. Partnerships and system actors are open to learning and improvement and have effective mechanisms in place to do so.
System barriers are addressed strategically and removed.

Category 3: Changes to commissioning structures and practices

- a. Commissioning is more integrated and less siloed.
- b. Commissioning and monitoring of services focuses more on shared outcomes and on quality.
Commissioning structures and processes incentivise service activities and attitudes that work for people experiencing multiple disadvantage.

Category 4: Improved systems of support

- a. Better identification of people experiencing multiple disadvantage.
Stakeholders highlighted in particular that this includes identifying people from groups who might currently be less well-identified, such as women and people with uncertain immigration status.
 - b. Earlier intervention and more preventative work.
 - c. More support is available that meets the needs of people experiencing multiple disadvantage, including different support and approaches to suit different people.
Reduced stigma and judgement within services of people experiencing multiple disadvantage.
-

1 Introduction

1.1 Background to this report

This report presents the findings of a feasibility study to inform the upcoming evaluation of the [Changing Futures programme](#), which aims to improve outcomes for adults experiencing multiple disadvantage⁵ by testing a more joined up, person-centred approach in local areas and across government.

The Changing Futures programme is a cross-government programme led by the Ministry of Housing, Communities and Local Government (MHCLG). It received £46m of funding from the [Shared Outcomes Fund](#)⁶ in order to undertake a pilot in up to 15 local areas to deliver new interventions and systems change to better support people experiencing multiple disadvantage. The pilot will run from late 2020-21 until 2022-23.

The feasibility study was commissioned by MHCLG and delivered in partnership by Cordis Bright, CFE Research, Revolving Doors Agency and the National Expert Citizens Group (NECG). Its purpose was to establish an appropriate evaluation design to enable MHCLG and partners to understand whether the Changing Futures programme is effective. It was delivered in a nine-week period from January to March 2021.

1.2 Key research questions

The feasibility study aims to provide information relating to the following key research questions:

1. What real-world outcomes should the programme seek to influence, for clients who engage with the programme?
2. What data is available which can be used as metrics to measure the effect of the programme on the aforementioned real-world outcomes?
3. One aim of the programme is to affect a systemic change, joining up local services to work together, sharing information. What evaluative approaches can be taken to measure this outcome?
4. What other data will need to be collected to support the evaluation and the programme, and what data systems will need to be in place in Local Authorities to facilitate this?

⁵ The Changing Futures programme considers that people are experiencing multiple disadvantage if they experience three or more of the following issues: homelessness, substance misuse, mental health issues, domestic abuse and contact with the criminal justice system. The explicit inclusion of domestic abuse represents a change from previous definitions of multiple disadvantage (such as those used by the Fulfilling Lives programme and Making Every Adult Matter coalition), which have focused on the other four issues included in the Changing Futures definition.

⁶ Since 2020 the Shared Outcomes Fund has provided a total of £400m to fund pilot projects to test innovative ways of working across the public sector, with an emphasis on thorough plans for evaluation.

5. What challenges exist in designing an impact evaluation with this client group, in particular with consideration to selecting a treatment and comparison group, and how can these challenges be overcome?

1.3 Approach to the feasibility study

Methodology

Figure 3 outlines the methodology for the feasibility study. The study was jointly delivered by Cordis Bright, CFE Research, Revolving Doors Agency and NECG but partners took the lead on different elements, as detailed in Figure 3. As a result of the tight timescales for the feasibility study, steps 3-5 were delivered simultaneously.

The findings from across steps 1-5 were synthesised to determine the key findings of the feasibility study and to consider their implications for evaluation design and implementation.

Figure 3: Feasibility study methodology

| Method | Description |
|--|---|
| <p>Step 1: Discussions with MHCLG and review of key documentation to develop understanding of the Changing Futures programme.</p> | <p>The feasibility study team met with MHCLG prior to starting the feasibility study, in order to better understand the Changing Futures programme and priorities for the evaluation and feasibility study. We also reviewed key documentation relating to the programme. We maintained regular contact with MHCLG colleagues to discuss developments for the programme and feasibility study, with the aim of ensuring that the study was as relevant and useful as possible.</p> |
| <p>Step 2: Internal workshops to collate experience within the feasibility study team.</p> | <p>Workshop A: Facilitated by Revolving Doors Agency with 9 NECG members, identifying shared insights into consultation with people experiencing multiple disadvantage, experiences of research and evaluation and potential challenges.</p> <p>Workshop B: Involving Cordis Bright, CFE and Revolving Doors Agency teams and 2 NECG members who had attended workshop A, drawing together insights on: evaluation approaches and challenges, priority outcomes, outcome measures, data availability.</p> |
| <p>Step 3: Focussed evidence review</p> <p>Lead partner: CFE Research</p> | <p>Focussed evidence review to summarise the evidence base on: a.) effective or previously-used evaluation approaches for programmes addressing multiple disadvantage (or programmes involving other marginalised or excluded groups) and/or systems change; b.) previous or recommended outcomes of focus for these programmes/evaluations (at individual, service and system level); and c.) outcomes measurement approaches and data availability. We agreed the search protocol and initial bibliography with MHCLG and built upon this using citation searches within initial sources and by including further sources</p> |

| | |
|--|--|
| | <p>recommended by stakeholders consulted in step 4 below. The search protocol and bibliography of reviewed evidence are included in Appendix D.</p> |
| <p>Step 4: Consultation with policy, programme and evaluation stakeholders</p> <p>Lead partner: Cordis Bright</p> | <p>Semi-structured interviews with 44 stakeholders with insight into one or more of the following⁷:</p> <ul style="list-style-type: none"> • The Changing Futures programme and its priority intended outcomes. • The types of evidence and approaches of interest to key government departments. • Design and delivery of other programmes aiming to improve support and outcomes for people experiencing multiple disadvantage. • Evaluation of programmes focused on multiple disadvantage or its constituent parts. • Specific research and evaluation methodologies and frameworks, such as trials and quasi-experimental designs or complexity theory. • Current services and systems which might be the focus of the Changing Futures programme, such as the criminal justice system or domestic abuse support provision. <p>We agreed the stakeholder list and topic guide in advance with MHCLG.</p> |
| <p>Step 5: Consultation with people with lived experience of multiple disadvantage</p> <p>Lead partner: Revolving Doors Agency and NECG</p> | <p>Six focus groups involving a total of 38 people with lived experience of multiple disadvantage, including members of the NECG, Revolving Doors Lived Experience Members and people involved in Fulfilling Lives projects (either as experts by experience or as people using services)⁸. These were co-facilitated by Revolving Doors Agency and NECG members and focused primarily on views on: a.) priority outcomes for the Changing Futures programme; b.) effective practice in co-producing evaluations with people with lived experience; and c.) effective practice in involving people using services in evaluations. Again we agreed the topic guide in advance with MHCLG.</p> |

⁷ This mainly included 1-to-1 interviews but there were also a number of small group interviews with 2 or 3 stakeholders from the same organisation.

⁸ More information on the gender and ethnicity of focus group participants, the areas in England in which they were based, and the approach taken in the focus groups is available in the full account of the lived experience consultation findings, which is available under separate cover.

Challenges for the feasibility study

We sought to deliver a feasibility study that was as robust as possible within the timescales and resource available. Key challenges are outlined below.

Specific information about the Changing Futures programme:

The Changing Futures programme prospectus and internal information shared by MHCLG were very useful in helping the feasibility study team to understand the programme's nature and overarching priorities. However, the expression of interest and shortlisting process for local areas to be included in the Changing Futures programme took place at the same time as the feasibility study. This meant that we only had access to information about shortlisted local areas and their proposals during the latter phases of report drafting. Earlier access to this information may have enabled us to take greater account of contextual factors, any previous work relating to multiple disadvantage and relevant systems change, and the specific planned priorities, activities and target cohorts in local areas.

In addition, a central theory of change for the programme was being drafted towards the end of the feasibility study period. This was used to inform the focus of the findings in this report but was not available to inform the development of the topic guides or evidence review protocol. This meant that we aimed to explore the broad range of outcomes which might form the focus of the programme and evaluation, rather than honing in on specific, pre-determined programme outcomes. The latter approach would have allowed for a more systematic exploration of the evidence on measuring a smaller number of possible outcomes.

Determining the best use of resource for the evidence review

Within the timescales and available resource it was not possible to review all of the literature which might usefully inform the evaluation of Changing Futures. For example, the specific priority outcomes for the programme and evaluation priorities were still under development and it was therefore not possible to focus in detail on measurement approaches for a small number of outcomes. Equally, the preferred balance was yet to be determined between evaluating individual-level outcomes, which might primarily result from local interventions funded by Changing Futures, and examining system-level outcomes, which might result from the programme's focus on funding systems change activity. This precluded the option of focussing in greater detail on the benefits and challenges of a smaller number of approaches potentially suited to a specific evaluation focus. Instead, it was necessary to examine a wide range of evaluation approaches and methods in as much depth as possible within the resource and timescales available.

Therefore, we took a focused approach in the evidence review, beginning with an initial bibliography agreed with MHCLG supplemented with literature identified from string searches on key terms relating to multiple disadvantage and evaluation. We then expanded to include additional sources suggested by stakeholders during consultation and purposive searches to address gaps in the initial bibliography.

Consulting with key stakeholders within the available timescales

We aimed to consult with stakeholders with wide-ranging relevant expertise that might inform the feasibility study, but there are likely to be many other stakeholders whose views

and experiences could have been equally valuable. The majority of stakeholders who were invited to participate in the feasibility study were able to take part within the study timescales. However, in a small number of cases we were unable to consult with people who had been originally approached or to arrange follow-up consultation resulting from snowball sampling.

1.4 Approach to reporting

This report synthesises findings from the evidence review, consultation with programme and evaluation stakeholders and consultation with people with lived experience of multiple disadvantage.

1.5 Report structure

The remainder of this report is structured as follows:

- **Chapters 2 to 4** detail suggested priority outcomes at the three main levels on which the Changing Futures programme focuses – individual level, service level and system level. These chapters also discuss possible measurement approaches for the outcomes.
- **Chapter 5** discusses approaches to impact evaluation. It proposes methods which might be feasible and suitable for evaluating impact of specific interventions/local programmes of work for individuals and impact of the programme on systems. It also outlines possible approaches to understanding economic impact.
- **Chapter 6** summarises findings on the role of process evaluation.
- **Chapter 7** outlines key considerations for engaging stakeholders in the evaluation. This includes discussion of co-producing the evaluation with people with lived experience of multiple disadvantage, engaging people using services in the evaluation, and engaging key local partners in the evaluation.
- **Chapter 8** presents our conclusions on the key challenges in designing and delivering an evaluation of the Changing Futures programme.
- **Appendices A-D** present summary findings from the evidence review, along with the search protocol and bibliography.

2 Individual-level outcomes

1.1 Framing outcomes and links across outcome levels

In some cases it is difficult to determine the level at which different outcomes should best be framed. For example:

- Changes in engagement and experiences of support are outcomes for individuals but have corresponding outcomes at a service or system-level, such as services being more accessible and better able to engage people.
- Changes in support flexibility can take place at service-level but increased flexibility across the system might also constitute a system-level outcome.

We have based judgements on the level at which to frame outcomes on the draft theory of change for the Changing Futures programme and on the level at which data is most likely to be collected.

There is a logic linking activities and outcomes within and between different outcome levels, as outlined in the draft theory of change. However, it is important to note the possibility that deciding to prioritise some outcomes at programme or evaluation level might influence programme delivery in local areas in ways that detract from the ability to achieve other outcomes. For example, if more stable accommodation is prioritised as a key measure of success, this could result in resource and activity being focused on reducing rough sleeping in the short-term, rather than on partnership working to co-produce longer-term solutions which prevent people from becoming homeless or address systemic barriers to effective support for people experiencing multiple disadvantage.

2.1 Summary of suggested priority outcomes

Figure 4 outlines a range of potential priority individual-level outcomes for the programme and the evaluation, along with possible measurement approaches⁹. Decisions on the impact evaluation (discussed in Chapter 5) will determine whether it is necessary to prioritise outcomes b and c in category 1.

It would likely be infeasible to include all outcomes in categories 3a and 3b if relying on data collection by local services with their clients, because this would exceed the resources available locally and require too much direct work with clients on data collection. However, they could still be explored in qualitative components of the evaluation. It will therefore be important to focus in on a smaller number of category 3 outcomes to be supported by local data collection, based on priorities for MHCLG and partners and on the relevance of the outcomes to the specific local programmes of work being undertaken.

⁹ These suggested outcomes are based on the assumption that meaningful outcomes frameworks of individually-determined outcomes (discussed in section 2.5.1) will not be desirable to MHCLG and partners.

Figure 4: Summary of suggested priority individual-level outcomes

| Priority outcomes | Measurement approaches |
|---|--|
| <p>Category 1: Changes in service use</p> <ul style="list-style-type: none"> a. Increased engagement with specialist support services. b. Increased engagement with planned support. c. Changes in the use of emergency or crisis support and unplanned service use. | <ul style="list-style-type: none"> • Linked administrative datasets. • Area-level/local population-level indicators. • Aggregated data on engagement with services. |
| <p>Category 2: Changes in experiences of support</p> <ul style="list-style-type: none"> a. Feeling better informed about support and the options available to them. b. Having increased freedom to make choices about support and goals. c. Greater trust of services, or individual people working within them. d. Building and maintaining more open and honest relationships with a consistent worker. | <ul style="list-style-type: none"> • Qualitative consultation with people receiving support and staff members/ volunteers delivering support. |
| <p>Category 3a: Changes in wellbeing, health, behaviour and circumstances – shorter-term</p> <ul style="list-style-type: none"> a. Feeling more hopeful about the future. b. Supportive relationships with trusted workers in services. c. Feeling a greater sense of purpose. d. Having more opportunities to take part in meaningful activities. e. Increased sense of agency and control. f. More stable accommodation. g. Improved financial situation (including access to appropriate benefits). h. Access to appropriate medication (for physical health, mental health and/or substance use treatment). <p>Category 3b: Changes in wellbeing, health, behaviour and circumstances – longer-term</p> <ul style="list-style-type: none"> a. Improved or maintained relationships with supportive friends. b. Improved or maintained relationships with children and/or other family members. c. Feeling safer. d. Feeling less lonely and/or isolated. e. Improved education and employment outcomes or employability. f. Improved wellbeing. | <ul style="list-style-type: none"> • Standardised and validated tools for self-completion or completion with workers. • Qualitative consultation with people receiving support and staff members/ volunteers delivering support. |

2.2 Introduction

Evaluating progress against individual-level outcomes is of high importance because the primary aim of the programme is to improve outcomes for adults experiencing multiple disadvantage. In this section we have focused on individual-level outcomes which the consultation and evidence review indicate are frequently important to people experiencing multiple disadvantage and/or to those commissioning, designing and delivering services, and/or which have been the focus of previous evaluations of similar programmes. These fall into three overarching areas:

- Changes in service use (Category 1).
- Changes in experiences of support (Category 2).
- Changes in wellbeing, health, behaviours and circumstances (Category 3).

We have also referred to the draft theory of change for the programme, to understand the extent to which these outcomes are likely to be applicable. Before final selection of priority outcomes for the evaluation, it will be important to confirm that they are likely to be impacted by the programmes of work in funded local areas and that they are of relevance and importance to local partners involved in the Changing Futures programme.

Although we have focused on commonly-mentioned outcomes, there is an important argument for defining priority outcomes on a case-by case basis. The programme emphasises flexible, person-centred and asset-based support. In order to understand whether this type of personalised support model is effective, ideally priority outcomes would be determined with each individual and then progress measured against these. People with lived experience of multiple disadvantage who contributed to this feasibility study highlighted the importance of individuals setting their own priority outcomes for support, as did stakeholders with experience of designing, delivering and evaluating programmes for people experiencing multiple disadvantage. It might be possible to develop meaningful outcomes frameworks to enable the measurement of individually-determined outcomes, although this would pose some specific challenges in terms of collating and aggregating data within and between areas.

Whether they are pre-selected at programme level or local area level or determined at individual level (or a combination of the three), it is also important to bear in mind that priority individual-level outcomes included in the theory of change for the programme may not be achieved or evidenced within the timescales for the initial pilot and evaluation. We have therefore tried to indicate those outcomes in which progress appears more feasible in a two-year period. Even for these outcomes, however, individuals' progress within the evaluation timescales may be limited and non-linear. It is also important to recognise that for many people experiencing multiple disadvantage, initial stabilisation and no further deterioration in key areas could represent significant progress.

Focusing on individual-level outcomes within the Changing Futures programme evaluation likely means discussing the impact of direct interventions delivered with people experiencing multiple disadvantage. This is because in a two-year period it is more feasible to link progress against individual-level outcomes to specific services or interventions funded under the Changing Futures programme than to draw explicit connections between any early systems change activities and these outcomes.

2.3 Category 1: Changes in service use

Outcomes

During the feasibility study consultation, engaging and staying engaged with services was highlighted as an important early outcome of improved support by both people with lived experience of multiple disadvantage and those with insight into programmes to support them. Previous evaluations of programmes relating to multiple disadvantage have also found that engagement improves as a result of offering higher-quality and more intensive support to individuals (Cordis Bright, 2019; Broadbridge, 2018; Isaac et al., 2017).

Changes in wider service use have been priority outcomes in previous evaluations of programmes for people experiencing multiple disadvantage (Cream et al., 2020; Cordis Bright, 2020; Lamb et al., 2019a; Battrick et al., 2014). This is partly because they are also the primary area in which economic and fiscal benefits of the programme may be seen. This is therefore an important outcome area of focus for the Changing Futures evaluation, although it was not highlighted as a priority by people with lived experience during the feasibility study consultation and both literature and some stakeholders with insight into multiple disadvantage programmes reported that this is not a meaningful outcome measure for individuals (Field et al., 2019).

The specific types of service which people might use more/less will vary in each Changing Futures site, depending on the specific target cohort(s) for programme interventions, the nature of the interventions and the focus of local systems change work. Broadly, the most relevant service use outcomes which might be achieved within the programme and evaluation timescales are likely to be:

- **Increased engagement with specialist support services, such as those specifically funded via Changing Futures.** Achieving this outcome relies predominantly on individual clients and specialist services and is thus relatively independent of the wider system.
- **Increased engagement with planned support,** such as primary or secondary healthcare, substance use treatment and mental health services. Improvements in these areas can result from better support by a specialist service, but are likely to be enabled or inhibited by the availability of these wider services and the extent to which they are currently using approaches that work for people experiencing multiple disadvantage. This might affect the extent to which they are seen within a two-year period.
- **Changes in the use of emergency or crisis support and unplanned service use,** such as A&E attendance, unplanned hospital admissions, police call-outs/arrest and nights in prison¹⁰. In the longer term, the use of these types of service might be expected to decrease if clients are accessing planned support that enables them to improve their health, wellbeing, circumstances and behaviours that might otherwise result in the involvement of emergency and unplanned services. However, it is possible that some aspects of emergency service use might rise in the shorter term if clients'

¹⁰ Both the Fulfilling Lives and MEAM Approach evaluations found evidence of reduced interactions with the criminal justice system (Cordis Bright, 2020; Lamb et al., 2019b).

confidence and motivation to access them increases and if input from specialist services enables them to engage with required emergency support. This might represent a positive outcome for these clients¹¹.

- A further potential longer-term change in service use is a reduction in the intensity of specialist support required over time, but this is unlikely to take place in the timescales for the programme or evaluation.

Measurement approaches

The gold standard approach to gathering data on service use is a linked administrative dataset on service use by people receiving local interventions or those in the target cohort for local work under Changing Futures (with or without a control group). This could potentially be collated by local areas or – for some service use indicators – centrally by MHCLG and partners. There is also an option to include worker-reported or self-reported data within the dataset as well as or instead of administrative data. Methodologies and analytical approaches using this dataset are discussed in section 5.4.

Other potential measurement approaches include:

- The use of area-level indicators of service use which might be impacted by the programme¹². Approaches to analysis are discussed in section 5.5 but some potential sources of data which might be of relevance (if the target cohort for local programmes is large enough relative to the local population who might be included in these indicators) are:
 - [Alcohol and drug use and treatment statistics](#).
 - [PHE local authority dashboard](#): indicators such as % not in treatment, waiting times, successful completion, mortality - for alcohol dependency and substance use.
 - [Public health outcomes framework](#).
 - [MHCLG Rough Sleeping Statistics](#).
 - [Live tables on homelessness](#).
- The use of aggregated data on engagement with services, such as attendance, drop-out, completion of programmes. We understand that a similar approach is being used in the evaluation of project ADDER. In the Changing Futures programme evaluation, this approach might only be feasible for specialist services for the target cohort, because any changes in aggregated engagement data for more mainstream services will not be specific to the cohort supported by the local programme, and any specific changes for this cohort would not be visible in the data for these services.

¹¹ For example, if their trust in services has increased then people experiencing domestic abuse may report incidents to the police when they would not have done so in the past. Equally, people whose mental or physical health requires emergency treatment may be more likely to be admitted to hospital and remain there if they have the support of a trusted worker in a specialist service.

¹² For example, Centre for Social Justice (2017) suggests comparison of changes in homelessness data by region as an indicator of the impact of Housing First e.g. in Denmark, overall homelessness increased between 2009-2012 (43% nationally) but in areas that implemented a homeless strategy, the rise was 'much less steep' (4%).

Challenges in collating administrative data

Accessing routinely collected monitoring or administrative data can be beneficial to an evaluation, reducing the need to collect additional data from people using services and even offering the potential of creating a statistical control. There is a range of administrative datasets that cover aspects of multiple disadvantage, such as the National Drug Treatment Monitoring System, NHS Digital and Police National Computer. However, gaining access to administrative data can be challenging. The feasibility study consultation and evidence review and our previous experience of working with local areas to collate administrative data highlights a number of reasons for this:

Data about individuals is held by a range of services and is not linked. It can therefore be hard to piece together, particularly when organisations' data recording systems do not always flag other relevant needs and circumstances. For example, Lankelly Chase commissioned research that made use of administrative data to estimate the population of people experiencing multiple disadvantage but the exercise was challenging, with extensive gaps and overlaps in the data from different sources (Bramley and Fitzpatrick, 2015).

Information sharing agreements and processes are complicated. Accessing individual-level administrative data on sensitive topics like health and offending requires robust data security infrastructure and the ability to evidence this. Information Sharing Agreements and processes need to be bespoke to different organisations and/or local areas. If data collation and sharing is required at a local level, it can be particularly time-consuming to set up information sharing agreements, both between local agencies between the lead agency/data controller and evaluators.

Consent from individuals is required to collate and use the data and it can be challenging to obtain this from people experiencing multiple disadvantage (see section 7.4).

Dedicated resource is required at local level to set up and coordinate consent processes and data collation. It is therefore important that this is built into local areas' resource planning for the programme.

Local areas' success in collating data tends to be dependent on individual relationships within organisations in the partnership. This can result in patchy data if the data collator does not have relationships with key organisations or if relationships are lost when staff leave organisations.

It may be more efficient for MHCLG and partners to coordinate data collation and sharing, and this may result in higher-quality data. However, this would require significant resource for MHCLG. The evaluation of the Troubled Families programme (MHCLG, 2019) provides a unique example of administrative data being linked and used to conduct a robust impact evaluation. Data linkage and analysis was conducted by government departments. However, Troubled Families was a much larger and longer-term programme than Changing Futures.

2.4 Category 2: Changes in experiences of support

Outcomes

More positive experiences of support were also recognised as a likely early outcome of programmes such as Changing Futures by people with lived experience of multiple disadvantage and those with insight into programmes to support them, with previous evaluations identifying improvements in this area (Crisp et al., 2020; Cordis Bright, 2019; Isaac et al., 2017)¹³. These changes are closely linked to service-level outcomes detailed in Chapter 3 but have been included here because they are experienced by individual clients. They include experiences like:

- Feeling better informed about support and the options available to them.
- Having increased freedom to make choices about support and goals.
- Greater trust in services, or individual people working within them.
- Building and maintaining more open and honest relationships with a consistent worker.

Measurement approaches

The primary mechanism for measuring these outcomes is consultation with clients receiving support in local areas participating in the Changing Futures programme, which might be supplemented by consultation with staff members or volunteers delivering support (Denzin and Lincoln, 2017). Recommended approaches to consultation are discussed further in section 7.4.

2.5 Category 3: Changes in wellbeing, health, behaviours and circumstances

Outcomes

The Changing Futures programme has the potential to contribute to a wide range of outcomes related to changes in individuals' wellbeing, health, behaviours and circumstances. These outcomes have been the focus of previous evaluation and research relating to multiple disadvantage and its constituent parts. The specific nature of desired outcomes will vary extensively for individual clients, depending on what is of most importance to them. It will also differ substantially in different local areas, depending on the target cohort and nature of interventions by the local programme.

¹³ These could also be framed as reduced negative experiences of support.

Recommendation

In order to make best use of the evaluation resource, we recommend prioritising category 3 individual-level outcomes which:

1. Cut across the specific nature of individual needs/preferences and intervention types.
2. Are likely to precede or catalyse progress in other areas.
3. Are more likely to be achievable and feasible to demonstrate within the programme and evaluation timescales (although this may vary in individual local areas, depending on the system context and baseline).

These outcomes predominantly fall into four themes associated with the concept of recovery within mental health and substance use. That said, the deeply personal nature of recovery is a key tenet of this concept so it is important not to assume that all outcomes will apply to everyone (Shepherd et al, 2008; Anthony, 1993). In addition, many individuals who begin a recovery journey during the programme will still be in the early stages of this journey at the end of the programme and evaluation timescales (Onken et al., 2007)¹⁴.

Recovery-associated outcomes were also identified as a high priority by people with lived experience of multiple disadvantage who were consulted as part of the feasibility study, and were amongst the most commonly-identified priority outcomes by stakeholders with insight into programmes to support them. Figure 5 outlines how possible priority outcomes for the Changing Futures programme relate to the four key elements of recovery. Outcomes highlighted in bold are those in which the evidence review and consultation indicated progress would be seen first.

Figure 5: Key elements of recovery and relevant outcomes for the Changing Futures programme

| Recovery theme | Relevant outcomes: Category 3a: Shorter-term | Relevant outcomes: Category 3b: Longer-term |
|--|--|---|
| Hope for the future and positive aspirations (Korcha et al., 2011) | <ul style="list-style-type: none"> • Feeling more hopeful about the future. | |
| Connection to others, including supportive and positive networks (social | <ul style="list-style-type: none"> • Supportive relationships with trusted workers in services. | <ul style="list-style-type: none"> • Improved or maintained relationships with supportive friends. • Improved or maintained relationships with children |

¹⁴ Definitions of recovery vary between the fields of mental health and substance use, but all place emphasis on wellbeing and quality of life (see, for example, Best and Laudet (2010)). The evidence review for this feasibility study explored the concept of recovery and different ways in which it is conceptualised. A summary of the findings can be found in Appendix B.

| | | |
|--|--|---|
| capital) (Ruiu, 2016; Dingle et al., 2015; Longabauch et al., 2010) | | and/or other family members ¹⁵ . <ul style="list-style-type: none"> • Feeling safer¹⁵. • Feeling less lonely and/or isolated. |
| A sense of purpose and engagement in meaningful activities (community capital) (Best et al., 2015; Del Vecchio, 2012) | <ul style="list-style-type: none"> • Feeling a greater sense of purpose. • Having more opportunities to take part in meaningful activities. | <ul style="list-style-type: none"> • Improved education and employment outcomes or employability¹⁶. |
| Personal capital, including material resources, skills, physical health and self-esteem (Hennessey, 2017; Best et al, 2010; White and Cloud, 2008) | <ul style="list-style-type: none"> • Increased sense of agency and control. • More stable accommodation¹⁷. • Improved financial situation (including access to appropriate benefits). • Access to appropriate medication (for physical health, mental health and/or substance use treatment). | <ul style="list-style-type: none"> • Improved self-esteem and awareness of strengths. • Improved life skills. • Improved wellbeing¹⁸. |

Achieving key outcomes of importance to them from amongst those outlined in Figure 5 might ultimately contribute to one or more of several key overarching outcomes for individuals. They are therefore useful to measure both in their own right and as possible early indicators of longer-term change.

These overarching outcomes include:

- Sustained improvements in mental health.
- Sustained improvements in physical health.

¹⁵ Stakeholders with specialism in domestic abuse noted that this outcome is often of particular importance to victims/survivors of domestic abuse.

¹⁶ The evidence review and stakeholder consultation suggests that this outcome in particular is likely to be more longer-term than many of the others (Welford et al., 2021; Cattell and Mackie, 2011).

¹⁷ Both the Fulfilling Lives and MEAM Approach evaluations found substantial improvements in accommodation for individuals who were supported (Cordis Bright, 2020; Lamb et al., 2019b).

¹⁸ Social Impacts Task Force (no date) recommend adding a measure of subjective wellbeing to evaluations because it is useful in its own right but also associated with other outcomes, such as health and educational attainment.

- Sustained stability of accommodation.
- Sustained reduction or stabilisation in drug/alcohol use and/or reduced harm associated with this.
- Reduced reoffending.

For outcomes such as these, sustainment and stability are key (see, for example, Sandu et al, 2021; Friel et al., 2020; Centre for Social Justice, 2017; UK Drugs Policy Commission, 2012). Equally, previous evaluations have found that the pace of progress towards some of these outcomes might be slower, sometimes linked to challenges in accessing appropriate mental health and/or substance use support (CFE Research, 2020; Cordis Bright, 2020). Partly for these reasons, consultees suggested that they are less likely to be achieved by a significant proportion of clients and/or demonstrable within the timescales for the programme and evaluation. They should therefore be lower-priority outcomes of focus for the evaluation of the Changing Future programmes¹⁹.

Measurement approaches

Validated scales and other self-assessment or worker-assessment tools

There is a vast array of tools that could potentially be used to measure changes in wellbeing, health, behaviour and circumstances, including a range of Patient Reported Outcome Measures (PROMs) and Clinician Reported Outcome Measures (CROMs). Many of these are standardised or validated, and a number are used in a wide range of clinical settings and have been used in previous research and evaluation, which might provide opportunities for comparison and benchmarking if it does not prove feasible to identify a comparator group for the evaluation (Social Impacts Task Force, no date).

Given that the specific priority outcomes for the evaluation are not yet determined, it was not possible to review and summarise the full range of possible tools within the resource available for the feasibility study. A sample of tools are outlined in Appendix A, alongside benefits, challenges and examples of evaluations in which they have been implemented. Four of the tools that have been used in similar evaluations or that we have experience of using elsewhere are discussed below. However, there was no clear consensus amongst consulted stakeholders or within the reviewed literature as to which tools might be most appropriate for the Changing Futures evaluation.

During the set-up stage for the programme and evaluation, evaluators will need to work with MHCLG and local areas to select the most relevant and useful tools for measuring the selected outcomes. Key considerations will include whether relevant tools are currently in use in local areas, whether training and/or licensing is required prior to use²⁰, whether a single tool can capture data on key priority outcomes or whether it is more appropriate to build a bespoke tool (ideally using stand-alone questions from other validated or commonly-used tools which permit question selection in this way). Using tools which are

¹⁹ This judgement is based on the feasibility of reporting robust findings in relation to these outcomes in the timescales available. However, we are aware that some of these outcomes may be of priority importance to MHCLG and partners and a decision may thus be made to include them within the evaluation in any case.

²⁰ For example, services are required to train their staff in the use of Homelessness Outcomes Star as part of the licensing process.

applicable to the wider population may also be useful in enabling comparisons which shed light on the likely lower baselines for people experiencing multiple disadvantage.

As with administrative data, responses to tools will need to form part of a linked dataset.

Example outcomes measurement tools

Four example tools are described and discussed below. These are tools that are commonly used in evaluations of multiple disadvantage interventions or that we have experience of using in other evaluations which may be relevant to the Changing Futures programme. Outcomes measurement tools should be selected based on the key outcomes the evaluation seeks to measure. See Appendix A for summary of a larger range of potentially relevant tools.

Homelessness Outcomes Star

The [Homelessness Outcome Star](#) is designed to support and measure change within care settings against ten outcome areas including motivation and taking responsibility, managing tenancy and accommodation, emotional and mental health, and meaningful use of time. The tool is designed to be completed by support workers and clients together. The Homelessness Outcome Star is widely used by homelessness agencies and is internationally recognised (Mackeith, 2014). As a result, it has been used in the MEAM Approach and Fulfilling Lives evaluations (CFE Research and University of Sheffield, 2020; Cordis Bright, 2019 and 2020).

However, there are some key concerns about using this tool for the Changing Futures evaluation:

- The tool tends towards low inter-rater reliability, particularly if workers do not have sufficient guidance on using the tool (Johnson and Pleace, 2016; Mackeith, 2014). As a result, the Homelessness Outcome Star cannot be assumed to provide robust statistical evidence of effectiveness and is not recommended as an evaluation tool.
- The tool does not facilitate measurement or prioritisation of outcomes that are important to individual clients. The tool explores 10 outcome areas but these may not be outcome areas of relevance to those individuals. In addition, the Homelessness Outcome Star focusses on outcomes deemed important to the homeless population, but not all beneficiaries of interventions under the Changing Futures programme will be experiencing homelessness. Alternative Outcomes Stars are available, such as the Recovery Star, but we do not recommend these for similar reasons.

There are also costs involved with using the tool: services must purchase a license from Triangle Consulting Social Enterprise and practitioners should attend training delivered by Triangle.

The New Directions Team Assessment

Formerly known as the “Chaos Index”, the [New Directions Team Assessment](#) (NDTA) was designed as an assessment criteria/client identification tool for referrals to one of the Adults facing Chronic Exclusion programme pilot areas (Merton). The NDTA covers 10 areas deemed relevant to people experiencing multiple disadvantage, including self-harm, risk to and from others, stress and anxiety and housing. The NDTA is widely used for assessing eligibility for Housing First services (Homeless Link, 2020) as well as eligibility

for support from interventions under the MEAM Approach network and Fulfilling Lives programme.

There are also several concerns with regards to using this tool for the Changing Futures evaluation. First, as with the HOS, the tool provides no scope for adaptation to outcomes of importance to the individual. Second, the tool is deficit-focussed, which is against recovery-oriented principles of being strengths-based. Third, the tool is designed for completion by workers as a referral tool, and therefore gives no agency or control to the client. Fourth, it is designed as an assessment support tool rather than for measuring change over time.

Warwick Edinburgh Mental Wellbeing Scale²¹

The [Warwick Edinburgh Mental Wellbeing Scale](#) (WEMWBS) is a 14-item scale for measuring mental wellbeing in the general population and the impact of projects, programmes and policies which aim to improve mental wellbeing. A shorter 7-item version is also available (SWEMWBS). Both are psychometrically validated and are sensitive to change, however it is recommended that studies concerned with gender differences use the longer version (Fat et al., 2017). Both versions are widely used, meaning results can be benchmarked against national survey data and population norms. This also means (S)WEMWBS provides a “common currency” enabling comparison across interventions (Social Impacts Task Force, no date). Monetised social value can be assigned to scores using the [HACT wellbeing valuation approach](#).

Goal attainment scaling

People have different goals in life. Personalised goal setting is regularly used by mental health and rehabilitation professionals as a way of focussing support towards outcomes that are important to the client (Cairns et al., 2019). Assessment of progress towards these goals is a well-established person-centred approach (see, for example, Kiresuk and Sherman, 1968) which, unlike the tools set out above, focuses on the outcome areas (goals) that are important to individual clients.

Goal attainment scaling is in itself a recovery focussed way of working that helps people to be more hopeful about their life and helps their recovery. As described above, recovery outcomes can in turn be early indicators of longer-term change against some of the broader outcomes measured by tools like HOS. Research has also found that goal setting may help improve engagement and retention in mental health services (Cairns et al., 2019).

Considerations for tool completion²²

²¹ There are a range of other tools which measure wellbeing, recovery and/or quality of life. For example, [ReQoL](#) is designed for use with populations experiencing mental health needs but this may not be applicable to all clients supported under Changing Futures. It is also longer than SWEMWBS and the areas it asks individuals to comment on may be less applicable. Equally, Mansa (see Priebe et al, 1999) is a possible quality of life measurement tool but is longer than SWEMWBS. [DIALOG](#) measures subjective quality of life and treatment satisfaction but it is designed for mental health settings and its treatment focus feels inappropriate in this context.

²² Specific considerations for engaging people using services in tool completion are discussed in section Engaging people using services in the evaluation

Timing of baseline data collection and client-led outcomes identification (if self-reported data needed): The evaluation needs to allow for baseline data collection with individuals to take longer than it might for some interventions or with some client groups (Lamb et al., 2019c). This is because important building of relationships and trust takes place early on in engagement and asking individuals to complete multiple assessments and tools can undermine this (see Friel et al, 2020). In addition, there is the potential that individuals will report more accurately once they have built trust with staff members.

Timescales for review: If set timepoints for review of tools are specified this might be incompatible with the support being delivered (e.g. if an individual is experiencing particular challenges at the time of review and support needs to be prioritised over tool completion) (see for example, Cattell and Mackie, 2011). However, not specifying timepoints poses challenges for later analytical approaches and risks the possibility that key data is not collected if clients move on or disengage from services. In addition, it introduces the possibility that staff members complete tools at the most positive point in an individual's engagement/progress, thereby positively skewing assessments of change over time (Murphy et al., 2018).

Consistency of use by workers: It is important to reliability and validity that tools are completed consistently (see for example, Mackeith, 2014). Therefore resource may be required for staff training and quality assurance.

Administrative data

A linked individual-level administrative dataset could also be used to record changes in accommodation situation, or area-level/local population-level variables such as rough sleeping counts could be used as a measure for this outcome.

Qualitative consultation

As with changes in experience of support, another important approach to measuring these outcomes is consultation with people receiving support in local areas participating in the Changing Futures programme, which might be supplemented by consultation with staff members or volunteers delivering support. This can complement quantitative data collection approaches by enabling the evaluation to understand change in more detail and to unpick the factors that made change possible (Denzin and Lincoln, 2017) Recommended approaches to consultation with people using services are discussed further in section 7.4.

3 Service-level outcomes

3.1 Summary of suggested priority outcomes

Figure 6 outlines suggested priority service-level outcomes for the programme and the evaluation, along with possible measurement approaches. If qualitative methodologies and bespoke surveys are used, it may be possible to include measures relating to all or most of these outcomes. However, if there is a desire to use standardised/validated tools for some outcomes, it will probably be necessary to narrow the focus to a smaller number of outcomes to make it feasible for staff to complete tools for the highest-priority outcomes.

Figure 6: Summary of suggested priority service-level outcomes

| Priority outcomes | Measurement approaches |
|---|---|
| <p>Category 1: Changes in how support is delivered</p> <ul style="list-style-type: none"> a. Support is more trauma-informed and/or psychologically-informed. b. Support is more flexible, person-centred and strengths-based. c. Support is more holistic. | <ul style="list-style-type: none"> • Fidelity tools designed to support service improvement. • Standardised and validated tools on worker-client relationships. • Bespoke surveys of staff. • Qualitative consultation with staff members/ volunteers delivering support, their managers and individuals receiving support. |
| <p>Category 2: Changes in access and engagement</p> <ul style="list-style-type: none"> d. Relevant support is offered at the time and place where people need it. e. Services persevere and are creative in engaging people. f. Fewer people are declined support without an alternative support mechanism in place. g. Fewer people drop out of support, including maintaining support during key transitions. | <ul style="list-style-type: none"> • Change over time in output measures. • Qualitative consultation with staff members/ volunteers delivering support, their managers and individuals receiving support. |
| <p>Category 3: Changes within the workforce delivering support</p> <ul style="list-style-type: none"> h. People delivering support have a greater and shared understanding of multiple disadvantage and its constituent parts. i. People delivering support have a better awareness of the range of services and support available and how to enable people to access them. j. There are more support work roles for people with lived experience of multiple disadvantage. k. People delivering support have a greater sense of efficacy and autonomy. l. People delivering support receive appropriate support themselves, resulting in reduced staff burn-out. | <ul style="list-style-type: none"> • Bespoke surveys of staff. • Standardised and validated tools on staff burn-out. • Qualitative consultation with staff members/ volunteers delivering support, their managers and individuals receiving support. |

3.2 Introduction

During the feasibility study consultation, there was a relatively strong consensus amongst people with lived experience of multiple disadvantage and those with insight into programmes to support them about the highest-priority service-level outcomes for the Changing Futures programme. In fact, people with lived experience generally gravitated towards describing service-level outcomes before they discussed individual-level or system-level outcomes, suggesting that these might be of particular importance to individuals. The outcomes resonated with the draft theory of change for the programme and divide into three main areas:

- Changes in how support is delivered (Category 1)
- Changes in access and engagement (Category 2)
- Changes within the workforce delivering support (Category 3).

Previous literature reviews suggest that these features of services are important to effective support for people experiencing multiple disadvantage (McCarthy et al., 2020; Diamond et al., 2013). A number of them have also been the focus of previous evaluations of similar programmes (Friel et al., 2020; Crisp et al., 2020; Ipsos Mori, 2019).

The navigator role/individual staff members in specialist services are key to many of these service-level outcomes, and previous evaluations have found that individuals often associate improved support with their worker rather than services as a whole (McCarthy et al., 2020; Cordis Bright, 2019; MHCLG, 2019; Moreton et al., 2018).

Implications for the evaluation

A strong process evaluation is important in order to be able to describe the work of specialist roles and services in detail, and to understand how and in what circumstances they impact on outcomes for individuals.

It will be important for the evaluation to distinguish between outcomes which can be achieved and maintained by specialist workers and those that need to be achieved on a wider scale across the system if any improvements are to be embedded and sustained.

3.3 Category 1: Changes in how support is delivered

Priority outcomes in this area are²³:

- **Support is more trauma-informed and/or psychologically-informed.** Those experiencing multiple disadvantage often report experiences of trauma and/or adverse childhood experiences and difficulties managing their emotions, impulsivity and behaviours. Trauma and psychologically-informed support is therefore key (Richard and

²³ As with the outcomes in section 2.4, these could also be framed as reduced provision of support which individuals experience negatively or which has a negative impact for them.

Paquette, 2019; Ritchie, 2015; Hayes et al., 2013; Felitti and Anda, 2010; Hopper et al., 2010). In addition to a safe environment and working cultures, an important feature is trusting relationships with staff in services, which is recognised separately in the literature and is also captured as an individual-level outcome in section 2.4.1 (Sandu et al., 2021; Johnsen, 2013; Cattell and Mackie, 2011).

- **Support is more flexible, person-centred and strengths-based.** This is linked to the individual-level outcome of increased agency and control discussed in section 2.5.1, and here people with lived experience included making choices about your support, such as picking the gender of your support worker. Some stakeholders noted that this also includes offering support at varying levels of intensity to respond to changes in individuals' circumstances and need for support.
- **Support is more holistic.** Support takes account of the inter-related nature of individual's needs and circumstances, rather than being structured around a single issue.

3.4 Category 2: Changes in access and engagement

Priority outcomes were predominantly identified during the feasibility study consultation, by both people with lived experience and a wide range of other stakeholders with insight into multiple disadvantage and related programmes. The outcomes are:

- **Relevant support is offered at the time and place where people need it.** This encompasses the provision of outreach support and also the concept of "no wrong door" (see for example Revolving Doors Agency, 2019).
- **Services persevere and are creative in engaging people.**
- **Fewer people are declined support without an alternative support mechanism in place.**
- **Fewer people drop out of support, including maintaining support during key transitions.**

3.5 Category 3: Changes within the workforce delivering support

Given the importance of well-trained and effective staff in developing relationships and what is known about the stresses of work in this field, it is important for the Changing Futures evaluation to track staff outcomes (Moreton et al., 2021; Moreton et al., 2018; Maguire et al., 2017). Priority outcomes here are:

- People delivering support have a greater and shared understanding of multiple disadvantage and its constituent parts.
- People delivering support have a better awareness of the range of services and support available and how to enable people to access them.
- There are more support work roles for people with lived experience of multiple disadvantage.

- People delivering support have a greater sense of efficacy and autonomy.
- People delivering support receive appropriate support themselves, resulting in reduced staff burn-out.

3.6 Measurement approaches

There appear to be fewer standardised measures of service-level outcomes. There are, however, tools that can be used to assess:

- The degree of fidelity to trauma-informed approaches (e.g. [the TICOMETER](#), Bassuk et al., 2017) and psychologically-informed environments (e.g. [the Pizazz](#), PIE link NET, 2021), although these are generally designed to provide a resource to support service improvement rather than programme evaluation.
- The client-worker relationship (e.g. the [Working Alliance Inventory](#), Horvath, 1984, 1981).
- Staff burn-out (the [Maslach Burnout Inventory](#), Maslach et al., 1986)

A sample of tools are outlined in more detail in Appendix A, alongside benefits, challenges and examples of evaluations in which they have been implemented.

For outcomes relating to changes in access and engagement, change over time in output measures also be appropriate, for example numbers of people seen by services, numbers sustaining engagement, numbers declined a service, numbers requiring repeat support etc. Ideally, output monitoring data would also enable the evaluation to understand individuals' progression within relevant services.

In the main, however, evaluations focusing on service-level outcomes tend to use bespoke surveys and qualitative methods (Cordis Bright, 2020; Cordis Bright, 2019; Cream et al., 2020; Maguire et al., 2017; Bretherton and Pleace, 2015; Bell, 2014). This might include survey or qualitative consultation with staff/volunteers working in services, their managers and individuals receiving support. Qualitative elements are important to move beyond describing outputs but can themselves sometimes result in relatively descriptive accounts of people's perspectives on service-level changes. If possible, it would be useful to explore whether the evaluation can benchmark good practice (e.g. in terms of the quality of the relationship between workers and clients or trauma-informed practice) and use mixed methods understand the extent to which areas are delivering this or getting closer to it.

4 System-level outcomes

4.1 Summary of suggested priority outcomes

Figure 7 outlines suggested priority system-level outcomes for the programme and evaluation. Systems thinking suggests that these may be more usefully conceptualised as changes than as outcomes (Egan et al. 2019a. See section 4.2.1 for more detail). We suggest prioritising changes in categories 1 and 2 as they are more likely to take place within the timescales for the programme and evaluation and therefore progress is more likely to be demonstrable. However, depending on the system starting point and intended system activities in different local areas, some changes in categories 3 and 4 may also be feasible.

Figure 7: Summary of suggested priority system-level outcomes

| Priority outcomes | Measurement approaches |
|---|--|
| <p>Category 1: Changes to multi-agency partnerships and working</p> <ul style="list-style-type: none"> a. Increased collaboration and investment in the agenda by systems actors, including people commissioning, designing, delivering and using services. b. Greater diversity of voices in partnerships, including more effective involvement of people with lived experience of multiple disadvantage. c. More effective information sharing, at both operational and strategic levels. d. Better coordinated support with clear pathways into services. <p>Category 2: Changes to system learning and development processes</p> <ul style="list-style-type: none"> a. Partnerships and system actors are open to learning and improvement and have effective mechanisms in place to do so. b. System barriers are addressed strategically and removed. <p>Category 3: Changes to commissioning structures and practices</p> <ul style="list-style-type: none"> a. Commissioning is more integrated and less siloed. b. Commissioning and monitoring of services focuses more on shared outcomes and on quality. c. Commissioning structures and processes incentivise service activities and attitudes that work for people experiencing multiple disadvantage. <p>Category 4: Improved systems of support</p> <ul style="list-style-type: none"> a. Better identification of people experiencing multiple disadvantage. Stakeholders highlighted in particular that this includes identifying people from groups who might currently be less well-identified, such as women and people with uncertain immigration status. | <ul style="list-style-type: none"> • Qualitative and participatory approaches. • Use of existing tools to audit partnerships and system maturity. • Use of outputs as a proxy for outcomes. |

- | | |
|--|--|
| <ul style="list-style-type: none">b. Earlier intervention and more preventative work.c. More support is available that meets the needs of people experiencing multiple disadvantage, including different support and approaches to suit different people.d. Reduced stigma and judgement within services of people experiencing multiple disadvantage. | |
|--|--|

4.2 Introduction

Systems thinking and evaluation

A systems perspective is useful for the evaluation of the Changing Futures programme for three key reasons (Bicket et al 2020; Egan et al., 2019b):

1. The programme is aiming to catalyse changes to wider structures or systems.
2. MHCLG and partners and local areas are interested in the processes and mechanisms which lead to changes.
3. Traditional evaluation methods may not be easily applicable for a range of reasons, such as system-level activities having no obvious start date and being difficult to disentangle from other activities and not being easily replicable.

Systems thinking can be a useful way to consider the bigger picture and think about how different people, services and organisations interconnect and influence each other (Egan et al., 2019a). There is no agreed definition of a system or a complex system (Bicket et al, 2020; Abercrombie et al., 2015). However, public health systems are recognised as complex and adaptive (Egan et al, 2019a). This means that they are characterised by (Bicket et al, 2020; Egan, 2019a):

- Many diverse, interacting components.
- Non-linear and non-proportional interactions between these components.
- Adaptation or learning by the components in response to change.
- Unpredictability.
- Emergence and change over time.

As systems are always in a process of changing, it may be more helpful to focus on evaluating system changes rather than system-level outcomes; the latter implies that outcomes will be fixed at the end of the programme or evaluation, whereas in fact change will be ongoing, with or without additional specific interventions at system-level (Egan et al. 2019a).

Although the systems which the Changing Futures programme seeks to influence are complex, evaluating its impact on these systems doesn't have to be. Evaluating the whole system is infeasible and it is likely to be more effective to focus available resource on examining specific aspects of the system and its complexity (HM Treasury, 2020a; Bicket et al 2020; Egan et al., 2019b; Moore et al., 2018).

Implications for the evaluation

1. Understanding the programme interventions and context in detail at the outset is even more crucial than in some other types of evaluation, so a comprehensive theory of change or logic map is an important first step and participatory approaches are valuable (Bicket et al, 2020; Moore et al., 2018).

2. Although the evaluation might seek to explore particular areas of systems change in more detail, it should remain open to identifying other changes which might emerge as important, and particularly to unplanned or unintended changes (Bicket et al, 2020; Egan et al., 2019a).
3. Due to the unpredictability of the system and how it may change, the evaluation needs to take a flexible approach, allowing opportunities to review and refine the approach in stages (Bicket et al, 2020; Egan et al., 2019a). This may necessitate a closer-than-usual relationship between the local areas and the evaluators (HM Treasury, 2020a).
4. The evaluation should be formative and form part of a continuous learning process for the programme and systems so needs to build in approaches to share early learning (HM Treasury, 2020a; Davidson Knight et al., 2017).

Challenges in defining system-level outcomes/changes

In addition to the challenges of evaluating in a complex system, the feasibility study consultation identified several challenges in defining specific priority system-level outcomes/changes. These relate to difficulties in agreeing what a “good” system looks like and how the system compares to this, which helps to focus in on what might need to change. In addition, it will be important to ensure that local areas share MHCLG and partners’ understanding of any systems change outcomes. There may also be a need to home in from overarching outcomes to more specific changes at a local level, depending on the nature of local systems change activity and context.

Overview of priority outcome categories

As with service-level outcomes, there was general agreement from the feasibility study consultation on the primary system-level outcomes that the Changing Futures programme should aim to catalyse, several of which have been a focal point in previous evaluations of similar programmes or of research into effective support for people experiencing multiple disadvantage (Cordis Bright, 2020; MHCLG, 2019; Moreton et al., 2018; Bell, 2014; Diamond et al., 2013). Again, these were broadly aligned with the draft theory of change and they fall into four overarching categories:

- **Changes to multi-agency partnerships and working.** In this area, more appears to be known about the features of effective partnership working and less on how different models of collaboration affect service user outcomes (Atkinson et al., 2007).
- **Changes to system learning and development processes.**
- **Changes to commissioning structures and practices.**
- **Changes to support enabled by improved systems.**

The latter two categories predominantly describe longer-term system changes, which are less likely to be seen in the timescales for the programme and evaluation. We therefore

suggest prioritising the first two categories of outcome (highlighted in bold above)²⁴. This is also in keeping with advice on systems evaluation discussed in section 4.2.1 in that it might enable the evaluation to focus on elements of the system rather than all aspects of its complexity, in order to make the evaluation more manageable within the resource available.

4.3 Changes to be prioritised

Category 1: Changes to multi-agency partnerships and working

Priority outcomes in this area are:

- Increased collaboration and investment in the agenda by systems actors, including people commissioning, designing, delivering and using services.
- Greater diversity of voices in partnerships, including more effective involvement of people with lived experience of multiple disadvantage.
- **More effective information sharing, at both operational and strategic levels.** This includes collating and sharing data to facilitate support for individual clients, but also to enable learning for service and system improvement.
- **Better coordinated support with clear pathways into services.** A particular feature highlighted by people with lived experience in relation to this outcome was not having to tell your story repeatedly to different people.

Category 2: Changes to system learning and development processes

Priority outcomes in this area are:

- **Partnerships and system actors are open to learning and improvement and have effective mechanisms in place to do so.** Literature on whole systems working and systems change emphasises the importance of ongoing learning and improvement (Davidson Knight et al., 2017; Abercrombie et al, 2015; Hough, 2014)
- **System barriers are addressed strategically and removed.** At present, there appears to be a gap in the evidence relating to the nature of systemic barriers to access and how they can be addressed, particularly in relation to minority groups experiencing multiple disadvantage (McCarthy et al., 2020). This may therefore also be a useful area of focus within any process evaluation accompanying the impact evaluation.

²⁴ Depending on the starting point and specific nature of programme delivery in local areas, some of these outcomes may be more relevant and achievable within the programme timescales.

4.4 Longer-term changes of lower priority for the evaluation

Category 3: Changes to commissioning structures and practices

Outcomes in this area include:

- Commissioning is more integrated and less siloed.
- Commissioning and monitoring of services focuses more on shared outcomes and on quality.
- Commissioning structures and processes incentivise service activities and attitudes that work for people experiencing multiple disadvantage.

Category 4: Improved systems of support

Outcomes in this area include:

- **Better identification of people experiencing multiple disadvantage.** Stakeholders highlighted in particular that this includes identifying people from groups who might currently be less well-identified, such as women and people with uncertain immigration status.
- Earlier intervention and more preventative work.
- More support is available that meets the needs of people experiencing multiple disadvantage, including different support and approaches to suit different people.
- Reduced stigma and judgement within services of people experiencing multiple disadvantage.

4.5 Measurement approaches

The primary methods for evaluating systems change are qualitative and participatory, though there are some possible quantitative approaches that could be applied too. The range of approaches is discussed in detail in section 5.6.

In terms of specific tools, there are tools that can be used to audit partnership working, although these are generally designed to provide a resource to support service improvement rather than for outcome evaluation. Examples include: the [EIF maturity matrix](#), [Lankelly Chase system behaviours framework](#) and the [MEAM Approach principles](#) and related tools²⁵.

²⁵ [Collaborate CIC](#) also offers a range of diagnostic tools and frameworks but these are not publicly available. The [SHARE framework](#) from the Centre for Homelessness Impact also provides an overview of elements and principles required for a system to be able to support reduced homelessness.

There is also the possibility of using outputs as a proxy for some outcomes relating to multi-agency working, such as:

- The development of formal structures for partnership working, including whether Memoranda of Understanding or similar protocols are in place and the number of partnership meetings that take place.
- The development of information sharing agreements or other protocols to support sharing of information.
- The development of tools to support joint working, such as common needs assessments.

5 Measuring impact

5.1 Summary of potential approaches to impact evaluation

Figure 8 summarises the three key levels at which the impact of the Changing Futures programme might usefully be explored by the evaluation, along with our suggestions for the most robust of the feasible approaches to conducting impact evaluation at this level, and the key impact evaluation questions at each level. Exploration of service-level outcomes could be incorporated at any of the three levels.

It would be possible to conduct impact evaluation at one or more of these levels, depending on the resource available for the evaluation. The decision on which level(s) to include depends on whether MHCLG and partners wish to place greater emphasis on evaluating outcomes at individual-level (attributable directly to interventions or otherwise) or at system-level (contributed to by the programme funding and model). There were divergent views amongst consulted stakeholders about which of the levels was most important and useful, depending on stakeholders' own methodological expertise and on their views on the priority focus for the evaluation.

It is our recommendation that the evaluation focuses as a minimum on level C, the impact at local system-level. This is because one of the key distinguishing characteristics of Changing Futures is the central focus on enabling systems change via an injection of funding and support. Focusing on an impact evaluation at this level may be the most effective way to produce a useful evaluation at national level, given the bespoke nature of local programmes and implementation.

If resource allows, we then recommend pursuing an evaluation of level A, intervention-level impact, as a second priority. If well designed, this could add value in improving the evidence base for the impact of specific interventions in different systems and contexts. However, many of the challenges outlined in the previous section are particularly applicable to developing and delivering an experimental or quasi-experimental evaluation design to understand individual-level impact, which also builds in approaches to unpick the nature of and context for interventions in more detail in order to say more about which elements of an intervention are important in determining impact.

The decision on which level to focus on may also be determined by the extent to which economic evaluation is a priority, because this is more likely to be feasible in approaches centred on impact for individuals (levels A and B) than those centred on impact for the system (level C). At this level, unit costs can be attached to outputs and outcomes to gain an understanding of economic impact.

If impact evaluation is required at multiple levels, there is an argument for commissioning these aspects separately as they will require evaluators with different specialist skillsets. In this case, careful thought will be required as to how to integrate the findings from both evaluations to avoid findings which fail to explore the important connections between the impact of funded interventions and the maturity and features of systems in which they are delivered.

Within any of these levels of impact evaluation and irrespective of the specific approaches taken within them, we recommend a theory-based, mixed methods overall approach to

understanding impact. This should be accompanied by a strong and detailed process evaluation which provides supporting information to explore the mechanisms and context for change at individual, service and system level. Alongside focusing on one or more of these levels, a key purpose of the programme is to understand whether and how central funding and resource can be used to catalyse changes at a local and national level, and to generate learning about how to do so. Therefore the evaluation should focus on generating and capturing learning about these aspects.

Neither the consultation nor the evidence review provided definitive findings on the impact evaluation methods which would be most suitable at any of the three levels. Final approaches and methods would need to be determined with the selected evaluators, in consultation with local areas.

Exploring specific methods in detail

At the time of the feasibility study, the preferred balance was yet to be determined between evaluating individual-level outcomes, which might primarily result from local interventions funded by Changing Futures, and examining system-level outcomes, which might result from the programme's focus on funding systems change activity. This precluded the option of focussing in greater detail on the benefits and challenges of a smaller number of approaches potentially suited to a specific evaluation focus. Instead, it was necessary to examine a wide range of evaluation approaches and methods in as much depth as possible within the resource and timescales available.

Figure 8: Summary of different levels and approaches to impact evaluation

| | Level A: Intervention-level impact | Level B: Local population-level impact | Level C: Local system-level impact |
|--|--|--|---|
| Main focus of impact evaluation | The impact for individuals receiving support directly from interventions funded or catalysed by the Changing Futures programme | The impact of the whole local programme of work relating to the Changing Futures programme (not just specific interventions) for all individuals experiencing multiple disadvantage in the local area (not just those receiving support from interventions) | The impact of the Changing Futures programme model on systems change activity and systems in the local areas included in the programme |
| Most robust of the feasible approach(es) | <p>Interrupted time series / change over time data study with no comparison group.</p> <p>Data for beneficiaries at two or more time points is compared in order to understand change in key indicators of outcomes.</p> <p>For change over time studies, data for a baseline time point is compared to data at one or more later time point (typically before/at the start of an intervention and after an intervention).</p> | <p>The feasibility of these approaches would need to be established by further exploration of the likelihood of any changes for the local multiply disadvantaged population being visible in the selected lead variables.</p> <ul style="list-style-type: none"> • Difference in difference with either one or two lead variables or synthetic controls. • Qualitative comparative analysis. <p>This could be delivered as part of a mixed methods approach, which could include qualitative approaches and additional quantitative methods.</p> | <p>Qualitative and participatory approaches, to be determined by selected evaluators. These could possibly be combined with quantitative approaches using frameworks for assessing system maturity or specific changes and outputs as proxies for some outcomes.</p> |

| | | | |
|--------------------|---|--|---|
| | <p>For interrupted time series, data at multiple time points before an intervention and multiple points after an intervention is required²⁶.</p> <p>The feasibility of these approaches should be confirmed once more when the specific outcomes and indicators of interest to the evaluation have been finalised.</p> <p>This could be delivered as part of a mixed methods approach, which could include qualitative approaches and additional quantitative methods.</p> | | |
| Our recommendation | We recommend pursuing an evaluation at this level as the second priority. | We do not recommend pursuing an evaluation at this level, unless further programme and evaluation planning establishes a clear lead variable at population level, in which change is likely to be visible given the scale of the programme activity and the evaluation timescales. | We recommend pursuing an evaluation at this level as the first priority |

²⁶ As a result, an interrupted time series design would likely require the use of administrative data as this is the only type of data likely to be available for multiple time points prior to an intervention.

| | | | |
|---------------------------------|---|--|--|
| <p>Key evaluation questions</p> | <ol style="list-style-type: none"> 1. Are the direct interventions with individuals that are funded via the Changing Futures programme delivering their intended outcomes for and with individuals who receive them? 2. Are any changes in outcomes experienced equally by different groups of individuals (e.g. based on presenting needs, demographic characteristics or any other differentiating factors)? If not, what might explain these differential outcomes? 3. Which aspects of the intervention(s) enable individuals to achieve positive outcomes? 4. Which aspects of the intervention(s) inhibit or limit positive outcomes? | <ol style="list-style-type: none"> 1. Is the whole programme of work (i.e. funded interventions and systems change activity) funded via the Changing Futures programme delivering positive outcomes for and with individuals experiencing multiple disadvantage in the funded areas? 2. Are any changes in outcomes experienced equally by different groups of individuals (e.g. based on presenting needs, demographic characteristics or any other differentiating factors)? If not, what might explain these differential outcomes? 3. Which aspects of the programme of work are contributing to these outcomes? 4. Which aspects of the programme of work inhibit or limit positive outcomes? | <ol style="list-style-type: none"> 1. Does the Changing Futures model catalyse systems change activity in local areas included in the programme? 2. If so, do these changes contribute to the intended system-level outcomes of the programme? 3. Are there other unanticipated changes to the system to which the Changing Futures programme contributes? 4. Which aspects of the programme of work contribute to systems change or activity to promote it? 5. Which aspects of the programme of work inhibit or limit systems change or activity to promote it? |
|---------------------------------|---|--|--|

5.2 Introduction

Impact evaluation at different levels

In this section we set out the key approaches for measuring impact – including economic impact – and understanding causality/attribution/contribution. There are three key levels at which the impact of the Changing Futures programme might usefully be explored by the evaluation. These are:

- **Level A: Intervention-level impact.** The impact for individuals receiving support directly from interventions funded or catalysed by the Changing Futures programme.
- **Level B: Local population-level impact.** The impact of the whole local programme of work relating to the Changing Futures programme (not just specific interventions) for all individuals experiencing multiple disadvantage in the local area (not just those receiving support from interventions).
- **Level C: Local system-level impact:** The impact of the Changing Futures programme model on systems change activity and systems in the local areas included in the programme.

Impact evaluation at levels A and B would involve a greater focus on quantitative measurement of harder outcomes experienced by individuals and a greater emphasis on the direct attribution of these outcomes to aspects of the Changing Futures programme. Impact evaluation at level C would draw more heavily on complexity theory and on qualitative and participatory approaches to understanding systems change and its relationship to the Changing Futures programme.

It would be possible to conduct impact evaluation at one or more of these levels, depending on the resource available for the evaluation. The decision on which level(s) to include primarily comes down to the question of whether MHCLG and partners would most like to place greater emphasis on evaluating outcomes at individual-level (attributable directly to interventions or otherwise) or at system-level (contributed to by the programme funding and model). It may also be determined by the extent to which economic evaluation is a priority, because this is more likely to be feasible in approaches centred on impact for individuals (levels A and B) than those centred on impact for the system (level C).

Impact evaluation at any of these levels could be delivered across all areas involved in the Changing Futures programme, but could also focus on a smaller number of areas of interventions of particularly high interest. This might enable more resource-intensive approaches to be delivered in selected areas.

Challenges in attribution

The supplementary Magenta Book guidance (Campbell and Harper, 2012) states that good quantitative impact evaluation should provide evidence that change can be attributed to an intervention. The Changing Futures programme is seeking to effect change through varied and complex interventions that involve multiple organisations and that vary between programme sites and between individual clients. Differences in interventions may, however, be less of a concern if the evaluation is primarily aiming to establish the impact of funding and delivering additional interventions as opposed to business as usual without

the Changing Futures programme. There may also be less of a concern if the interventions are based on common underlying principles or characteristics, and the evaluation seeks to determine the impact of delivering interventions which include these.

In addition, activities and interventions are being introduced into a complex system, which poses significant challenges for attribution (Bicket et al., 2020). Specific challenges which were recognised by stakeholders during the feasibility study consultation include:

- **Capturing additionality of the programme model.** This is challenging in relation to individual-level impact, where it is difficult to unpick any additional value of the programme's wider systems change work from the value of simply funding specific intervention/service in the local area (such as navigator roles). It is also challenging in relation to system-level impact, where it is difficult to determine the extent to which the Changing Futures funding has caused changes which might otherwise not have occurred (particularly in areas where systems change work relating to multiple disadvantage has already been taking place prior to Changing Futures).
- **Change overload.** There is so much change underway throughout the system that it will be difficult to disentangle any impact of Changing Futures from the impact of other changes. For example, there will be concurrent changes to the probation system, and there is considerable crossover with other programmes such as Project ADDER, the Rough Sleepers Initiative, and the Troubled Families programme and associated funding streams.
- **Difficulties for stakeholders to comment on attribution.** Even if change is attributable to the programme, system actors might not view it in this way or may not even recognise the programme.

These challenges add to the difficulties in applying counterfactual approaches to measuring impact, which we set out in section 5.4.2.

Generalisability

It will be difficult to say whether any impact identified for the Changing Futures programme would be achieved at a similar scale in other local areas or in similar programmes in the future. This is because of the level of variation across the programme, e.g. differences in selection process for clients included in the intervention/dataset, and because local contexts and systems will be varied and will not be precisely replicable elsewhere (Burchett et al., 2020; Egan et al., 2019a).

Accounting for timescales

Most stakeholders emphasised the importance of being realistic about the type of impact evaluation which can be delivered in the timescales for the programme and evaluation. In part, this relates to realism about the types of outcomes which might be achieved and demonstrable within the timescales (as discussed in Chapters 2-4). It also means realism about the type of evaluation which can be set-up and delivered in a two-year period when

evaluating new interventions and activities which have themselves yet to be set up²⁷. This may preclude the use of some approaches; where this is a concern it is indicated in the discussion throughout this chapter.

Considering differential impact

Reflecting on the diversity of people experiencing multiple disadvantage, both the evidence review and consulted stakeholders highlighted the importance of investigating whether and how services and the system are meeting the needs of different groups of individuals (e.g. based on presenting needs, demographic characteristics or any other differentiating factors), and whether these groups experience differential outcomes as a result of interventions/the programme (McCarthy et al., 2020; McNeish et al., 2016)²⁸. The inclusion of domestic abuse in the definition of multiple disadvantage potentially changes the make-up of target cohorts in local areas and finding ways to unpick this further will be useful (Sosenko et al., 2020).

Exploring any differential impact using quantitative approaches will be more feasible in the intervention-level impact approach (level A) but it could still be explored qualitatively in the local population-level approach (level B). The feasibility of quantitative approaches to this will also depend on sample size and group sizes within the sample.

5.3 Role for theory-based, mixed methods and qualitative approaches

Recommendation

Irrespective of whether the evaluation deploys a counterfactual approach to measuring impact or one of the alternative approaches suggested, we recommend a theory-based, mixed methods overall approach to understanding impact. Again, this makes it crucial to develop and maintain a strong theory of change for the programme. This can be used as a starting point for theory-based methods, which must also remain open to identifying other changes which might emerge as important, and particularly to unplanned or unintended changes (Bicket et al, 2020; Egan et al., 2019a).

For all approaches, a strong process evaluation will also be crucial to understanding how different aspects of the programme have been delivered and why different approaches and interventions have been successful or less successful. This is discussed further in Chapter 6.

Many stakeholders highlighted how it may be more appropriate to use theory-based approaches to understand impact and causality. This is supported by the Magenta Book, which sets out how theory-based approaches may be more suitable when trying to evaluate a combination of interventions in a complex landscape or interventions designed to make a change in a complex system, and when there is a need to understand the role

²⁷ For example, for interventions being introduced as part of Changing Futures, the delivery period coinciding with the evaluation timescales may in fact be 18 months or less because an initial lead-in period will be required to commission and/or set up services and interventions.

²⁸ For example, Fulfilling Lives found some evidence of less good outcomes for women compared to men (Lamb et al., 2019d)

of place and context in achieving outcomes (HM Treasury, 2020a). All of these are features of the Changing Futures programme. The Magenta Book also describes how theory-based approaches may be appropriate when it is not possible to develop a counterfactual. In fact, the Changing Futures evaluation might provide an important opportunity to build the evidence base and acceptance of theory-based, mixed methods systems evaluation approaches as an alternative to approaches drawing primarily on experimental or quasi-experimental designs.

In a complex system the “cause” will rarely be the intervention alone. What is important is how the intervention works in relation to other parts of the system (Byrne, 2013). Theory-based approaches can therefore be used to test causal chains to help understand the extent of change and why it has occurred. These approaches can be applied rigorously when supported by the following: a well-defined and coherent Theory of Change, specific evidence to test theory, triangulation of multiple sources, approaches to ruling out alternative causes, critical reflection, and peer and external scrutiny. Theory-based methods allow for an understanding of attribution or contribution but do not give precise effect sizes (Bicket et al, 2020). In particular, two main generative causation designs could be used to complement or deliver impact evaluation at any of levels A-C outlined in sections 5.4 to 5.6. These are:

- **Realist approaches, based on a theory of change and mixed methods.** Data collected through both quantitative and qualitative methods are analysed and triangulated to assess change and understand causes for change, grounded in a rigorous theory of change (Bicket et al., 2020; Salter and Kothari, 2014; Pawson and Tilley, 1997; Better Evaluation, no date)²⁹.
- **Contribution Analysis.** This is a step-by-step approach for exploring if an intervention has contributed to observed outcome, which also helps to explain the process of change (Bicket et al., 2020). It is an evidence-based line of reasoning rather than concrete proof and is appropriate if there is large variability in implementation or outcomes (HM Treasury, 2020a; Bicket et al, 2020; Befani and Mayne, 2014; Mayne, 2008)³⁰. It may be useful in understanding how to improve implementation but can be harder to undertake if there is a complex or distant relationship between process and outcomes (Bicket et al., 2020; Biggs et al, 2014).

In addition to or as part of theory-based approaches, mixed methods are particularly important when evaluating complex interventions in a complex system and can help to understand attribution/contribution (Bicket et al., 2020; Moore et al., 2018). Using a range of methods helps to overcome the limitations of any one approach; it enables the triangulation of data to build confidence in conclusions, and it provides vital insight into the context, and the how and why of impact. The Centre for the Evaluation of Complexity Across the Nexus (CECAN) provides an [interactive tool](#) for choosing appropriate evaluation methods. This is particularly applicable for impact approaches seeking to understand the impact of the programme (rather than at the level of specific funded interventions) – in this case levels B and C.

²⁹ This approach was used in the evaluation of the Government’s Drug Strategy 2010 (HM Government, 2017) and was suggested in the feasibility study for the evaluation of Violence Reduction Units (MacLeod et al., 2020).

³⁰ We understand from consultation that this approach may be used in the evaluation of Project ADDER.

5.4 Level A: Intervention-level impact for individuals receiving support directly from interventions

Key evaluation questions for level A impact evaluation

1. Are the direct interventions with individuals that are funded via the Changing Futures programme delivering their intended outcomes for and with individuals who receive them?
2. Are any changes in outcomes experienced equally by different groups of individuals (e.g. based on presenting needs, demographic characteristics or any other differentiating factors)? If not, what might explain these differential outcomes?
3. Which aspects of the intervention(s) enable individuals to achieve positive outcomes?
4. Which aspects of the intervention(s) inhibit or limit positive outcomes?

Overview

In this section we explore approaches for measuring the impact for individuals receiving support directly from interventions funded or catalysed by the Changing Futures programme, with a particular emphasis on exploring possible counterfactual approaches.

Challenges for counterfactual approaches

A good counterfactual is often viewed as essential in attributing impact to specific interventions, as causality can rarely be confidently attributed without this (Campbell and Harper, 2012). However, both the evidence review and consultation exercise show that it will be very challenging to robustly apply counterfactual approaches to evaluating the impacts of the interventions funded or catalysed by the Changing Futures programme. Counterfactual approaches are likely to be very resource-intensive and, in isolation, will not be able to reliably attribute impact to the interventions. There are several reasons for this:

- **Complexity of the system, interventions and the cohort.** The experts we consulted with highlighted the complexity of the system and the cohort, and the possibility of multiple interventions with one individual, rather than a single, identifiable 'treatment'. Byrne (2013) argues that the complexity of the social world is such that the value of randomised controlled trials (RCTs) in establishing causality of complex social interventions is limited. In a complex system the cause will rarely be the intervention alone. What is important is how the intervention works in relation to other parts of the system. The complexity therefore reduces the extent to which a counterfactual approach is feasible and meaningful. This is supported by the Magenta Book (HM Treasury, 2020a, p.35 to 36), which sets out how experimental and quasi-experimental designs are most appropriate when the intervention does not involve several activities or a varied implementation. They are less feasible if the intervention is difficult to define or disentangle from other interventions, adapts over time, is applied to a complex system, builds up gradually over extended time, involves consolidation of existing good practice, and there is subjective allocation (comparison and target group are different from the outset), and if the effect size is small.

- **Differences between local areas.** Stakeholders also highlighted how there are likely to be considerable differences between local programmes and interventions. This is again against the criteria set out in the Magenta Book which make experimental and quasi-experimental designs more feasible (see above).
- **Focus on innovation and development.** Stakeholders also highlighted how the Changing Futures programme places emphasis on innovation, exploration and learning through delivery. This also makes RCTs and quasi-experimental designs unsuitable (see above).
- **Absence of “control” conditions in potential comparator areas.** Stakeholders highlighted that some areas are likely to already be doing some of the work included in the bid, and unsuccessful areas (i.e. possible comparator areas) will continue to do some of the work anyway. This reduces the ability to identify a “control” group in another area. Moreover, areas outside of the Changing Futures programme with significant populations of people experiencing multiple disadvantage are likely to be Fulfilling Lives or MEAM Approach areas in any case, leading to risk of cross-contamination. Areas that are not already participating in these programmes are likely to be those with smaller populations of people experiencing multiple disadvantage, which would make the unsuitable comparators.
- **Difficulty of identifying a well-matched control group.** This is particularly challenging given the target beneficiary group, who tend to not be otherwise connected to support services. Consulted stakeholders highlighted that this reduces the feasibility of identifying a control group who are not connected to at least some form of service, thereby meaning that one needs to compare outcomes for recipients of the treatment service with outcomes for recipients of another model of service (rather than no-treatment). For example, the Fulfilling Lives evaluation sought to establish a matched comparison group from people accessing support in non-Fulfilling Lives areas. However, while the comparison group members all met the basic criteria of experiencing need across at least two of the multiple disadvantage areas, overall, they had lower levels of need and risk, were less likely to have needs across three or four needs, were less likely to be disabled or have a long-term health condition, were more likely to be in employment have some form of qualification and were less likely to have a problem with literacy (Lamb et al, 2019a).
- **Use of resources.** In a number of cases, stakeholders reported that it would be a poor use of resource to try to pursue counterfactual options and that there are better ways to make the best use of evidence likely to be available to the evaluation.

Counterfactual approaches involving an external comparator group

We understand that MHCLG view a counterfactual approach involving an external comparator group as important and we have therefore made every effort to explore a range of options.

A pure experimental method (RCTs) is not appropriate or desirable for the Changing Futures evaluation. Stakeholders highlighted the obvious ethical issues with withholding service from people randomly allocated into a control group. Using waiting list control groups also poses ethical issues because clients will need a service as soon as they are identified and because there is not necessarily a guarantee of later service given the funding timescales. Waiting list control groups also pose methodological challenges

because the control group sample size shrinks as people move from the waiting list into the service.

However, there is a range of other quasi-experimental evaluation designs and methods which attempt to mimic randomisation through identifying or creating a well-matched distinct counterfactual comparison group. These methods could potentially be applied to the Changing Futures programme if the key challenges set out above and in Figure 9 in section 5.5.5 are not considered sufficient barriers to pursuing a counterfactual approach. Non-beneficiaries in the comparison group would likely need to be from areas not participating in the Changing Futures programme because non-beneficiaries within the area may still be impacted by the systems change work.

The evaluation could then apply a difference in difference analysis to compare the mean change over time in the lead outcome variable for the intervention group with the mean change over the same period for the comparison group, or regression discontinuity analysis.

We set out the principal counterfactual approaches involving a comparator group identified through stakeholder consultation and the evidence review below. Figure 9 in section 5.5.5 provides more detail.

- **Propensity score matching.** The propensity score is the probability for a subject to receive treatment conditional on a set of characteristics. It is commonly estimated using logistic regression. Once participants and non-participants are matched using propensity scores (assuming all influencing factors can be controlled for), any difference in outcomes can be attributed to the intervention.
- **Stepped wedge design.** If the intervention can be rolled out in stages, potential beneficiaries in the later joining areas can act as the control group in the early stages. In larger local areas involved in the Changing Futures programme, this could involve phased roll-out in different localities to provide within-area comparator groups. However, phased roll-out might not be feasible within the timescales for the programme.
- **Un-matched comparator.** In this approach a group of people identified as experiencing multiple disadvantage, most likely from a non-participating area, would form the comparator group.

Most robust and feasible approaches in absence of external comparator group

All consulted experts in quantitative and economic analysis recognised the challenges in identifying an external comparator group for use in the evaluation and none were able to identify a feasible solution to this. As a result of this, several recommended that the evaluation should not pursue quasi-experimental evaluation designs with a distinct counterfactual group of people who are not programme beneficiaries. Instead, they recommended alternative quantitative approaches for measuring impact at level A. These approaches are likely to be more feasible than the approaches above and/or will be more effective at developing an understanding of what changes are happening for whom, and how and why these changes are occurring.

We set out the principal quantitative approaches without a distinct comparator below. Figure 9 in section 5.5.5 provides more detail.

- **Interrupted time series / change over time data study.** Data for beneficiaries at a two or more time points is compared in order to understand change in key indicators of outcomes.
- **Simulation modelling.** The modelled continuation of the beneficiaries' pre-support trend is the counterfactual. Markov and Monte Carlo modelling are commonly used in health economics.
- **Comparisons between cohorts across local areas in the programme.** Comparing outcomes between different types of areas can help understand impact, as well as potential causal factors. For example, comparisons could be made between areas taking different approaches to implementing support, or areas with different levels of funding.
- The most promising alternative quantitative method is interrupted time series or change over time analysis, with no distinct counterfactual comparison group. This may typically be perceived as a less robust evaluation design than designs with distinct counterfactual comparison groups. However, given the challenges with counterfactual approaches set out in section 5.4.3 many stakeholders recommended an interrupted time series or change over time data study as the most feasible and proportionate quantitative approach to measuring impact for complex interventions with this client group. In addition, stakeholders with insight into alternative approaches without a distinct comparator group (e.g. simulation modelling and comparison between cohorts across local areas) ultimately concluded that these were likely to be infeasible with this cohort and likely outcomes of focus and over this relatively short programme and evaluation timescale.

Considerations for interrupted time series or change over time data studies

Please note: We have not made proposals for the indicators and data sources that would be included in this type of data study. This is because we do not yet know the chosen priority outcomes for the evaluation. The specifics of any data study would need to be determined by the selected evaluators, MHCLG and partners and the local areas involved in Changing Futures. However, an interrupted time series design would likely require the use of administrative data as this is the only type of data likely to be available for multiple time points prior to an intervention.

Consulted stakeholders highlighted the following key considerations for an effective design:

Baseline/pre-intervention timings: Baseline data from the start of support is easier to obtain than pre-service use data. The latter is generally only possible with administrative data but it is usually labour-intensive for local areas to collate (though it may be more feasible if MHCLG and partners have capacity to collate it centrally). If using self-reported measures, the baseline period will need to allow time for accurate self-reporting by clients once trust is built.

Temporal dimension: The length of time a person experiences multiple disadvantage is significant – longer experience often equates to greater intensity of problems.

Time series: We know that the path of progress and required length of support varies quite substantially between clients. Given this, some stakeholders question whether it is meaningful to examine progress for all clients after a fixed time period.

Non-linear nature of progress for clients experiencing multiple disadvantage: This means fluctuations in outcomes data are likely and a longer time period is required to be confident about trajectory.

Time required for initial progress, especially in harder outcome measures: Clients need time to socialise into the intervention, engage and build trust before they are likely to make progress.

Trading off sample size and data quality: Coordinating data collection represents a reasonable amount of work and often the data ends up being too patchy to be robust and useful. Asking people to collect data for a random selection of clients in the sample might improve data quality.

5.5 Level B: Local population-level impact of the whole local programme of work (not just specific interventions)

Overview

In this section we explore approaches for measuring the impact for all individuals experiencing multiple disadvantage in the local area of the local programme of work relating to the Changing Futures programme. These approaches would use area-level/local population-level variables rather than individual-level variables.

Key evaluation questions for level B impact evaluation

1. Is the whole programme of work (i.e. funded interventions and systems change activity) funded via the Changing Futures programme delivering positive outcomes for and with individuals experiencing multiple disadvantage in the funded areas?
2. Are any changes in outcomes experienced equally by different groups of individuals (e.g. based on presenting needs, demographic characteristics or any other differentiating factors)? If not, what might explain these differential outcomes?
3. Which aspects of the programme of work are contributing to these outcomes?
4. Which aspects of the programme of work inhibit or limit positive outcomes?

Benefits of local population-level impact evaluation

Some stakeholders recommended taking a local population-level approach to understanding impact, either in addition to or as an alternative to the intervention-level approach set out in section 5.4. This is also identified as a potentially valid approach by the Violence Reduction Unit impact evaluation feasibility study (Macleod et al., 2020). A focus on local population-level impact has several potential advantages over the evaluation of intervention-level impact:

- It enables **exploration of programme-level impact**, which we understand from consultation to be important to MHCLG. This enables testing of whether the whole Changing Futures model is effective, i.e. whether centrally funding sites to develop and implement a programme locally and including an emphasis on systems change brings about system improvements and better services and outcomes. In contrast, exploration of intervention-level impact only enables testing of whether the interventions funded or catalysed by the programme bring about better outcomes for individuals who access that support.
- It would be based on **publicly available aggregate data**, which is comparable between areas and can easily be accessed by the evaluation team.

Challenges of local population-level impact evaluation

- Many of the challenges outlined in section 5.4.2 also apply to local population-level impact evaluation. In addition, it faces some distinct challenges:
- **Risk that impact is not visible in local population-level variables.** A number of stakeholders asserted that the small cohort size and short timescales of the programme would mean that any impact on local population-level administrative data would be too small to see. Others reported that an evidence review is required in order to understand whether it would be feasible to see an effect on these variables, with particular consideration of effect size.
- **Inference of impact restricted to local programme level.** Stakeholders highlighted that it won't be possible to make inferences about individual projects within an area based on its aggregate outcomes.
- **Significant noise in system.** Areas may be systematically different to one another. Without very careful matching to comparators, there may be a lot of noise in the data, which can mask effects of the intervention in question (Campbell and Harper, 2012).

Additional approaches to measuring impact at level B

Many of the approaches outlined in relation to measuring impact at level A (see sections 5.4.3 to 5.4.4) might also apply at level B, but with a focus on area-level/local population-level comparators and variables rather than individual-level ones. Relevant approaches are indicated in Figure 9 in section 5.5.5. Three further approaches could also theoretically be used at level B, but only if a suitable lead variable can be identified and if it appears feasible that changes for people experiencing multiple disadvantage would be visible in this variable. The most likely variable is probably rough sleeping counts, but further investigation would be required to make a judgement on its suitability. These are:

- **Synthetic controls.** This methodology enables the construction of a counterfactual by selecting a weighted average of the outcome variable from a group of units similar to the treated unit (Bouttell et al, 2018).
- **Qualitative comparative analysis.** This is an approach to identifying the necessary conditions (e.g. aspects of an intervention and the wider context) to achieve an outcome of interest. The methodology works back from an outcome to identify how different

combinations of factors combine to affect the outcome (Bicket et al., 2020; Byrne, 2016)³¹.

- **Process tracing.** The main purpose of process tracing is to establish whether, and how, a potential cause or causes influenced a specified change or set of changes. This is done by working backwards from an outcome and identifying a chain of causal mechanisms and then applying formal tests of probability to assess the strength of evidence for each link in the chain. Process tracing also involves testing alternative ideas about how change might have come about (Bicket et al, 2020; Befani and Stedman-Bryce, 2017; Puntton and Welle, 2015).

Approaches to measuring impact at levels A and B

Figure 9 summarises tools, methods and approaches for measuring impact at levels A and B, any key benefits and challenges, and examples of evaluations where they have been used.

³¹ QCA could also potentially be used in conjunction with other methods applicable to impact level A, in order to explore any differential impact of interventions in different local areas.

Figure 9: Approaches for measuring the intervention-level and local population-level impact of the Changing Futures programme
Experimental design

| Approach | Level A | Level B | Benefits and challenges ³² | Example evaluations using this approach |
|--|---------|---------|---|---|
| Randomised controlled trials (RCTs) | ✓ | | <ul style="list-style-type: none"> • Seen as ‘gold standard’ of impact evaluation (Campbell and Harper, 2012). • Strong method for attributing causality (Campbell and Harper, 2012). • May take longer to set up and needs close management throughout (Campbell and Harper, 2012). • May be less feasible for evaluating complex interventions in a complex system (Byrne, 2013). • Little used in relation to multiple disadvantage in UK context. <p>Unlikely to be perceived as an ethical approach to take for interventions with this cohort.</p> | Housing First (Canada) IPS Mental Health and Employment Support Evaluation |

Quasi-experimental design – with distinct counterfactual comparison group

| Approach | Level A | Level B | Benefits and challenges | Example evaluations using this approach |
|----------------------------------|---------|---------|--|--|
| Propensity score matching | ✓ | | <ul style="list-style-type: none"> • Good method for attributing causality (Campbell and Harper, 2012). • Particularly effective with datasets with a large number of variables. | Justice Data Lab evaluations |

³² Where references are not provided, this information was drawn from the feasibility study consultation.

| Approach | Level A | Level B | Benefits and challenges | Example evaluations using this approach |
|-----------------------------|---------|---------|--|---|
| | | | <ul style="list-style-type: none"> • Requires good data on treatment and comparator groups (Campbell and Harper, 2012). • Only controls for factors for which data is available (Campbell and Harper, 2012). • Collecting/sourcing data on comparator groups may be challenging (Campbell and Harper, 2012). • Unlikely to be feasible in this evaluation due to the difficulties in establishing an external comparator group (see sections 5.4.2 and 5.4.3). | <p>Homelessness Social Impact Bond</p> <p>Troubled Families</p> |
| Waiting list designs | ✓ | | <ul style="list-style-type: none"> • Can be a solution in situations where RCTs are not appropriate as participants still receive the intervention. • For this client group it remains ethically challenging as those requiring support are likely to need to access it as soon as possible. For this reason it is unlikely to be feasible. • Still requires consent and/or data collection from comparator group – may be challenging with people experiencing multiple disadvantage. • Control group sample size shrinks as people drop out of the original service group and those on the waiting list access the intervention. | <p>Troubled Families</p> |
| Stepped wedge design | ✓ | ✓ | <ul style="list-style-type: none"> • Over time, all eligible individuals in all areas still get the intervention. • As with waiting list designs, over time the controls become part of the treatment group so it is not possible to understand the long-term implementations. • Loss of statistical power when focus on clusters (areas). | |

| Approach | Level A | Level B | Benefits and challenges | Example evaluations using this approach |
|---------------------------|---------|---------|---|--|
| | | | <ul style="list-style-type: none"> • Need to consider possibility of cross-area contamination if areas are close to each other. • Short timescales of Changing Futures programme means this design is unlikely to be feasible. | |
| Synthetic controls | | ✓ | <ul style="list-style-type: none"> • Offers relevant comparison when few or no other comparators exist (Bouttell et al., 2018; HM Treasury 2020a). • Most commonly applied to evaluating population-level interventions (Bouttell et al., 2018; HM Treasury 2020a). • Requires consistent data from treatment and control units prior to intervention, and for data to show relationship between two (Bouttell et al., 2018; HM Treasury 2020a). So, likely only to work if using routine data collected by statutory bodies. There is no clear lead variable collected by statutory bodies that would be most relevant to the Changing Futures evaluation, and for this reason the approach is unlikely to be feasible. • Not widely used in the UK and not in evaluations of sufficiently similar programmes. Also not yet widely understood. • Would be difficult to use if interventions and target cohorts in each local area in the programme are not standardised. For this reason, it is likely to be infeasible for this evaluation. • For area-level analysis, would produce a separate indication for each local area individually, rather than an average across all areas - might show where it is working and where it's not. | <p>California's Tobacco Control Program</p> <p>Considered for evaluation of Violence Reduction Units (Feasibility study)</p> <p>Consultation suggests this approach will be taken in the evaluation of Project ADDER</p> |

| Approach | Level A | Level B | Benefits and challenges | Example evaluations using this approach |
|------------------------------|---------|---------|---|--|
| Un-matched comparator | ✓ | ✓ | <ul style="list-style-type: none"> • Weaker approach to attributing causality (Campbell and Harper, 2012). • Comparator group may be very different from the treatment group (Campbell and Harper, 2012). • Collecting/sourcing data from comparator can still be resource intensive (Campbell and Harper, 2012). • In the Fulfilling Lives evaluation, the differences in the profile of the treatment and comparator cohort proved too substantial for meaningful comparisons to be made. This would likely be the case if the approach were taken for Changing Futures, and for this reason it is not a good use of evaluation resource. | Fulfilling Lives Rough Sleeping Initiative³³ |

Quasi-experimental design – with no distinct counterfactual comparison group

| Approach | Level A | Level B | Benefits and challenges | Example evaluations using this approach |
|--|---------|---------|---|---|
| Interrupted time series/change over time analysis | ✓ | ✓ | <ul style="list-style-type: none"> • Weaker approach to attributing causality (Campbell and Harper, 2012). • Less onerous data collection requirements (Campbell and Harper, 2012). • Relies on the assumption that the previous trend in the treated unit would have continued but for the intervention | Fulfilling Lives MEAM Approach |

³³ The Rough Sleeping Initiative impact analysis used a difference in differences approach, whereby mean change in rough sleeping in the 83 areas in receipt of RSI funding was compared to mean change in rough sleeping in the 83 areas with the next highest levels of rough sleeping. This is less feasible for Changing Futures because: (a) there's no obvious lead variable and (b) it would be difficult to find comparable areas that didn't have some work going on to address multiple disadvantage.

| Approach | Level A | Level B | Benefits and challenges | Example evaluations using this approach |
|--|---------|---------|---|---|
| | | | <p>only. (This assumption could be tested via theory-based methods).</p> <ul style="list-style-type: none"> Fulfilling Lives and MEAM Approach evaluations have shown that it is possible to obtain reliable pre- and post-participation data on a range of relevant variables, e.g. A&E visits, arrests, hospital admissions. This could be a combination of self-report and administrative data. However, identifying a counterfactual is difficult (Lamb et al, 2019a). Several stakeholders recommended this as the best option for measuring impact. | <p>Adults facing Chronic Exclusion pilots</p> <p>Tackling Multiple Disadvantage</p> <p>Impact of CBT training</p> |
| Simulation modelling | ✓ | ✓ | <ul style="list-style-type: none"> Addresses the difficulty in recruiting and obtaining data for individuals in a counterfactual group, but applies some robustness to the concept of an alternative trend. | |
| Comparisons between local areas in the programme | ✓ | ✓ | <ul style="list-style-type: none"> Stakeholders reported this approach to be useful in understanding what worked well and less well. Must be coupled with decent investment of resource in the process evaluation and in understanding local contexts and systems in which interventions are situated. | |
| Individual trials on selected local interventions | | | <ul style="list-style-type: none"> May be a useful way of delivering some impact evaluation at Level A if taking this approach in all areas is not feasible. Could focus on interventions thought to be most similar/comparable, promising, mature or potentially scalable. Might enable more resource-intensive work in this smaller number of areas | <p>National Collaborative Outreach Programme</p> |

| Approach | Level A | Level B | Benefits and challenges | Example evaluations using this approach |
|----------|---------|---------|---|---|
| | | | <ul style="list-style-type: none"> Requires strategic buy-in and careful management. Likely to be a resource-intensive and expensive exercise (Tazzyman and Bowes, 2019), and may not produce generalisable findings. However, if more experimental or quasi-experimental methods are considered important by key stakeholders this may still provide highly-valued evidence, albeit at a smaller scale. | |

Qualitative approaches

| Approach | Level A | Level B | Benefits and challenges | Example evaluations using this approach |
|---|---------|---------|---|---|
| Qualitative Comparative Analysis | | ✓ | <ul style="list-style-type: none"> Can be used to understand individual and service characteristics associated with outcomes (Bicket et al, 2020; Byrne, 2016). Can be used to identify how interventions can be implemented most effectively (Warren et al, 2014) Can account for complex causation in a systematic way (Byrne, 2016) and be used to evaluate systems change interventions (e.g. Matheson et al, 2019). Works best with 10-50 cases (HM Treasury, 2020a) Useful where interventions are similar but with different contexts (Byrne, 2016). Useful where stakeholders are disparate but is resource intensive and requires robust facilitation. | A 'health first' case management service Healthy Families NZ |
| Process tracing | | ✓ | <ul style="list-style-type: none"> Adds rigour and transparency to qualitative theory-based evaluation (Punton and Welle, 2015; Befani and Stedman-Bryce, 2016). | |

| | | | | |
|---------------------------------|---|---|---|---|
| | | | <ul style="list-style-type: none"> • Helpful in evaluating complex interventions, where pathways of change is unpredictable and dependent on changing circumstances (Kay and Baker, 2015; Punton and Welle, 2015). • Most appropriate for single-case studies, though Kay and Baker (2015) suggest it can also be used for between case analysis. • Can be time intensive and may need substantial volume of evidence to avoid inconclusive results. May be challenging where outcomes are not known until the end of the evaluation (Punton and Welle, 2015). • Approach can be strengthened by using in combination with contribution analysis (Befani and Mayne, 2014) | |
| Case studies | ✓ | ✓ | <ul style="list-style-type: none"> • Can provide a detailed understanding of how change mechanisms operate and how participants experience the intervention. • Can be produced at area level, intervention level or individual level. • Can take a thematic approach to delve into specific outcomes or processes of interest. • Are specific to the case in question so may not be representative or produce any generalisable findings. | Troubled Families MEAM Approach Homelessness Social Impact Bond |
| Participatory approaches | ✓ | ✓ | <ul style="list-style-type: none"> • Participatory approaches involve professional researchers and project participants / people with lived experience working together on the evaluation. • Involving a range of stakeholders helps ensure multiple perspectives are captured and conflicting views identified (Bicket et al., 2020). • Specific methods could include semi-structured interviews, focus groups, participant observation, diaries | |

| | | | | |
|--|--|--|---|--|
| | | | and more innovative methods such as mobile ethnography or Photovoice. <ul style="list-style-type: none">• Helps build a shared understanding (Bicket et al., 2020). | |
|--|--|--|---|--|

5.6 Level C: System-level impact of the funding/programme model

Key evaluation questions for level C impact evaluation

1. Does the Changing Futures model catalyse systems change activity in local areas included in the programme?
2. If so, do these changes contribute to the intended system-level outcomes of the programme?
3. Are there other unanticipated changes to the system to which the Changing Futures programme contributes?
4. Which aspects of the programme of work contribute to systems change or activity to promote it?
5. Which aspects of the programme of work inhibit or limit systems change or activity to promote it?

Overview

In this section we explore approaches for measuring the impact of the Changing Futures programme model on systems change activity and systems in the local areas included in the programme.

Considerations for evaluating complex systems

The Changing Futures programme has many features of a complex intervention in a complex system. The implications of this for evaluation are discussed in section 4.2. In addition to these considerations, complex systems evaluations often challenge traditional notions of evaluation and robustness, can be time-consuming, require specialist expertise and involve methods which have not yet been widely used (Bicket et al. 2020).

Approaches to measuring impact at level C

Counterfactual approaches are unlikely to be appropriate given the scope of systems change and the complexity of the system. Instead, a range of alternative approaches have been identified by the evidence review and consultation with stakeholders. With all approaches, understanding the intervention and context in detail, through a theory of change or logic map, is an important first step (Bicket et al. 2020; Moore et al., 2018).

Qualitative and participatory approaches

Most stakeholders and literature stress the important role of qualitative and participatory methods and approaches that involve engagement with a wide range of stakeholders and sources of evidence to identify both indicators and drivers of change (Bicket et al., 2020; Egan et al., 2019a and 2019b). These approaches tend to focus on understanding the system context and baseline, understand change as a process, and seek to capture the progress made and the programme's role in this. Stakeholders highlighted the importance

of building a sense of the system's trajectory and building a narrative around what is happening in/to the system.

Key methodologies for assessing systems change/ impacts are set out below. Hybrid designs which draw on more than one of these methodologies may also be appropriate (Bicket et al., 2020). Each approach will involve a range of research and analysis methods. Stakeholders highlighted qualitative interviews and focus groups with system actors, observations of partnerships in action, and content/discourse analysis as likely to be particularly useful. More information on key benefits and challenges to these approaches is provided in Figure 10.

In addition to engaging partners and individual staff members who are more closely involved in the local implementation of the programme, it will also be important to engage wider stakeholders who are more removed from the programme, in order to understand whether they are observing changes in systems or outcomes and whether they link these to the Changing Futures programme.

- **Qualitative research with a systems lens.** This would likely involve a range of qualitative methods. It can involve wide sampling of participants, a focus on relationships and change, analysing how different parts of the system affect one another (Egan et al 2019b).
- **System mapping.** This is a methodology for creating a visual map of the system. This is typically developed using participatory methods, bringing together a range of people representing different perspectives and organisations (Egan et al., 2019a; Hough, 2014). However, stakeholders highlighted how system mapping can also be built on document review, which is cheaper and easier than participatory workshops. System mapping can provide its own insight and build shared understanding across the system, as well as inform the design of further evaluation research. This can be understood as a key part of “doing” systems change, as well as evaluation (Abercrombie et al., 2015). Stakeholders highlighted how this can help identify other activity in the system that might complement or detract from the programme activity.
- **Network analysis.** This is a methodology for mapping how different people or organisations connect to one another. It can be repeated over time to show change in relationships (Egan et al 2019b).
- **Qualitative comparative analysis.** This is an approach to identifying the necessary conditions (e.g. aspects of an intervention and the wider context) for an outcome of interest. The methodology works back from an outcome to identify how different combinations of factors affect the outcome (Bicket et al, 2020; Byrne, 2016).
- **Process tracing.** The main purpose of process tracing is to establish whether, and how, a potential cause or causes influenced a specified change or set of changes. This is done by working backwards from an outcome and identifying a chain of causal mechanisms and then applying formal tests of probability to assess the strength of evidence for each link in the chain. Process tracing also involves testing alternative ideas about how change might have come about (Bicket et al, 2020; Befani and Stedman-Bryce, 2017; Punton and Welle, 2015).

Quantitative approaches

- Some stakeholders highlighted the need to move beyond qualitative conversations about systems. The following quantitative approaches can build on the qualitative and participatory methodologies above. However, given the challenges of assessing impact on system change in complex systems, the following methods are most appropriately applied as part of the broader, process-focussed approaches described above in order to understand the role of the Changing Futures programme in any changes identified.
- **Frameworks** for assessing system maturity or specific types of system change, such as collaboration. (See section 4.5 for details on specific tools). This can be used to monitor and assess scale and nature of change across the system.
- **Survey of actors across the system.** Questions focussing on behaviour outcomes such as partnership working at several points in time can also help monitor and assess scale and nature of behaviour-related change across the system.
- **Computational modelling** can be used to help understand complex systems and test the impact of different scenarios. These methods are highly skilled and require reliable data (Egan et al 2019b). Stakeholders recommended Systems Dynamics and Agent Based Modelling as two potentially relevant techniques.
- **Outputs as proxies for outcomes.** Stakeholders recommended a number of output metrics which could serve as proxies for specific system changes or outcomes. These include the number of ISAs or MOUs in place, and the number of partnership meetings and/or learning events. (See section 4.5 for more discussion).

Figure 10: Approaches for measuring the system-level impact of the Changing Futures programme

Qualitative and participatory approaches

| Approach | Benefits and challenges ³⁴ | Example evaluations using this approach |
|--|--|--|
| Qualitative research with a systems lens | <ul style="list-style-type: none"> • Appropriate for understanding complex interventions in a complex environment (Egan et al, 2019a). • Brings together evidence drawn from a range of sources. • Helps with understanding how change happens. | Troubled Families MEAM Approach Newcastle and Gateshead Fulfilling Lives |
| Systems mapping / concept mapping / network analysis | <ul style="list-style-type: none"> • Can help identify barriers and opportunities (Egan et al., 2019b; Mouser and Bowers, 2017). | Fulfilling Lives Lambeth, Southwark and Lewisham |

³⁴ Where references are not provided, this information was drawn from the feasibility study consultation.

| Approach | Benefits and challenges ³⁴ | Example evaluations using this approach |
|----------------------------------|--|--|
| | <ul style="list-style-type: none"> • Can inform qualitative or computational methods (Egan et al., 2019a). • Resulting map may be too simplistic (Moore et al., 2019). • Danger of getting ‘bogged down’ into trying to capture complexity before acting (Abercrombie et al., 2015). • Stakeholders highlighted how any mapping would need to focus on a smaller number of areas if to be used for Changing Futures programme. • | |
| Qualitative comparative analysis | <ul style="list-style-type: none"> • Can be used to understand individual and service characteristics associated with outcomes (Bicket et al, 2020; Byrne, 2016). Can be used to identify how interventions can be implemented most effectively (Warren et al, 2014). • Can account for complex causation in a systematic way (Byrne, 2016) and be used to evaluate systems change interventions (e.g. Matheson et al, 2019). • Works best with 10-50 cases – useful where interventions are similar but with different contexts (Byrne, 2016). • Useful where stakeholders are disparate but is resource intensive and requires robust facilitation. | A ‘health first’ case management service Healthy Families NZ |
| Process tracing | <ul style="list-style-type: none"> • Adds rigour and transparency to qualitative theory-based evaluation (Punton and Welle, 2015; Befani and Stedman-Bryce, 2016). • Helpful in evaluating complex interventions, where pathways of change is unpredictable and dependent on changing circumstances (Kay and Baker, 2015; Punton and Welle, 2015). • Most appropriate for single-case studies, though Kay and Baker (2015) suggest it can also be used for between case analysis. • Can be time intensive and may need substantial volume of evidence to avoid inconclusive results. May be challenging where outcomes are not known until the end of the evaluation (Punton and Welle, 2015). | Consultation suggests this approach will be an option in the evaluation of Project ADDER |

| Approach | Benefits and challenges ³⁴ | Example evaluations using this approach |
|----------|---|---|
| | <ul style="list-style-type: none"> • Approach can be strengthened by using in combination with contribution analysis (Befani and Mayne, 2014) . • | |

Quantitative approaches

| Approach | Benefits and challenges | Example evaluations using this approach |
|---------------------------------|---|--|
| System maturity frameworks | <ul style="list-style-type: none"> • Stakeholders were unaware of a directly relevant framework - existing frameworks would likely need to be adapted for Changing Futures. • Unable to account for attribution in isolation. | |
| Survey of system actors | <ul style="list-style-type: none"> • Stakeholders highlighted this as useful for understanding change in attitudes, practice and views. • Unable to account for attribution in isolation. • Potential source of evidence to support qualitative and theory-led approaches. | Troubled Families MEAM Approach |
| Computational modelling | <ul style="list-style-type: none"> • Can help understand complex systems and test the impact of different scenarios (Egan et al., 2019b; Bicket et al. 2020). • Highly skilled. • Require reliable data (Egan et al, 2019a. • Stakeholders recommend that this requires prolonged relationship with stakeholders and is a big, difficult and expensive project that is not feasible within the two year timescales of Changing Futures. | |
| Outputs as proxies for outcomes | <ul style="list-style-type: none"> • May be useful in indicating progress and key milestones in relation to areas such as multi-agency working but are not in themselves indicators of changes in ways of working so would need to be supplemented with qualitative methodologies. • Unable to account for attribution in isolation. | |

5.7 Economic impact

Methods for measuring economic and fiscal impact

Within the programme and evaluation timescales, it is most feasible to examine economic impact at the level of impact for individuals receiving support directly from interventions (see section 5.4). At a cohort level, longer-term fiscal impact may be seen over a longer period of time.

In all of the quantitative approaches to measuring intervention-level impact listed in section 5.4, a unit cost can be attached to inputs, outputs and outcomes to gain an understanding of economic impact. For example, if the lead variable is hospital admissions a tariff cost can be used. Social outcomes, such as change in wellbeing, can also be monetised, as can health outcomes (using QALYs). See the box on tariffs and tools for recommended sources for such unit cost tariffs.

Reporting of data from local areas on spend could be aligned to quarterly monitoring submissions to MHCLG. To collate this data in a consistent format, we would suggest developing a standardised template for submission by local programme leads. This would ideally include Changing Futures programme funding, matched funding from local areas or other sources, and in-kind contributions such as staff time in posts whose salaries are not funded as part of Changing Futures. As part of the set-up phase for the programme and evaluation, it will be important to involve local areas in discussions about the range of potential sources of funding and resource in their local area. The findings from these discussions should be factored into the development of the spreadsheet, which should include guidance for local areas on which costs should be included in which category and how to monetise in-kind resources.

Once unit costs are applied, economic impact can be assessed through a range of approaches. However, the same challenges to measuring impact as set out earlier in this chapter apply to the assessment of economic impact. For this reason, approaches to economic impact assessment that rely on a counterfactual were not widely recommended by stakeholders. A more proportionate and realistic approach may be to assess costs and benefits based on the beneficiaries' service use before and after service use, and for this assessment to focus on changes in a limited set of outcome areas. However, this approach is still likely to be relatively resource-intensive and challenges in generating and collating high-quality service use data might affect the robustness of the analysis and findings. Further investigation would be required by the appointed evaluators to determine whether likely levels of validity justify the use of evaluation resource.

We set out the principal approaches identified through stakeholder consultation and the evidence review below. See Figure 11 for more detail on the advantages and benefits of each method.

- **Cost-benefit analysis (CBA).** This is often used as part of economic evaluations, as many benefits can be assessed in monetary terms (reduction in demand on public services). This approach compares the costs of implementing the programme to the costs/savings achieved through a change in observed outcomes. The GMCA CBA model and accompanying unit cost data base provides a valuable and well-used resource (HM Treasury et al 2014). Running the full CBA model requires comprehensive data on costs and benefits. This includes full costs of intervention (including in-kind costs) and cost of business as usual, population at risk, level of engagement with intervention, level of retention, scale of impact (change in outcomes) and a way to calculate deadweight (for example, a counterfactual, but other methods are sometimes used).

- **Social return on investment (SROI).** SROI provides a framework for incorporating wellbeing impacts into an assessment of costs and benefits. The methodology is also grounded in direct stakeholder engagement to ensure that what matters to the people affected by an intervention is counted in the assessment (NEF Consulting, 2013).
- **Before and after analysis.** Relevant evaluations to date have tended to take a more limited approach to measuring economic impact, exploring change in service use and associated costs between the start of their support and end of support/a fixed time point (rather than comparison with a control group or distinct comparator group) (Cordis Bright, 2020; Crisp et al., 2020; Lamb et al, 2019b).
- **Simulation modelling.** This would involve modelling likely increases in, for example, rough sleeping, non-elective admissions, arrests and calculating associated costs. These would then be compared to the actual incidence and associated costs in areas included in the programme.

It may also be possible to use the results of previous studies to contextualise the findings of economic evaluation, in the absence of built-in comparator groups.

The study identified the following sources for unit cost tariffs and guidance:

- *Tariffs and tools*
- *Greater Manchester Combined Authority (GMCA, formerly New Economy): [Unit cost database](#)*
- *Home Office: [The economic and social costs of crime](#)*
- *Housing Association Charitable Trust (HACT) have a tool for applying [economic value](#) to wellbeing, based on WEMWBS.*
- *Possible methodological guidance*
- *GMCA: [Cost Benefit Analysis guidance](#)*
- *HM Treasury: [The Green Book annex](#)*

Figure 11: Methods for assessing economic impact of the Changing Futures programme

| Economic evaluation methods | Benefits and challenges ³⁵ | Example evaluations using this method |
|-----------------------------|---|---------------------------------------|
| Cost benefit analysis (CBA) | <ul style="list-style-type: none"> • Green Book recommended approach. • Requires comprehensive data including full costs of intervention, business as usual, outcomes and | Troubled Families |

³⁵ Where references are not provided, this information was drawn from the feasibility study consultation.

| | | |
|------------------------------------|--|---|
| | <p>deadweight (counterfactual) HM Treasury et al., 2014.</p> <ul style="list-style-type: none"> • GMCA cost benefit model offers guidance and unit cost resource. • Many stakeholders argued against the use of a cost-benefit analysis approach with this client group: very challenging to find a proportionate counterfactual approach, and complexity theory indicates that you cannot produce reliable return on investment figures for this type of programme. | |
| Social return on investment (SROI) | <ul style="list-style-type: none"> • Takes into consideration social outcomes that are typically difficult to quantify such as improved mental health and wellbeing (Allcock and Smith, 2018; NEF, 2014). • Most commonly used method of measuring cost-benefit of women's specific interventions in addressing multiple needs (Allcock and Smith, 2018). • Stakeholders' concerns in relation to CBA also apply to SROI. • May risk overlooking the real financial and fiscal savings from the programme. | <p>Women's voluntary and community organisations</p> <p>Birmingham Changing Futures Together Lead Worker Peer Mentor Programme</p> |
| Simulation modelling | <ul style="list-style-type: none"> • Addresses the difficulty in recruiting and obtaining data for individuals in a counterfactual group, but applies some robustness to the concept of an alternative trend. | |
| Before and after analysis | <ul style="list-style-type: none"> • No counterfactual so causal attribution limited. • Lower data requirements than full CBA. | <p>MEAM Approach</p> <p>Fulfilling Lives</p> <p>Pathway Homelessness Team</p> <p>Impact of CBT training</p> |

Costs to consider

- The specific costs to be considered will depend on decisions about priority outcomes to include in the evaluation and broader approaches to understanding impact. MHCLG and partners can then work out an appropriate approach to measuring economic impact, and which costs to include, based on this. However, stakeholders made several recommendations with regards to which costs should be considered in the evaluation:
- For many, the key interest was in **changes in costs to the system**. This encompasses wide-ranging measures that might be indicators of multiple disadvantage such as arrests, nights in prison, hospital attendance, substance use treatment, number of appointments not attended, and benefits accessed. Healthcare and the criminal justice system were felt to be the most likely places to see cost reductions relating to reduced service use.
- Some highlighted the importance of also considering **wider impacts** in financial terms, beyond immediate service use costs. Areas for exploration include the financial impact of the reduction in crime beyond criminal justice system costs, for example reduced costs to victims of theft, and application of costs to substance use (as project ADDER is currently seeking to do). However, these costs may be challenging to establish³⁶. The economic impact of changes to health or wellbeing outcomes can also be monetised, for example using QALYS (health) or WEMWBS (wellbeing).
- **Displaced costs** would also ideally be captured. Any cost benefit analysis approach should also include all cost increases for services related to the additional support to beneficiaries which may have led to cost reductions elsewhere in the system. This would be challenging to do for two reasons. Firstly, the full costs of the Changing Futures programme would be likely to exceed the amount of grant funding from MHCLG and may be complicated to identify and calculate. This means it is important to develop a systematic and consistent approach to local areas reporting project spend for evaluation purposes. Secondly, there may be displaced costs of which partners are not aware (e.g. increases in use of services that are not directly involved in the programme and whose use is not being monitored for programme purposes).

Key considerations

In addition to the methodological challenges to counterfactual and cost benefit analysis approaches set out in section 5.4.2 and 5.7.1, there are some further considerations for any assessment of economic and fiscal impact for the Changing Futures programme:

- **Limited likelihood of cashable savings.** Stakeholders highlighted how any reductions in service use by individuals do not reduce levels of service delivery overall, and that

³⁶ The GMCA Unit Cost database does include estimates of the fiscal, economic and social costs of crime.

any identified savings are therefore not likely to be cashable (see also HM Treasury et al., 2014).

- **Sector-specific savings.** Stakeholders reported it may be that increased spending in one sector, e.g. housing or health, saves money in another sector, e.g. criminal justice. Any economic evaluation will need to take account of this.
- **Possible increase in costs.** The evidence review shows that there is a possibility that increased engagement with services could lead to increase in service use and costs, at least in the short term (for example, Cordis Bright, 2020; Crisp et al., 2020; Field et al., 2019; Centre for Social Justice, 2017). Many stakeholders also highlighted the risk that an economic evaluation finds the costs go up. The analysis therefore needs to be contextualised, and consideration given to whether increases in costs indicate positive or negative outcomes. While monitoring costs is important, improved outcomes and better access to services need to remain the focus.

Several stakeholders commented that an assessment of economic impact would be hard but important, and that the programme should be ambitious with this despite the challenges. However, other stakeholders were sceptical of the value in pursuing economic analysis. Some of these commented that there was no point in pursuing something too complicated at the cost of being able to implement it; others argued that the economic analysis should not be the priority given the limited time available and the fact that cost reduction is not the primary purpose of the programme. Field et al (2018) argue that equity and parity of care should be key outcomes for evaluations rather than seeking cost reductions.

6 Process evaluation

6.1 Summary of findings

There is an important role for process evaluation of the Changing Futures programme. It can complement theory-based impact evaluation approaches and provide a more in-depth understanding of the detail of implementation and how this links to impact, providing supporting information to explore the mechanisms and context for change at individual, service and system level. Formative and iterative approaches may be useful, especially at the systems level. There is also scope to focus in on specific themes or areas of interest, in order to unpick these in further detail.

6.2 Role of the process evaluation

There was a strong consensus during the feasibility study consultation that a high-quality process evaluation should be a key priority for the evaluation of Changing Futures. Stakeholders highlighted the importance of this process evaluation regardless of their views of the feasibility and desirability of different impact evaluation levels and approaches. High-quality process evaluation can serve a number of important functions (HM Treasury, 2020a):

- Understanding whether an intervention or programme is being implemented as intended.
- Understanding whether the intervention or programme design is effective.
- Unpicking what is working well or less well and what might explain this.
- Exploring how context has influenced delivery.

The process evaluation is closely linked to theory-based, realist approaches outlined in section 5.3. As with the impact evaluation, it can be applied at both the level of specific interventions being delivered with people experiencing multiple disadvantage or at a systems level. It can also be used to understand processes relating to outcomes at service level.

In particular, stakeholders highlighted the crucial role of process evaluation in enabling judgements about:

1. The elements of the interventions, system activities or the programme which most contribute to any changes in outcomes, to move towards understanding how and why things work³⁷.
2. The replicability of the interventions, system activities and programme in other local areas, based on understanding the role of local contexts.

³⁷ Some stakeholders reported that this has been lacking in previous evaluations which focus on robust impact evaluation but are less able to comment on the specifics of what works.

3. The practicalities of replication, based on understanding the detail of how the interventions, system activities and programme have been set up/delivered.
4. Common barriers and enablers to implementation, and potential solutions to barriers.

Recommendation

We recommend that a high-quality process evaluation is prioritised alongside MHCLG and partners' chosen impact evaluation approach.

6.3 The benefits of formative or iterative approaches

One benefit of a process evaluation is that it can support formative or iterative approaches, which are likely to be of value for the Changing Futures programme. It is unlikely that a full action research model could be implemented, unless extensive resource is available to conduct this in all local areas involved in the programme. However, action research principles could be applied to some aspects of the evaluation. For example, findings from formative evaluative work could be fed into the learning networks which are intended part of the Changing Futures programme model and thus support the ongoing development and implementation of the programmes of work in local areas. Several stakeholders recognised a key role for formative findings to enable ongoing improvement but also provide early indication if the model is not working as successfully as anticipated so that any final impact and economic evaluation findings are not unexpected.

Implication for the evaluation

To enable useful formative findings and feedback, close links will need to be established between the central policy, delivery and evaluation teams. If resource allows, ideally the evaluation team would also build close links with local areas participating in the programme.

6.4 Scope for closer focus on specific themes or local areas

Through a combination of process evaluation and focused impact evaluation using theory-based approaches it would be possible to home in on particular aspects of the programme or interventions which might be of high interest in understanding what works. These might include, for example:

- Partnership structures³⁸.

³⁸ These were explored in detail in the MEAM Approach year 3 evaluation (Cordis Bright, 2020).

- Workforce development and support³⁹.
- System leadership.
- Specific interventions, such as the navigator role⁴⁰ or one-stop shop approaches.

Similarly, it might be possible to focus in on the experiences of sub-groups of individuals experiencing multiple disadvantage. This would be a way of exploring important potential differences in experiences and outcomes, as well as possible variation in what works for different individuals (based on common characteristics, previous experiences or presenting needs).

A further way to direct resource towards understanding some aspects of the programme in detail, would be to focus resource on specific local areas. Additional process evaluation or systems-level impact evaluation could be conducted in individual areas in order to be able to gather a wider range of perspectives and deliver more evaluation activity than might be feasible if the same methods were replicated across all local areas included in the programme.

It was not possible to review evidence of what works in supporting people facing multiple disadvantage systematically as part of this feasibility study but there is a range of existing research and evaluation available. This includes specific evidence of roles and approaches that are effective when working with this cohort (such as assertive outreach and care navigator roles). It also includes wider evidence of the efficacy of approaches or interventions that are currently used in relation to people experiencing multiple disadvantage when they are used with a range of other target beneficiaries (i.e. multi-disciplinary teams which are not specifically for those experiencing multiple disadvantage).

Appendix C collates a number of sources which discuss the efficacy of interventions (“what works”) and/or the specific features of effective interventions (“how it works”). As part of the programme or its evaluation it may be useful to produce a systematic review of the existing evidence base in order to benchmark interventions and identify key gaps in the evidence. This would help in understanding where specifically the programme and evaluation could add to the evidence base.

6.5 Key process evaluation questions

Potential evaluation questions for the process evaluation

1. Has the Changing Futures programme been implemented as intended?
2. How quickly were local areas able to mobilise to deliver the project?
3. Has the systems change activity in local areas been implemented as intended?

³⁹ The Fulfilling Lives evaluation recently published a series of briefings on different aspects of workforce development (Moreton et al, 2021a).

⁴⁰ A number of Fulfilling Lives partnerships and the national evaluation have explored this – see Moreton et al, 2021b; Broadbridge, 2018 and CRESR; 2016.

4. Have the interventions with people with lived experience funded in local areas been implemented as intended?
5. Which aspects of the programme delivery have been effective (at the level of the central programme, local areas/systems and specific interventions of interest)? Why is this?
6. Which aspects of the programme delivery have been less effective (at the level of the central programme, local areas/systems and specific interventions of interest)? Why is this?
7. What are the enablers to effective delivery at the three levels? How do they support effective delivery?
8. What are the challenges to effective delivery at the three levels? How do they inhibit effective delivery? Is there evidence of any solutions that have been identified to these challenges?
9. How has the national context for the Changing Futures programme influenced delivery of the model?
10. How have the different local contexts (including systems) for the Changing Futures programmes influenced

7 Engagement of stakeholders

7.1 Summary of findings

The quality of the evaluation is likely to be improved if it is co-produced at all stages with people with lived experience of multiple disadvantage. This is a resource-intensive process which requires adequate investment if it is to be delivered meaningfully and safely (Gujit, 2014; Zukosi and Luluquisen, 2002).

Whilst this lived experience input is valuable throughout the evaluation process, peer researcher models may be particularly valuable in improving the engagement of people using services in the evaluation (Terry and Cardwell, 2016). In addition to involving peer researchers, other key mechanisms for promoting the engagement of people using services include transparency about the purpose and nature of involvement, offering a range of ways to engage, using qualitative consultation methods and taking a tailored, strengths-based approach.

Effectively engaging local programme partners in the evaluation is crucial to its success. It is therefore important to generate buy-in by ensuring the evaluation is relevant to local partners and by being realistic about the extent to which they can support the evaluation. Building early relationships centred on the development of local theories of change may be effective, and will also support the theory-based approaches which are important to evaluations in complex systems (Bicket et al., 2020; Moore et al., 2018).

Ensuring dedicated local resource is available for evaluation activities is crucial, particularly in relation to the collation and sharing of individual-level data. This element of the evaluation is likely to be challenging and may also benefit from dedicated central resource from MHCLG.

7.2 Introduction

This chapter summarises the key findings of the feasibility study in relation to engaging stakeholders in the evaluation. This includes co-production of the evaluation with people with lived experience, engaging people as research participants, and engaging local partners who play an important role in supporting delivery of the evaluation. We draw primarily on views shared during consultation though this is supplemented by information from the evidence review. In particular, views and experiences shared by people with lived experience of multiple disadvantage have informed section 7.3 and 7.4.

7.3 Co-production

Definition of co-production

Co-production is a collaborative research practice. Wider stakeholders highlighted how a co-produced evaluation should be peer-led with peer researchers co-producing all elements of the research, from agreeing the research questions, to identifying outcomes and how to measure change, to gathering and analysing the data and presenting findings. This is importantly different to an evaluation that simply uses peer research methodologies

(e.g. peer researchers conducting surveys or interviews), which in isolation are a form of peer involvement rather than co-production.

The key principles of co-producing research include sharing power, trusting relationships, including and valuing multiple perspectives and ensuring everyone benefits (INVOLVE, 2018). Co-production recognises that involving the people the research is about in the evaluation process is ethically and morally the right thing to do (Gujit, 2014), and can also improve the quality of the evaluation. Wider stakeholders with experience of co-produced evaluations confirmed that it had improved the quality of the evaluation and changed the nature of the evidence generated.

Despite a wealth of literature which focuses on co-production, less evaluative work has been carried out to identify how co-production works in practice, or on the impact of co-production on research, practice, policy or population outcomes (Gagliardi et al., 2017).

Approaches to co-production

Co-production can be built on both peer research and participatory evaluation approaches.

Peer research

Peer research is a participatory research method in which people with lived experience of the issues being studied take part in the design, development and delivery of research and become the researchers, based on the assumption that such shared experiences bring a “unique quality to research” (Terry and Cardwell, 2016). Peer research can add high value to the evaluation of the Changing Futures programme:

- It can help to make the **research more inclusive** (for example, Terry and Cardwell, 2016). Stakeholders also highlighted how the involvement of peer researchers can engage people in the research who would not otherwise have contributed.
- People with lived experience and wider stakeholders highlighted how peer research can also lead **to very different and better data**. Peer researchers are often seen as more approachable and can have a greater understanding of the experiences and views of others with lived experience (Vaughn et al., 2018; Green and South, 2006).
- Peer research can also help to improve the relevance of the research and its acceptability to key stakeholders (Vaughn et al., 2018).

Participatory evaluation

There are other participatory approaches to research that could be considered for co-producing the evaluation, such as participatory action research where all stakeholders, including current beneficiaries, participate in the development and the delivery of the evaluation (Gujit, 2014). However, this approach may be less appropriate for the Changing Futures programme, as beneficiaries of interventions funded or catalysed by the Changing Futures programme may not yet be in a position to participate in the delivery of the evaluation and because delivering this across multiple sites would be highly resource-intensive.

Key considerations for effective co-production of evaluation

Recruitment

The feasibility study consultation elicited the following suggestions for recruiting people with lived experience to the evaluation peer research team:

- It can be helpful to **tap into existing groups and forums** for people with lived experience. Examples include Expert Link, Revolving Doors Agency/NECG, User Voice, local Fulfilling Lives expert groups and (for specific experience of domestic abuse) SafeLives and Women's Aid survivor forums.
- **Word of mouth** can also be a powerful approach to recruitment.
- **People like clarity**, so it will be important to provide examples of what has been done before or what people can get involved in.
- The MEAM Approach evaluation has worked consistently over several years with a **small core group**. Wider stakeholders commented that this has been an effective approach for maintaining consistent involvement from people and building trusting relationships.
- People with lived experience highlighted the importance of having a **diverse team involved**, to get a range of perspectives, and being proactive with recruitment to try and engage a mixture of people.

Involvement, training and support

The people with lived experience we consulted with identified two priority recommendations for co-production:

- **Involve individuals with lived experience from start to finish.** People with lived experience need to be involved throughout evaluation projects, rather than just for a specific task or length of time. The people we consulted with highlighted how lived experience input can be beneficial in the design of evaluations, when collecting data, when considering the implications of the findings and when presenting these findings to a wider audience. For example, when the researchers with lived experience are involved in reporting research findings, this can reduce the risk of information getting 'lost in translation' or findings peer researchers felt were important not being prioritised. In another example, the MEAM Approach and Blackpool Fulfilling Lives evaluations interview topic guides have been significantly improved by the input of peer researchers.
- **Provide training and support to peer researchers.** This demonstrates that peer researchers are valued, and helps to professionalise their role and ensure they have the skills and confidence to design and deliver research. The people we consulted with think that training should cover facilitating and conducting research, personal boundaries and safe sharing and the importance of neutrality. Wider stakeholders also highlighted the importance of training, particularly given that existing forums might focus

on advisory or lobbying functions, and therefore research training will likely be needed for people recruited via these routes⁴¹.

Other elements to involvement, training and support identified as important are:

- **Support for peer researchers.** Wider stakeholders highlighted the importance of ensuring effective support structures are in place. People with lived experience are highly motivated to support others and effect change. Support is needed to ensure that their involvement is a positive experience for them and does not jeopardise their safety and wellbeing, though obviously their choice to be involved should be respected and promoted. Specialist agencies may be required to help with the support function.
- **Ownership of work.** People with lived experience described the importance of peer researchers having ownership of the work being completed. This includes being involved in disseminating its findings and being kept informed about its use and impact.
- **Information and feedback.** People with lived experience and wider stakeholders both recommended the importance of a feedback loop to keep team members aware of project activity and progress, and so that they are aware of how their contributions are put to use.
- **Enabling and encouraging participation.** The people with lived experience we consulted identified the following as important:
 - Ongoing rather than one-off activities enable people to become engaged and commit.
 - Flexibility in how people engage with the research is required to accounting for individuals' capacity to be involved and avoid overwhelming anyone.

Approaches to payment or recognition

Peer researchers should receive recognition of their time and contribution. In peer research projects this is often by way of vouchers though in some cases peer researcher are paid directly; this latter approach requires support and infrastructure to ensure that any paid work does not jeopardise people's existing financial arrangements, such as benefits they receive. In some organisations, peer researchers are paid. However, wider stakeholders have also highlighted the role of upskilling peer researchers and longer-term benefits such as internet connectivity, which can help ensure that peer researchers are benefitting from the research in return for contributing their time and expertise.

Skills and resourcing

Co-production is "time-consuming, ethically complex, emotionally demanding, inherently unstable, vulnerable to external shocks, subject to competing demands and expectation" (Flinders et al, 2016, p. 266). There are some key considerations in relation to this:

- **Sufficient time is required.** Co-production takes more time (Flinders et al., 2016). This will need to be built into project timescales. For example, people with lived experience highlighted how they need time to reflect and deliberate. In our experience of delivering

⁴¹ There are accredited peer research training courses which can be delivered, such as OCN NVQ Level 2 peer research training.

co-production for the MEAM Approach evaluation, it also tends to necessitate additional steps in the planning and delivery of most activities (e.g. longer sign-off processes on research tools, methods and outputs; and more project administration time in setting up and preparing for field work).

- **Considerable resourcing is required** (Oliver et al., 2019). This is necessary to account for the longer time scales and additional work for the core evaluation team, to provide training and support to peer researchers, and to appropriately compensate peer researchers for their time and contributions. For this reason, wider stakeholders emphasised that the co-production element of the evaluation needs to be resourced properly if it is to be attempted at all.
- **Skilful facilitation.** Stakeholders highlighted how participatory approaches and co-production require skilful facilitation to ensure peer researchers can fully contribute and that the impartial evaluative focus is maintained (Oliver et al., 2019). Care also needs to be taken with power balances, for example if delivering consultation alongside non-peer researchers or during workshops. As a consequence, evaluators and any staff members involved in co-producing the evaluation may require additional training, guidance and support.

Other considerations

- **Representation.** People who are able to participate as peer researchers tend to be further on in their recovery journey than the people facing multiple disadvantage on whom the research will focus. Wider stakeholders therefore highlighted that there is a need to take care with claims about representation.
- **Experience and expertise of non-peer researchers.** People with lived experience and wider stakeholders highlighted the benefits of non-peer researchers and peer-researchers co-delivering interviews, so that non-peer researchers can help with the development of a more thematic and systematic approach during interviews where appropriate.

7.4 Engaging people using services in the evaluation

Most effective methods for consulting with people using services

A lack of involvement of people using services is a fundamental limitation of some evaluations (Bell, 2014). People with lived experience and wider stakeholders emphasised the importance of offering choice and flexibility when engaging people using services in evaluation. This might include offering a range of different consultation methods, such as group-based and 1-to-1 methods (face-to-face or remotely delivered), or the option of written or alternative feedback mechanisms, such as apps or dictaphones.

Most people with lived experience and wider stakeholders suggested that interviews and other in-depth, semi-structured methods result in better engagement and higher-quality data than surveys (although a small number of wider stakeholders pointed to the successful use of surveys in the past).

In some cases, it may also be beneficial for support workers to take part in the consultation, either to increase people's confidence to engage or to add their own

perspectives on people's experiences, progress and challenges. People with lived experience also pointed out that consultation could potentially be enriched by including people's family members, friends or others in their support network who could comment on whether and how things had changed as a result of the programme. Another method for enhancing the quality of people's feedback which was suggested by a small number of wider stakeholders was speaking to them alongside reviewing other sources of evidence about their circumstances and experiences, such as case notes.

Key considerations for effective consultation

Advance and informed consent

The need for informed consent is a given for any ethical research and evaluation but there are some specific points of relevance to gaining the consent of people experiencing multiple disadvantage who are using services (or could be) which were highlighted by people with lived experience and wider stakeholders:

- **Clarity and transparency about use of data.** For a variety of reasons, some people who experience multiple disadvantage might have particular concerns about confidentiality. In addition, they may already feel that they are asked to share information repeatedly by a range of different services. Understanding exactly what data is being collected, how this process is different to sharing information as part of engaging with services, and how this data will be used and shared is therefore especially important; clear explanations of this need to be built into the informed consent process. This applies equally to consent to collate administrative/service-collated data as to consent prior to consultation. For consultation, it includes making it clear that people's views on services will not be shared with those who deliver them; people often feel concerned that negative feedback will be shared with services.
- **Making it clear that consent can be withdrawn.** People may wish to withdraw consent at a later date and need to be able to do this easily. It is important to build in simple processes for withdrawing consent and to make it clear up-front that it is OK to do so. This includes making it clear that the choice not to participate will not affect the support someone receives.
- **Avoiding negative impact on engagement with services.** This is more applicable to consent for sharing administrative/service-collated data; asking for consent to do so can affect engagement with services if people have not built adequate trust before they are asked. This means that gaining advance consent at the start of engagement with services is not always practical.

Encouraging engagement

The people with lived experience and wider stakeholders we consulted identified several ways to encourage people using services to engage in consultation:

- **Use trusted sources to identify and broker contact with potential participants.** This might include support workers or volunteers in services where people are engaging or – for people experiencing multiple disadvantage who are not engaging with services – people in other community venues.

- **Provide clear information up-front about what is involved in participation.** In addition to enabling informed consent, transparent information about the consultation can also encourage engagement.
- **Emphasise how taking part can contribute to better support.** People may be more motivated to take part they understand how sharing their views fits into the bigger picture of improving support. Therefore it can be useful to explain that the purpose of evaluation is to understand what works and what does not, to help with future designing, delivering and funding of services. People with lived experience also highlighted that it is useful to provide concrete examples of how previous evaluation and research findings have informed decision-making.
- **Involve peer researchers.** As discussed in section 7.2, peer researchers conducting or co-conducting consultation can encourage participants to engage more readily. If peer researchers are local to the area, participants may already know them to some degree.
- **Consult people in a familiar, comfortable environment.** People are more likely to attend and engage in consultation if it takes place in a setting in which they are comfortable. This could be a familiar and supportive service. People with lived experience also emphasised how food and hot drinks, or walking and talking, could create a more relaxed atmosphere. In addition, they reported that the ratio of evaluators to participants is important as it can be intimidating to be approached by a group of researchers. This is an important consideration for co-delivered consultation, such as paired interviews by a peer and non-peer researcher.
- **Be friendly, respectful and open.** These are important ways to build people's confidence to share their views.
- **Use accessible research tools.** It is important to developing research tools which use straightforward and accessible language and ask questions that people can relate to. Co-production approaches with people with lived experience can help with this.
- **Meet people's accessibility needs.** This might include additional learning needs, languages other than English or accessibility needs relating to health issues or disabilities.
- **Share questions in advance.** Providing the questions in advance might boost people's confidence to engage by reducing the element of the unknown. It also enables them to prepare their thoughts in advance if they would like to, which might lead to more reflective responses.
- **Provide incentives.** Voucher incentives were seen to be a relatively simple and safe way to incentivise participation. People with lived experience reported that it works well when incentives are provided at the start of the consultation. This can be a means to build trust and also to reassure participants that the incentive is not contingent on how they respond in the consultation.
- **Cover expenses.** It is important to consider and cover any expenses that participants might incur, such as travel or childcare costs.
- **Consider childcare.** For parents/carers to take part, providing childcare may also be important.

Ensuring participation is a positive experience and does not cause harm

People with lived experience and wider stakeholders described a number of ways to help to make participation a positive experience. Taking person-centred and strengths-based approaches in consultation can help to make the experience a positive one, as can reassuring people that they can choose not to answer questions on topics they would prefer not to discuss. Reflecting back what has been discussed at the end of the interview is also an important way of clarifying and ensuring researchers and participants have a shared sense of the key discussion points to inform the evaluation. This may reinforce to participants that their views and experiences have been heard.

It is also important to check in with participants after consultation and to signpost them to support options if the consultation has brought up difficult issues⁴².

A final way of making participation a positive experience is completing the feedback loop so that participants understand how their participation has contributed to the evaluation findings and – where this is known – any impact of the evaluation on decision-making about future support.

Recognising selection/non-response bias

Wider stakeholders highlighted an important limitation of most approaches to consulting with people with lived experience as part of programme evaluation; in general, opportunity sampling via services is used as the most feasible and cost-effective method of recruiting participants. However, this creates selection or non-response bias because it precludes the involvement of people who are not engaging with services and it reduces the likelihood of engagement by people whose experiences of those services are more negative. Reflecting on ways to reduce this bias would be helpful. For example, people with lived experience felt that using local peer researchers might be a means of encouraging other local people not in direct contact with services to take part in the evaluation. In the event that it cannot be avoided, findings from consultation should be framed in ways that recognise this limitation.

7.5 Engaging local programme partners

Local programme partners are crucial to the success of the evaluation, taking roles in delivering key aspects of the evaluation but also participating in consultation and other key evaluation activities. Our experience of delivering evaluations of programmes involving central funding, local delivery and centrally-commissioned evaluations suggests that there are a number of important elements which need to be built into evaluation design and delivery in order to maximise the chances of successfully engaging local programme partners in the evaluation. These are:

Finding ways to encourage buy-in to the evaluation

MHCLG may wish to consider mandating local areas' involvement in some aspects of evaluation activity (such as supporting administrative/service-level data collation). In general, however, the quality of local areas' participation in and support for evaluations is

⁴² This in addition to reporting any safeguarding concerns via agreed routes.

determined by the extent to which local partners see it as relevant and useful to them. As a result, it is important to use the evaluation design and set-up phase as a means of building buy-in to the evaluation from local partners in both strategic and operational roles.

For example, consulting on local priorities for the evaluation and incorporating these where possible can be a useful way to encourage local areas to invest in the evaluation. In addition, evaluation activities which build local capacity (e.g. setting up data collection processes which will be meaningful and useful to local areas beyond the evaluation itself) or serve other functions alongside evaluation (e.g. worker-completed case studies that can support reflective practice) are useful ways to encourage buy-in.

Understanding differences in local programmes, systems and contexts

For evaluators of national programmes being delivered in multiple sites, it is often challenging to understand local differences in context and delivery in detail because the evaluation resource does not allow for adequate engagement with local areas outside of key field work activities. Time and resource ideally needs to be built into the evaluation to enable evaluators to work with local partners to build this understanding.

This is particularly important for the Changing Futures evaluation, which seeks to understand systems change activities and their impact on the system, and also to account for ways in which local systems might influence the implementation and success of direct interventions with people experiencing multiple disadvantage. As discussed in sections 5.3 and 5.6.3, theory-based, participatory and multi-perspectival approaches are important in systems evaluation and these need to take place at local level.

If this resource is not available, it is important to be realistic about the level of detail that the evaluation can achieve in relation to understanding systems change or how local contexts influence delivery and outcomes.

Recommendation

We recommend that an early requirement for the evaluation is to work with local areas to develop or refine their local theory of change for the Changing Futures programme activities. Ideally, if resource allows, this would involve participatory approaches and MHCLG representatives would be included in the process.

This would serve a dual purpose. Firstly, it would enable evaluators, local programme partners and MHCLG stakeholders to build relationships and a shared understanding of local programme models and how these relate to the central model up-front. Secondly, it would begin to build evaluators' understanding of the local system, thereby generating early evaluation findings and information to guide later evaluation activity.

It would also ensure that the format of theories of change across the local areas is standardised where possible and that the relationship between the central theory of change and local theories of change is clear. This will make it easier to compare theories of change and to ensure that they support theory-based approaches within the evaluation.

Overcoming challenges with data collection

In the Fulfilling Lives and MEAM Approach evaluations, generating high-quality, linked, individual-level data has proved challenging⁴³. Some of the main challenges are described in section 2.3.2.

If resource is available centrally, it may prove more efficient for centrally-available data to be collated by MHCLG and partners and for MHCLG to work closely with local partners to negotiate access to data which is held locally. Whilst this might be time-consuming, it may represent a good investment of time if it aids the evaluation whilst also contributing to longer-term improved information sharing in local areas (i.e. one of the priority system-level outcomes for the programme).

If local partners will be required to lead on the data collation, specific resource needs to be allocated to this activity. Building it into someone's role will help to clarify responsibility for this. Coordinating the data collation is often built in at an operational level and this makes sense in terms of obtaining consent and inputting data. However, this role is often allocated to someone with limited experience in data collation (e.g. a team leader in a frontline service) and who would therefore benefit from specialist input from local data analyst colleagues about how to request the specific data that is required and how to arrange data sharing.

In addition, support from senior and strategic stakeholders is crucial in arranging the required information sharing. This needs to be made clear in early discussions with local areas, to help them to plan and establish effective data collection processes to support the evaluation. Significant practical support with data collation is likely to be required from evaluators and resource needs to be built-in for this.

It is also important to find ways to ensure that that the volume of data is manageable. The two main approaches to this are reducing the number of variables included in the dataset (by focusing in on a smaller number of priority outcomes and indicators) or reducing the number of individuals included in the dataset (by selecting only a sample of clients from each local area to be included in this element of the evaluation).

Finally, the quality of data collection is likely to be higher if the data is also of use to the local projects responsible for collating it. For example, the Fulfilling Lives quarterly data dashboards have been helpful in providing an accessible tool for projects to explore their data and compare it with others.

7.6 Engaging evaluation audiences

A number of stakeholders highlighted two important ways in which the evaluation's target audiences might shape decisions about the evaluation reporting:

- **Maintaining the profile of the programme:** Regular, shorter reports might help to keep the profile of the programme high and provide digestible findings in real-time. This could work particularly well if more focussed evaluation is conducted into specific

⁴³This was also a challenge in the Troubled Families evaluation (MHCLG, 2019).

aspects of the programme, which might lend themselves well to shorter, thematic summaries.

- **Providing relevant findings to different audiences:** There are a range of target audiences for the evaluation, including central policymakers and funders as well as local commissioners and delivery partners. As well as informing the evaluation's focus and design, this also has implications for reporting because it will be important to convey findings in ways to ensure that different audiences receive and are able to act upon the key findings of most relevance to them.

8 Conclusions

This important evaluation is likely to be challenging to deliver, particularly from an impact evaluation perspective. There is a range of reasons for this, but the chief of them are:

- The multi-site nature of the programme, with flexibility for local areas to design and implement local programmes of work that they believe fit their local contexts. This means that interventions and programmes of work will differ in each local area.
- The desire for robust evaluation of impact at both individual- and system-level, and the possible tensions between the approaches required for these two different strands of impact evaluation.
- The personal nature of priority outcomes for individuals receiving support, which runs counter to the idea of an evaluation focusing on pre-determined individual-level outcomes.
- The multiplicity of individual-level outcomes which might be affected and the absence of a clear lead variable for understanding impact via experimental or quasi-experimental methods.
- The absence of an obvious counter-factual control group for experimental and quasi-experimental methods.
- The complex and “noisy” systems into which the programme is being introduced, which will all differ in structure, relationships and context. This poses challenges for understanding and attributing impact via traditional evaluation approaches.
- The innovative nature of interventions and related activities, where an emphasis is placed on learning by doing, meaning that interventions and activities are unlikely to remain fixed for the duration of the programme. This also means that there are limited examples of previous similar evaluations.
- The relatively short timescales for the programme and intervention, making it important to determine which outcomes are most likely to be achievable and demonstrable in the timescales and to be realistic about evaluation approaches that can be set up and delivered in the two-year window.

It may not be feasible to resolve all of these challenges within the resource available to the evaluation. As a result, MHCLG and partners will need to decide on the highest priorities for the evaluation. This will help to identify the most relevant approaches to pursue and evaluation resource can then be focused on solutions to challenges for these specific approaches and the appropriate methods. In order to support this decision-making, we have attempted to make suggestions throughout this report about feasibility of methods and recommendations on priority evaluation approaches.

Appendix A: Summary of outcome measurement tools identified during feasibility study

Please note that this is not an exhaustive list of the tools that could be used to measure all outcomes relevant to Changing Futures, because a more systematic review of tools was not in the scope of the feasibility study. Instead it provides further detail of the tools which were most commonly referenced by interviewed stakeholders or which were most evident in the reviewed literature.

| Data collection tool/approach | Applicable at Individual-level | Applicable at Service-level | Applicable at System-level | Example evaluations using this approach | Benefits and challenges |
|-------------------------------|--|-----------------------------|-----------------------------------|---|--|
| Administrative data | <ul style="list-style-type: none"> Health service use (CSUs, hospital trusts, NHS Digital) Offending (PNC) Homelessness / housing stability (local authority / MHCLG H-CLIC) Benefit claims and payments by type (DWP) Drug treatment use (NDTMS (PHE)) | | Changes in use of public services | Troubled Families Justice Data Lab evaluations Health and homelessness in Scotland (research) | <ul style="list-style-type: none"> Reduces need for additional data collection from participants. Potential for creating a statistical control (e.g. as in Waugh et al., 2018). Requires consent and personally identifiable data. Sourcing administrative data can be time-consuming and challenging. |

| | | | | | |
|--|---|--|--|---|--|
| | <ul style="list-style-type: none"> Time in prison (Prison National Offender Management Information System) | | | Violence Reduction Units (Feasibility study) | |
| Bespoke / novel questionnaires | <ul style="list-style-type: none"> Satisfaction with services / support | <ul style="list-style-type: none"> Staff confidence, autonomy and beliefs Fidelity to a service model (e.g. Housing First) Perceptions of service effectiveness | Service coordination and collaboration | Impact of CBT training MEAM Approach Housing First evaluations Troubled Families | <ul style="list-style-type: none"> Ability to collect data on specific issues / outcomes for which validated measures do not exist. Non-standard so unable to benchmark against other studies. |
| CORE-OM (Clinical Outcomes in Routine Evaluation – Outcome Measure) | 34 items of psychological wellbeing health covering 4 domains: <ul style="list-style-type: none"> Well-being Symptoms | <ul style="list-style-type: none"> Can also be used to measure staff wellbeing | | Psychologically Informed approach to homelessness prevention | <ul style="list-style-type: none"> Designed for measuring outcomes of psychological therapies (Evans et al., 2009. Said to be reliable and valid (Evans et al., 2009; Maguire et al., 2017). |

| | | | | | |
|--|--|--|--|---|---|
| | <ul style="list-style-type: none"> • Functioning • Risk | | | Impact of CBT training | <ul style="list-style-type: none"> • Shorter versions available (Evans et al.). |
| EQ5D | <ul style="list-style-type: none"> • Mobility • Self-care • Usual activities • Pain/discomfort • Anxiety/depression | | | Adults facing Chronic Exclusion pilots | <ul style="list-style-type: none"> • Well-established, reliable measure used in health research and clinical trials (EuroQol, 2018). • Can be used to calculate QALYs (HM Treasury 2020b). • Flexible methods for administration (EuroQol, 2018). • Fairly short (EuroQol, 2018). |
| Homelessness Outcomes Star | <ul style="list-style-type: none"> • Motivation and taking responsibility • Self-care and living skills • Managing money • Social networks and relationships • Drug and alcohol use • Physical health • Emotional and mental health | | | Fulfilling Lives MEAM Approach Tackling Multiple Disadvantage | <ul style="list-style-type: none"> • Designed as a keywork tool to support reflection and action planning. • Also used to measure outcomes. • Widely used by homelessness agencies (Mackeith, 2014). • Questions over inter-rater reliability (Mackeith, 2014; Johnson and Pleace, 2016). • May not be appropriate for evidencing effective services (Johnson and Pleace, 2016). |

| | | | | | |
|--|--|--|--|--|---|
| | <ul style="list-style-type: none"> • Meaningful use of time • Managing tenancy and accommodation • Offending | | | | |
| Intermediate Outcomes Measurement (IOMI) | <ul style="list-style-type: none"> • Resilience • Wellbeing • Agency / self-efficacy • Impulsivity / Problem-solving • Motivation to change • Hope • Interpersonal trust • Questions on practical problems • Optional questions on relationships with staff | | | Engager programme for prisoners with common mental health problems | <ul style="list-style-type: none"> • Covers a wide range of areas which are applicable to the Changing Futures programme, including both recovery-related outcomes and more practical areas, such as housing and finances (Liddle et al., 2019). • Focuses on interim outcomes which might be achieved and visible before longer-term outcomes (Liddle et al., 2019). • Has its roots in offending-related programmes, which might feel less applicable to some cohorts supported by Changing Futures. |
| Maslach Burnout Inventory | | <ul style="list-style-type: none"> • Staff burnout • Emotional exhaustion • Depersonalisation | | Impact of CBT training | <ul style="list-style-type: none"> • Long-standing, validated and widely used (Maguire et al., 2017). |

| | | | | | |
|---|--|---|--|---|--|
| | | <ul style="list-style-type: none"> • Personal accomplishment | | | |
| Recovery Star | <ul style="list-style-type: none"> • Managing mental health • Physical health and self-care • Living skills • Social networks • Work • Relationships • Addictive behaviours • Responsibilities • Identity and self-esteem • Trust and hope | | | | <ul style="list-style-type: none"> • Like the Homelessness Outcome Star, used as a keywork tool (Dickens et al., 2012). • Focuses on mental health and wellbeing (Dickens et al., 2012). • Some indication of validity and reliability (Killaspy et al., 2012; Dickens et al., 2012). • Requires paid for licence. (Triangle, 2012) • May not be applicable for use with people experiencing multiple disadvantage. |
| New Directions Team Assessment (Chaos Index) | <ul style="list-style-type: none"> • Engagement with services • Self-harm • Risk to/from others • Stress/anxiety • Social effectiveness • Alcohol/drug abuse • Impulse control | | | Fulfilling Lives MEAM Approach | <ul style="list-style-type: none"> • Designed as an assessment / referral tool (Rinaldi et al., 2008). • Deficit based. • Completed by worker – no client input. |

| | | | | | |
|---|--|--|--|---|---|
| | <ul style="list-style-type: none"> • Housing stability | | | | |
| REC-CAP | <ul style="list-style-type: none"> • demographic characteristics • quality of life and satisfaction • barriers to recovery • services involvement and needs • personal recovery capital • social recovery capital • involvement with recovery groups local community • social support • recovery commitment | | | REC-CONNECT pilot | <ul style="list-style-type: none"> • Designed as assessment and recovery planning tool (Best et al., 2018) • Psychometrically validated (Cano et al., 2017). • Strength-based (Best et al. 2017). • Captures barriers to recovery. (Best et al. 2016) • Online tool available (Advanced Recovery Management System, n.d.). |
| Residential Time-Line Follow-Back Inventory | <ul style="list-style-type: none"> • Homelessness • Housing stability • Accommodation types/patterns | | | Housing First in five Canadian cities | <ul style="list-style-type: none"> • Said to be valid and reliable (Tsembris et al., 2007). • Used with people affected by homelessness, substance use and with mental ill-health (e.g. Aubry et al. 2015). |

| | | | | | |
|---|---|--|--|---|---|
| | | | | | <ul style="list-style-type: none"> • May require research professional to administer (Tsembris et al., 2007). • Could be time-consuming (Tsembris et al., 2007). |
| TICOMETER | | <ul style="list-style-type: none"> • Level of trauma-informed care in working practices | | | <ul style="list-style-type: none"> • Designed as organisational assessment and monitoring tool, rather than outcome measure (Bassuk et al., 2017). • Can be used to determine training needs and trauma-informed policies (Bassuk et al., 2017). • Captures staff not service user perceptions (Bassuk et al., 2017). |
| Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS) | <ul style="list-style-type: none"> • Subjective mental wellbeing | | | Blackpool Fulfilling Lives Integrated Personal Commissioning Programme | <ul style="list-style-type: none"> • Well validated population-level measure (Warwick Medical School, 2020). • Widely used (Warwick Medical School, 2020). • Uses positive statements. • Shorter version available (Warwick Medical School, 2020). • Results can be “monetised” for use in economic evaluation (Social Impacts Task Force, no date). |

| | | | | | |
|--|--|--|--|--|--|
| <ul style="list-style-type: none">• Working Alliance Inventory | | <ul style="list-style-type: none">• Client-worker relationship | | | <ul style="list-style-type: none">• Well-recognised tool referenced by a number of interviewed stakeholders.• Long version includes 36 statements for client-completion using a 7-point scale (Horvath, 1984, 1981).• Shorter forms available. |
|--|--|--|--|--|--|

Appendix B: Summary of findings on the approaches to conceptualising recovery

Introduction

The concept of recovery, rooted in the fields of substance use and mental health, provides a useful way to think about outcomes that are important in supporting people affected by multiple disadvantage. Definitions of recovery vary across the two fields, but emphasise a purposeful, fulfilled and satisfying life (Davidson et al., 2008; Shepherd et al, 2008; Anthony, 1993) which maximises health and wellbeing (UK Drugs Policy Commission, 2012; Best and Laudet, 2010).

Key principles of recovery

Identifying common themes across recovery literature can be beneficial to aiding recovery-oriented practice. Originating from the mental health recovery field and based on a review of evidence, Leamy and colleagues (2011) used the acronym CHIME to highlight five principles of recovery. These five principles (Connectedness, Hope, Identity, Meaning and Empowerment) are discussed in turn below. This approach to mental health recovery fits well with what is known about recovery from substance use and as such, is now also commonly used within the substance use literature (Best, 2019; Best et al., 2018). A recent evidence review conducted on behalf of St Mungo's (Bennett, 2020) demonstrates that many aspects of these concepts, such as positive relationships, hope and empowerment, are also relevant to recovery from homelessness and to people affected by multiple disadvantage.

Connectedness

Mental-ill health and substance use are commonly associated with social isolation (Chou et al., 2011). More so, such isolation is often exacerbated by the stigmatisation associated with using substances, can put individuals at risk of relapse (Best and Lubman, 2017). With this in mind, providing individuals with the opportunity to feel a sense of connection and belonging (for example, to other peers in recovery, or through professional support services), should be a critical component of recovery trajectories (Dingle et al., 2015). Forming connection to others who are also in recovery, or supportive of an individual's recovery attempts, is a known predictor of recovery outcomes (Longabaugh et al., 2010).

Hope

Having positive aspirations and hopes about living a life free of substances is known to contribute to the motivation to remain sober (Korcha et al., 2011). If individuals can be connected to those who are supportive of their recovery attempts, then the positive social support this in turn provides helps to promote the "belief that change is possible, generating a sense of hope that energises attempts to manage change" (Best, 2019, p. 6).

Identity

For those experiencing multiple disadvantage, the development of a new recovery-oriented identity can help to combat issues of stigma and discrimination. The formation of

these new identities is internalised and learned through exposure to recovery peers and recovery groups that result in a sense of belonging and a set of social supports that protect against future relapse (Buckingham et al., 2013). Providing opportunities for individuals to participate in meaningful activities can help individuals to begin to form pro-social networks and identities (Collinson and Best, 2019; Best et al., 2017; Best et al., 2015).

Meaning and empowerment

Having an increased sense of purpose in life is a critical aspect of recovery trajectories (Del Vecchio, 2012) and having a sense of meaningfulness is important in overcome adversity. Such feelings are also associated with improved health and wellbeing (Martin et al., 2011). Having the opportunity to engage in meaningful activities not only contributes to feelings of meaningfulness but can also generate feelings of empowerment (Best, 2019). Empowerment (which is linked to self-esteem and self-efficacy), is a known predictor of sustained recovery (Moos, 2007).

Recovery capital

Recovery capital provides a framework for capturing the impacts of recovery-oriented practice outline above. Cloud and Granfield (2008) introduced the theoretical framework of recovery capital (RC) as a strengths-based metric for measuring recovery progress. RC is defined as "the breadth and depth of internal and external resources that can be drawn upon to initiate and sustain recovery from AOD [alcohol and other drug] problems" (Granfield and Cloud, 1999, p32) and has gained much traction over recent years amongst substance use recovery.

While there have been numerous classificatory systems of RC, Hennessey's (2017) systematic review provides an overview and historical context of these developments. For the purpose of this evidence review, White and Cloud's (2008) framework is used, categorising recovery capital into three components: personal, social and community capital. These can reflect outcomes on an individual (personal), meso (social) and macro (community) level. It is noted that RC is a great predictor of recovery success and it is believed that both the quality and quantity of RC that an individual retains play a major role in predicting recovery success, and engagement with treatment.

It has, however, been argued that specific factors such as age, gender, mental health and incarceration may be forms of 'negative capital' (Cloud and Granfield, 2008), hindering an individual's ability to accumulate RC.

Personal capital

White and Cloud's conceptualisation of personal capital brings together physical and human capital. Physical capital includes material resources such as money and having essential needs met such as food and housing (Hennessey, 2017). Human capital covers personal characteristics needed to achieve goals, such as knowledge, skills, mental health, problem-solving capacities, physical health, self-esteem and self-efficacy, hopes and aspirations (Hennessey, 2017). The aspects associated with personal capital can help to understand outcomes important for those experiencing multiple disadvantage as such changes are important to note within a person's process of change (Simoneau and Bergeron, 2003) and are known to predict sustained recovery (Best et al., 2010).

Social capital

Social capital includes access to information, social networks and mutual support through the development or maintenance of supportive relationships (Ruiu, 2016; Best et al, 2015). To support an individual's process of change, an individual must shift from 'deviant' or 'substance using' social networks to networks that are supportive of sobriety and recovery. It is known that post substance use treatment outcomes can be strongly predicted by the number of people in an individual's social network (Zywiak et al., 2002) and it has been demonstrated that increased support for sobriety within a social network is associated with reduced risk for relapse (Litt et al., 2009).

The formation of pro-social networks is an important outcome for those experiencing multiple disadvantage. Jetten et al. (2012) have demonstrated that belonging to groups in which members have shared pro-social aspirations benefits both wellbeing and physical health. However, subsequent work on homelessness (Johnstone et al., 2015) has shown that membership of marginalised and excluded groups does not confer the same wellbeing benefits and can indeed be a barrier to health and wellbeing.

A key component of recovery support should include the incorporation of predominantly supportive social networks (Dingle et al., 2015). Our sense of self is derived from our membership in certain groups and resulting positive identity formation can structure and change a person's perceptions and behaviour (Haslam, 2014). If social identity change can therefore be encouraged by surrounding an individual with recovery-oriented support networks, it may provide the individual with a greater chance of envisaging their identity change and working towards it.

Community capital

Community capital can be understood in terms of community resources, such as activities and transport links, groups and facilities, recovery communities, as well as non-stigmatising attitudes within the community (Best and Laudet, 2010). Whilst access to community capital may be reliant on communities more broadly, engagement in pro-social, meaningful activities should be an encouraged aspect of an individual's recovery journey and as such, individuals should be supported to do so. Community engagement can be defined as "involvement in interpersonal interactions outside the home, including social, leisure, community activities and work" (Goll et al., 2015, p. 2).

When individuals early in addiction recovery are effectively linked into positive community resources, it is anticipated that their levels of personal, social and community capital will grow as a result. Literature also demonstrates such engagement in groups that are supportive of the individual's recovery attempts is required for long-term recovery (Best et al., 2017).

An individual's level of community engagement may also increase in parallel with their recovery journey – with a shift from lower levels of engagement (such as simply attending a group) to higher levels of engagement where the individual is perceived to be an active contributor (such as voluntary work). Themes such as 'making good' and generativity are often noted within the desistance and recovery literature (Nugent and Schinkel, 2016). Individuals are described as wanting to 'redeem' their past selves by making good and giving back to communities and their families.

Appendix C: Examples of evidence sources on “what works” in supporting people experiencing multiple disadvantage

| Intervention ⁴⁴ | Example evidence sources |
|---|--|
| Multi-agency / multi-disciplinary working | <ul style="list-style-type: none"> • Literature review on multi-agency working identifies how multi-agency working improves access to services and increases the preventative focus of services for clients who access more than one service or organisation (Atkinson et al., 2007) • MEAM Approach evaluation highlights partnership working at operational and strategic levels as crucial to efficacy and outcomes (Cordis Bright, 2020). • Rapid evidence review on key factors for successful multi-disciplinary team working: (Cordis Bright, 2018) • Practical advice and guidance on establishing and maintaining multi-agency partnerships (Cheminais, 2009) • Systematic reviews on the efficacy and implementation of relevant multi-agency arrangements: <ul style="list-style-type: none"> ○ Integrated care (Baxter et al., 2018; Keane et al., 2018; Flanagan et al., 2017). ○ Dual diagnosis support (Fantuzzi et al., 2020). |
| Navigators and assertive outreach | <ul style="list-style-type: none"> • Research from WY-FI Fulfilling Lives project describes the key role that Navigators play in engaging beneficiaries and ensuring they access the necessary support (Centre for Regional Economic and Social Research, 2016). • National and local Fulfilling Lives evaluations have explored what makes an effective multiple disadvantage coordinator (Moreton et al., 2021b; Broadbridge, 2018). • Studies found evidence of positive impacts of assertive community treatment teams on: |

⁴⁴ Other approaches and interventions that it may be useful to explore within an evidence review of “what works” include co-production, treatment of co-occurring substance misuse and mental health problems (dual diagnosis) and personal budgets.

| | |
|---|---|
| | <ul style="list-style-type: none"> ○ Hospital bed use for patients, which reduced during the period of support by the treatment team when compared to the period prior to support (Sood and Owen, 2014). ○ Treatment flexibility, engagement with treatment, quality of life, meeting needs and hospital admissions (Nugter et al., 2016). |
| <p>Trauma informed practice / psychologically informed approaches</p> | <ul style="list-style-type: none"> ● UK government best practice recommendations for working with people experiencing homelessness suggest creating PIEs, where approaches and interventions “recognise and work with the levels of emotional trauma that accompany, and in many cases precede, an individual becoming homeless” (Johnson and Haigh 2010). ● Research has shown that PIEs achieve significant positive change for people experiencing multiple exclusion/deprivation and with histories of compound trauma. This includes improved housing outcomes, improved behaviours, improved use of services and improved mental health. Data also suggests that PIEs achieve more positive outcomes than services not run in this way, and that staff benefit alongside clients (Cockersell, 2016). ● The literature provides an overview of key elements of psychologically-informed environments, enabling features, their reception in the homelessness sector and learning in implementing them (Breedvelt, 2016; Turley et al, 2013). |
| <p>Peer support</p> | <ul style="list-style-type: none"> ● Systematic review has shown that peer support is associated with similar psychosocial, satisfaction, clinical and service outcomes to those achieved by other professionals (Pitt et al, 2013). ● A 2017 review of existing studies found that peer support can have a significant impact on the quality of life for people experiencing homelessness, their drug and/or alcohol use and their social support, due to factors such as shared experiences and role modelling (Barker and Maguire, 2017). ● Peer workers play a key role in improving other outcomes such as engagement, social outcomes, hope, control, self-esteem and confidence and how it can reduce the sense of there being a “them and us” divide between those delivering the service and those using it, encouraging engagement (Repper and Carter, 2011). ● Edited book provides an account of the development of peer support approaches and an overview of their current uses and applications (Watson and Meddings (eds), 2019). |

Appendix D: Evidence review protocol and bibliography

Below we set out our strategy and parameters for reviewing evidence to support the feasibility study. It includes proposed research questions, search and selection strategy. Our approach was informed by the need to complete the evidence review within a tight timeframe.

Research questions

We broke down the overarching study research questions to identify sub-questions to guide our evidence review.

What real-world outcomes should the Programme seek to influence for clients who engage with it?

- What outcomes are important for people experiencing multiple disadvantage? What are the key features of recovery?
- What outcomes have been monitored and/or achieved by other similar programmes and initiatives? What outcomes are realistic?

What data is available which can be used as metrics to measure the effect of the programme on real-world outcomes?

- How have other evaluations measured outcomes?
- What administrative data is available that could evidence individual, service and system level outcomes? What can be learnt from other evaluations that have made use of administrative data?
- What standardised tools/metrics exist to measure outcomes at the individual and service level (including indicators of recovery)? How appropriate are they for this target group?

What evaluative approaches can be taken to measure systemic change?

- How can the concept of systemic change be operationalised in relation to this client group?
- What methods can be used to evaluate systemic change? What tools are available to support this?
- How have other evaluations of similar programmes measured systemic change? What can we learn from these evaluations?

What other data will need to be collected to support the evaluation and the programme, and what data systems will need to be in place in Local Authorities to facilitate this?

- What methods can be used to measure fiscal impact? To what extent have these approaches been used with this client group? What can be learnt from other similar evaluations?

What challenges exist in designing an impact evaluation with this client group, in particular with consideration to selecting a treatment and comparison group, and how can these challenges be overcome?

- What methods can be used to attribute impact? To what extent have these been used with this target group? What are the pros and cons of different methods?
- What is the learning from evaluations of similar interventions on effective impact evaluation with this client group?
- How can evaluations be effectively co-produced with people with lived experience? What are the resource implications of this?

Search and review strategy

We identified an initial bibliography of sources to review. Using this as our starting point we snowballed sources by reviewing references for additional relevant material. We augmented this with recommendations from stakeholder interviewees. We also undertook a search of academic literature (Google Scholar) and grey literature (general Google search) to identify supplementary material.

We used search strings formed of a combination of synonymous terms for multiple disadvantage, specific aspects of multiple disadvantage and evaluation. By including terms for specific aspects of multiple disadvantage (e.g. substance misuse) we reduce irrelevant results; terms such as complex needs are used in other contexts in relation to different multi-morbidities. We set out search terms in the table below – we will select one term from each column (e.g. multiple disadvantage AND homelessness AND evaluation).

| Multiple disadvantage | Aspects of multiple disadvantage | Evaluation / research terms |
|----------------------------------|----------------------------------|-----------------------------|
| Multiple disadvantage | Homelessness | Evaluation |
| Multiple and complex needs | Substance misuse | Outcomes |
| Severe and multiple disadvantage | Offending/Criminal justice | Outcome measures/metrics |
| Chronic exclusion | Domestic abuse | Impact evaluation |
| | [Severe] mental health/illness | Economic/Fiscal evaluation |
| | | Cost benefit analysis |

Search results were scanned by title and abstracts of promising results screened. Only results that were written in English and published within the last 15 years (2005-2020) were screened. We prioritised articles related to the UK context, but also included relevant articles from the United States, Canada, Australia or western Europe; articles from other contexts were excluded. We reviewed abstracts to select a sub-set of articles/reports that were relevant to our research questions. Selected articles were reviewed in more detail and mapped to the above research questions. Given time constraints, we did not always read articles/reports in full but focused on sections that related directly to our research questions (such as descriptions/assessments of methods and outcomes). As the review progressed, we also purposively searched for and select articles that helped to fill gaps in the review e.g. articles on promising research methods.

Bibliography

Abercrombie, R. Harries, E. and Wharton, R. (2015) *Systems Change: A guide to what it is and how to do it*. Lankelly Chase/NPC

Advanced Recovery Management System (ARMS). (no date). Available at <http://www.recoveryoutcomes.com> [Accessed 17th March 2021]

Allcock, A. and Smith, L. (2018) *Impact of women specific interventions* Institute for Research and Innovation in Social Sciences (IRISS)

Anthony, W.A. (1993) Recovery from mental illness: The guiding vision of the mental health system in the 1990s. *Psychosocial Rehabilitation Journal*, 16, 11- 23

Atkinson, M. Jones, M. and Lamont, E. (2007) Multi-agency working and its implications: A review of the literature CfBT Education Trust

Aubry T., Tsemberis S., Adair C.E., Veldhuizen S., Streiner D., Latimer E., Sareen J., Patterson M., McGarvey K., Kopp B., Hume C. and Goering P. (2015) One-year outcomes of a randomized controlled trial of housing first with ACT in five Canadian cities. *Psychiatric Services* 66(5), 463-9

Barker S. and Maguire N. (2017). Experts by Experience: Peer Support and its Use with the Homeless. *Community Mental Health Journal*, 53, 598-612

Bassuk, E. L., Unick, G. J., Paquette, K., and Richard, M. K. (2017) Developing an instrument to measure organizational trauma-informed care in human services: The TICOMETER. *Psychology of Violence*, 7(1), 150-157

Battrick, T., Crook, L., Edwards, K. and Moselle, B. (2014) Evaluation of the MEAM pilots – Update on our findings. A report by FTI Consulting and Compass Lexecon for Making Every Adult Matter (MEAM). MEAM

Baxter, S., Johnson, M., Chambers, D., Sutton, A., Goyder, E. and Booth, A. (2018). The effects of integrated care: a systematic review of UK and international evidence. *BMC Health Services Research*, 18, 350

Befani, B. and Mayne, J. (2014) Process tracing and contribution analysis: A combined approach to generative causal inference for impact evaluation. *IDS bulletin*, 45(6), 17-36

- Befani, B. and Stedman-Bryce, G. (2017) Process Tracing and Bayesian Updating for impact evaluation. *Evaluation* 23(1), 42-60
- Bell, R. (2014) A multi-agency evaluation of the Leeds Dual Diagnosis care co-ordination protocol, *Advances in Dual Diagnosis*, 7(4), 162-184
- Bennett, L. (2020) St Mungo's Recovery Approach: Rapid evidence review St Mungo's
- Best D (2019) A model for resettlement based on the principles of desistance and recovery. *HM Inspectorate of Probation Academic Insights*
- Best, D., Albertson, K., Irving, J., Lightowlers, C., Mama-Rudd, A. and Chaggar, A. (2015) *The UK Life in Recovery Survey 2015: the first national UK survey of addiction recovery experiences*. Sheffield Hallam University
- Best, D., Bamber, S., Battersby, A., Gilman, M., Groshkova, T., Honor, S., and White, W. (2010) Recovery and straw men: An analysis of the objections raised to the transition to a recovery model in UK addiction services, *Journal of Groups in Addiction and Recovery*, 5, 264-288
- Best D., De Alwis S.J. and Burdett D. (2017) The recovery movement and its implications for policy, commissioning and practice. *Nordic Studies on Alcohol and Drugs*, 34(2)
- Best D., Edwards M., Cano I., Durrance J., Lehman J. and White W. (2018) Strengths planning for building recovery capital. *Counsellor Magazine*
- Best, D. and Laudet, A. (2010) *The potential of recovery capital*. Royal Society of Arts, London
- Best, D., and Lubman, D. (2017) Friends matter but so does their substance use: The impact of social networks on substance use, offending and wellbeing among young people attending specialist alcohol and drug treatment services. *Drugs: Education, Prevention and Policy*, 24(1), 111-117
- Better Evaluation (no date) *Realist Evaluation*. Available at: https://www.betterevaluation.org/en/approach/realist_evaluation [Accessed 17th March 2021]
- Bicket, M., Christie, I., Gilbert, N., Hills, D., Penn, A. and Wilkinson, H. (2020) *Magenta Book 2020 Supplementary Guide: Handling Complexity in Policy Evaluation*. London: HM Treasury
- Biggs, J.S., Farrell, L., Lawrence, G. and Johnson, J.K. (2014) A practical example of contribution analysis to a public health intervention. *Evaluation*, 20(2), 214-229
- Bouttell J., Craig P., Lewsey J., Robinson, M. and Popham, F. (2018) Synthetic control methodology as a tool for evaluating population-level health interventions *Journal of Epidemiology and Community Health*, 72, 673-678
- Bramley, G. and Fitzpatrick, S (2015) *Hard Edges: Mapping severe and multiple disadvantage* Lankelly Chase

- Breedvelt, J. (2016). *Psychologically Informed Environments: A Literature Review*. London: Mental Health Foundation
- Bretherton, J. and Pleace, N. (2015) *Housing First in England: An Evaluation of Nine Services* Centre for Housing Policy
- Broadbridge, A. (2018) *What makes an effective multiple and complex needs worker?* Fulfilling Lives Newcastle and Gateshead
- Buckingham, S. A., Frings, D., and Albery, I. P. (2013) Group membership and social identity in addiction recovery. *Psychology of Addictive Behaviors*, 27(4), 1132
- Burchett, H., Kneale, D., Blanchard, L. and Thomas, J. (2020) When assessing generalisability, focusing on differences in population or setting alone is insufficient. *Trials* 21, 286
- Byrne, D. (2013) Evaluating complex social interventions in a complex world *Evaluation* 19(3), 217-228
- Byrne, D. (2016). *Qualitative Comparative Analysis: a pragmatic method for evaluating intervention*. CECAN Evaluation and Policy Practice Note (EPPN) No. 1 for policy analysts and evaluators
- Cairns, A.J., Kavanagh, D.J., Dark, F., and McPhail, S.M. (2019) Goal setting improves retention in youth mental health: a cross-sectional analysis. *Child and Adolescent Psychiatry and Mental Health*, 13, 31
- Campbell, S. and Harper, G. (2012) Quality in policy impact evaluation: understanding the effects of policy from other influences (supplementary Magenta Book guidance) HM Treasury / DECC / DEFRA
- Cano I., Best D., Edwards M. and Lehman J. (2017) Recovery capital pathways: modelling the components of recovery wellbeing. *Drug and alcohol dependence*, 181, 11-19
- Cattell, J. and Mackie, A. (2011) Simple but effective: Local solutions for adults facing multiple deprivation. Adults facing Chronic Exclusion evaluation – Final report CLG
- Centre for the Evaluation of Complexity Across the Nexus (CECAN). (2020) Choosing Appropriate Evaluation Methods – A Tool for Assessment and Selection (Version Two)
- Centre for Regional Economic and Social Research (2016) *Navigator practice in the WY-FI project – briefing paper WY-FI*
- Centre for Social Justice (2017) *Housing First: Housing-led solutions to rough sleeping and homelessness* The Centre for Social Justice
- CFE Research and University of Sheffield (2020) Improving access to mental health support for people experiencing multiple disadvantage. CFE Research
- Cheminais, R. (2009). *Effective Multi-Agency Partnerships. Putting Every Child Matters into Practice*. London: Sage

- Chou, K. L., Liang, K., and Sareen, J. (2011). The association between social isolation and DSM-IV mood, anxiety, and substance use disorders: wave 2 of the National Epidemiologic Survey on Alcohol and Related Conditions. *The Journal of clinical psychiatry*, 72(11), 1468-1476
- Cloud, W. and Granfield, R. (2008). Conceptualizing recovery capital: Expansion of a theoretical construct. *Substance use & misuse*, 43(12-13), 1971-1986
- Cockersell P. (2016). PIEs five years on. *Mental health and social inclusion*, 20 (4)
- Cordis Bright (2020) MEAM Approach evaluation: year 3 report. MEAM
- Cordis Bright (2019) MEAM Approach evaluation: year 2 report. MEAM
- Cordis Bright (2018) Integrated health and social care evidence reviews. What are the key factors for successful multidisciplinary team working?
- Cream, J., Fenney, D., Williams, E., Baylis, A., Siham, D. and Wyatt, H. (2020) *Delivering health and care for people who sleep rough* London: The King's Fund
- Crisp, R., Fletcher, D., Parr, S. and Wilson, I. (2020). West-Yorkshire Finding Independence (WY-FI): Effectiveness, outcomes and impact. Final evaluation report 2020. Sheffield Hallam CRESR
- Davidson L., Andres-Hyman R., Bedregal L., Tondora J., Frey J. and Kirk T.A. (2008) From “double trouble” to “dual recovery”: Integrating models of recovery in addiction and mental health. *Journal of Dual Diagnosis*, 4(3), 273-290
- Davidson Knight, A., Lowe, T., Brossard, M. and Wilson, J. (2017) *A Whole New World: Funding and Commissioning in Complexity Collaborate*
- Del Vecchio, P. (2012) *SAMHSA's working definition of recovery updated*. *Substance Abuse and Mental Health Services Administration (SAMHSA)* Available at: <https://www.samhsa.gov/recovery> [Accessed 17th March 2021]
- Denzin, N. and Lincoln, Y. (2017) *The Sage Handbook of Qualitative Research* 5th edition London: Sage
- Diamond, A. Adamson, J. Moreton, R. Robinson, S, Spong, S. Howe, P. Bysshe, S. Sheikh-Latif, N. and Citarella, V. (2013) *Multiple and complex needs: a rapid evidence assessment* CFE Research
- Dickens, G., Weleminsky, J., Onifade, Y. and Sugarman, P. (2012) Recovery star: validating user recovery, *The Psychiatrist*, 36, 45-50
- Dingle, G. A., Cruwys, T., and Frings, D. (2015). Social identities as pathways into and out of addiction. *Frontiers in psychology*, 6, 1795
- Egan, M., McGill, E., Penney, T., Anderson de Cuevas, R., Er, V., Orton, L., Lock, K., Popay, J., Savona, N., Cummins, S., Rutter, H., Whitehead, M., De Vocht, F., White, M., Smith, R., Andreeva, M., Meier, P., Marks, D. and Petticrew, M. (2019a) *NIHR SPHR Guidance on Systems Approaches to Local Public Health Evaluation. Part 1: Introducing*

systems thinking. London: National Institute for Health Research School for Public Health Research

Egan, M., McGill, E., Penney, T., Anderson de Cuevas, R., Er, V., Orton, L., Lock, K., Popay, J., Savona, N., Cummins, S., Rutter, H., Whitehead, M., De Vocht, F., White, M., Smith, R., Andreeva, M., Meier, P., Marks, D. and Petticrew, M. (2019b) *NIHR SPHR Guidance on Systems Approaches to Local Public Health Evaluation. Part 2: What to consider when planning a systems evaluation*. London: National Institute for Health Research School for Public Health Research

EIF (2020) *Maturity Matrix Summary*. Early Intervention Foundation

EuroQol Research Foundation (2018) *EQ-5D-3L User Guide, 2018*. Available at: <https://euroqol.org/publications/user-guides> [Accessed 17th March 2021]

Evans, C., Mellor-Clark, J., Margison, F., Barkham, M., Audin, K., Connell, J. and McGrath, G. (2009) CORE: Clinical Outcomes in Routine Evaluation, *Journal of Mental Health*, 9(3), 247-255

Fantuzzi, C. and Mezzina, R. (2020). Dual diagnosis: A systematic review of the organization of community health services. *The International journal of social psychiatry* 66(3), 300-310

Fat, L., Scholes, S., Boniface, S., Mindell, J., & Stewart-Brown, S. (2017). Evaluating and establishing national norms for mental wellbeing using the short Warwick-Edinburgh Mental Well-being Scale (SWEMWBS): findings from the Health Survey for England. *Quality of Life Research*, 26(5), 1129–1144

Felitti, V. J. and Anda, R. F. (2010) The relationship of adverse childhood experiences to adult medical disease, psychiatric disorders and sexual behavior: Implications for healthcare. In R. Lanius, E. Vermetten and C. Pain (Eds.), *The impact of early life trauma on health and disease: The hidden epidemic*. Cambridge, UK: Cambridge University Press

Field, H., Hudson, B., Hewett, H. and Khan, Z. (2019) Secondary care usage and characteristics of hospital inpatients referred to a UK homeless health team: a retrospective service evaluation *BMC Health Services Research*, 19, Article number 857

Flanagan, S., Damery, S. and Combes, G. (2017). The effectiveness of integrated care interventions in improving patient quality of life (QoL) for patients with chronic conditions. An overview of the systematic review evidence. *Health and Quality of Life Outcomes* 15(1), 188

Flinders M., Wood M. and Cunningham M. (2016) The politics of co-production: risks, limits and pollution. *Evidence and Policy: A Journal of Research, Debate and Practice*, 12(2), 261–79

Friel, S., Murphy, H., Klenk, H. and Vaid, L. (2020) *Tackling Multiple Disadvantage – Final Evaluation Report* Learning and Work Institute

Gagliardi A., Kothari A., Graham I., Gagliardi A.R., Kothari A. and Graham I.D. (2017) Research agenda for integrated knowledge translation (IKT) in healthcare: What we know and do not yet know. *Journal of Epidemiology and Community Health*, 71(2), 105–6

- Green J. and South J. (2006) *Evaluation*. Maidenhead: Open University Press
- Goll, J. C., Charlesworth, G., Scior, K. and Stott, J. (2015). Barriers to social participation among lonely older adults: the influence of social fears and identity. *PloS one*, 10(2)
- Granfield, R. and Cloud, W. (1999). *Coming clean: Overcoming addiction without treatment*. NYU Press
- Guijt I. (2014) Participatory Approaches. *Methodological Briefs: Impact Evaluation 5*, Florence: UNICEF Office of Research
- Handran, J. (2015). Trauma-informed systems of care: The role of organizational culture in the development of burnout, secondary traumatic stress, and compassion satisfaction. *Journal of Social Welfare and Human Rights*, 3(2), 1-22
- Hayes, M., Zonneville, M., and Bassuk, E. (2013) *The SHIFT Study final report: Service and housing interventions for families in transition*. Newton, MA: National Center on Family Homelessness
- Hennessy, E. (2017). Recovery Capital: a systematic review of the literature. *Addiction Research & Theory*. 25(5), 349-360
- HM Treasury (2020a) Magenta Book: Central Government guidance on evaluation. London: HM Treasury
- HM Treasury (2020b) The Green Book: Central Government guidance on appraisal and evaluation. London: HM Treasury
- HM Treasury, Public Service Transformation Network, New Economy (2014) Supporting public service transformation: cost benefit analysis guidance for local partnerships. London: HM Treasury
- Homeless Link (2020) *Eligibility and referrals in Housing First, June 2020* Available at https://hfe.homeless.org.uk/sites/default/files/attachments/Eligibility%20and%20referrals%20briefing_2.pdf#:~:text=Eligibility%20and%20referrals%20in%20Housing%20First%20About%20This,what%20to%20consider%20beyond%20the%20point%20of%20referral [Accessed 1st April 2021]
- Horvath, A.O. (1981, 1984). Working Alliance Inventory. Available at: <https://wai.profhorvath.com/> [Accessed 1st April 2021]
- Hough, J. (2014) Changing systems for people with multiple needs: Learning from the literature. Summary report New Economics Foundation
- INVOLVE (no date) What is public involvement in research? Available at <https://www.invo.org.uk/find-out-more/what-is-public-involvement-in-research-2/> [Accessed 17th March 2021]
- Ipsos Mori (2019) Liverpool Waves of Hope evaluation: Final report. Ipsos Mori

Isaac, B., Bolden, R., Pawson, C. Gulati, A., Gasper, R., Plumridge, A., Stern, E. and Tischner, I. (2017) *Building Connections: Golden key local evaluation phase 2 report*. University of the West of England

Jetten, J., Haslam, C. and Alexander, S. H. (Eds.). (2012). *The social cure: Identity, health and well-being*. Psychology press

Johnson, G. and Pleace, N. (2016) How do we measure success in homelessness services? Critically assessing the rise of the homelessness outcomes star. *European Journal of Homelessness*, 10(1)

Johnson R. and Haigh R. (2010). Social psychiatry and social policy for the 21st century: new concepts for new needs: the 'psychologically-informed environment'. *Mental Health and Social Inclusion*, 14(4), 30-35

Johnstone, M., Jetten, J., Dingle, G.A., Parsell, C. and Walter, Z.C. (2015) Discrimination and well-being amongst the homeless: the role of multiple group membership. *Frontiers in psychology*, 6, 739

Kay, A. and Baker, P. (2015) What can causal process tracing offer to policy studies? A review of the literature. *Policy Studies Journal*, 43(1), 1-21

Keane, M., Coyle, C., Farragher, L., O'Nolan, G., Cannon, A. and Long, J. (2018). Integrated interventions to reduce pressure on acute hospitals. Evidence review. Dublin: Health Research Board

Kielburger, L. C., Rockman, P. and Baranek, P. (2018). Preventing burnout among service providers. In Kidd, S., Slesnick, N., Frederick, T., Karabanow, J. and Gaetz, S. (eds.) *Mental Health & Addictions Interventions for Youth Experiencing Homelessness: Practical Strategies for Front-line Providers*. Canada: Canadian Observatory on Homelessness Press, 171-187

Killaspy, H., White, S. and King, M. (2012) Psychometric properties of the mental health recovery star, *British Journal of Psychiatry*, 201, 65-70

Kiresuk, T.J. and Sherman, R.E. (1968) Goal Attainment Scaling: A General Method for Evaluating Comprehensive Community Mental Health Programs, *Community Mental Health*, December 1968

Korcha, R. A., Polcin, D. L., Bond, J. C., Lapp, W. M. and Galloway, G. (2011) Substance use and motivation: a longitudinal perspective. *The American Journal of Drug and Alcohol Abuse*, 37(1), 48-53

Lamb, H., Moreton, R., Leonardi, S., Welford, J., O'Donnell, J. and Howe, P. (2019c) *What makes a difference: method notes* CFE Research

Lamb, H., Moreton, R., Welford, J., Leonardi, S., O'Donnell, J. and Howe, P. (2019a) *What has Fulfilling Lives achieved* CFE Research

Lamb, H., Moreton, R., Welford, J., Leonardi, S. and Howe, P. (2019b) *Why we need to invest in multiple needs* CFE Research

Lamb, H., Moreton, R., Welford, J., Leonardi, S., O'Donnell, J. and Howe, P. (2019d) *What makes a difference* CFE Research

Leamy, M., Bird, V. J., Le Boutillier, L., Williams, J. and Slade, M. (2011) Conceptual framework for personal recovery in mental health: systematic review and narrative synthesis. *The British Journal of Psychiatry*, 199, 445-452

Liddle, M. in collaboration with Disley, E., Maguire, M., Meek, R. and Renshaw, J. (2019). *Intermediate Outcomes Measurement Instrument (IOMI) toolkit. Guidance notes*. Ministry of Justice Analytical Series 2019

Litt, M. D., Kadden, R. M., Kabela-Cormier, E. and Petry, N. M. (2009). Changing network support for drinking: network support project 2-year follow-up. *Journal of consulting and clinical psychology*, 77(2), 229

Longabaugh, R., Wirtz, P. W., Zywiak, W. H. and O'Malley, S. S. (2010). Network support as a prognostic indicator of drinking outcomes: The COMBINE study. *Journal of Studies on Alcohol and Drugs*, 71(6), 837-846

Maguire, N. Grellier, B. and Clayton, K. (2017) The impact of CBT training and supervision on burnout, confidence and negative beliefs in a staff group working with homeless people. *Behavioural and Cognitive Psychotherapy*

Mackeith, J. (2014) Assessing the reliability of the Outcomes Star in research and practice *Housing, Care and Support*, 17(4), 188-197

Maslach, C. Jackson, S. E. Leiter, M. P. (1986) *Maslach Burnout Inventory* Palo Alto, Cal.: Consulting Psychologists Press

Mayne, J. (2008). Contribution analysis: An approach to exploring cause and effect. ILAC Brief 16

McCarthy, L. Parr, S. Green, S. Reeve, K. (2020) *Understanding models of support for people facing multiple disadvantage: A literature review* Fulfilling Lives Lambeth, Southwark, Lewisham

McNeish, D., Scott, S., Sosenko, F., Johnsen, S. and Bramley, G. (2016). *Women and girls facing severe and multiple disadvantage*. Lankelly Chase

MHCLG (2019) National Evaluation of the Troubled Families Programme 2015- 2020: Findings Evaluation overview policy report. MHCLG

Moore, G.F., Evans, R.E., Hawkins, J., Littlecott, H., Melendez-Torres, G.J., Bonell, C. and Murphy, S. (2019) From complex social interventions to interventions in complex social

systems: Future directions and unresolved questions for intervention development and evaluation. *Evaluation*, 25(1), 23-45

Moreton, R., Hansel, M., Sandhu, J. and Welford, J. (2021b) *What makes an effective multiple disadvantage navigator?* CFE Research

Moreton, R., Welford, J., Collinson, B. Milner, C. Sandhu, J. and Hansel, M. (2021a) *Summary: Workforce development and multiple disadvantage.* CFE Research

Moreton, R., Welford, J., Mulla, I. and Robinson, S. (2018). *Promising practice Key findings from local evaluations to date.* CFE Research

Mouser, A. and Bowers, A. (2017) *Systems Changers: How frontline workers can contribute to and create systems change.* Lankelly Chase

Murphy, H., Friel, S., Dhillon, C. and Vaid, L. (2017) *Tackling Multiple Disadvantage: Year 1 Interim Report, Crisis*

NEF Consulting (2013) *Social CBA and SROI, Economics in policy-making 4.* NEF

Nugter, M. A., Engelsbel, F., Baehler, M., Keet, R. and van Veldhuizen, R. (2016). Outcomes of FLEXIBLE assertive community treatment (FACT) implementation: a prospective real life study, *Community Mental Health Journal*, 52(8), 898-907

Nutley, S., Powell, A., and Davies, H. (2013) *What counts as good evidence* Alliance for Useful Evidence

Oliver K., Kothari A. and Mays N. (2019) The dark side of coproduction: do the costs outweigh the benefits for health research? *Health Research Policy and Systems* 17, 33

Olivet, J., McGraw, S., Grandin, M. and Bassuk, E. (2010). Staffing challenges and strategies for organizations serving individuals who have experienced chronic homelessness. *The Journal of Behavioral Health Services & Research*, 37(2), 226-238

Onken, S.J., Craig, C.M., Ridgway, P., Ralph, R.O. and Cook, J.A. (2007) An analysis of the definitions and elements of recovery: a review of the literature. *Psychiatric Rehabilitation Journal*, 31, 9-22

Pawson, R. and Tilley, N. (1997). *Realistic Evaluation.* London: Sage

PIE link Net (2021) *Introducing: The Pizazz.* Available at: <http://pielink.net/pizazz/> [Accessed 17th March 2021]

Pitt V., Lowe, D., Hill, S., Prictor, M., Hetrick, S.E., Ryan, R. and Berends, L. (2013). Consumer-providers of care for adult clients of statutory mental health services. *Cochrane Database of Systematic Reviews* 2013 (3)

Priebe, S., Huxley, P., Knight, S. and Evans, S. (1999). Application and Results of the Manchester Short Assessment of Quality of Life (Mansa). *International Journal of Psychiatry*, 45(1), 7-12

- Punton, M. and Welle, K. (2015) Straws-in-the-wind, Hoops and Smoking Guns: What can Process Tracing Offer to Impact Evaluation? CDF Practice Paper 10, Brighton: IDS
- Repper, J. and Carter, T. (2011). A review of the literature on peer support in mental health services. *Journal of Mental Health*, 20(4), 392-411
- Revolving Doors Agency (2019) Highlights from the Evaluation of Birmingham Changing Futures Together. Revolving Doors Agency
- Ritchie, C. (2015) Creating a psychologically informed environment Westminster City Council
- Rinaldi, M., Linnell, D. and Clenaghan, M. (2008) *The New Directions Team Assessment (Chaos Index)* Available at: <http://www.meam.org.uk/wp-content/uploads/2010/05/NDT-Assessment-process-summary-April-2008.pdf> [Accessed 17th March 2021]
- Rogers, Patricia J. (2008). Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions *Evaluation* 2008, 14, 29
- Ruiu, M.L. (2016). The social capital of cohousing communities. *Sociology*, 50(2), 400-415
- Salter, K.L. and Kothari, A. (2014). Using realist evaluation to open the black box of knowledge translation: a state-of-the-art review. *Implementation Science* 9, 115.
- Sandu, R.D., Anyan, F. and Stergiopoulos V. (2021) Housing first, connection second: the impact of professional helping relationships on the trajectories of housing stability for people facing severe and multiple disadvantage. *BMC Public Health* 21, Article no. 249
- Shepherd, G., Boardman, J. and Slade, M. (2008). *Making Recovery a Reality*. Sainsbury Centre for Mental Health
- Simoneau, H. and Bergeron, J. (2003). Factors affecting motivation during the first six weeks of treatment. *Addictive behaviors*, 28(7), 1219-1241
- Social Impacts Task Force (no date) Adding subjective wellbeing to evaluations – Why and How? SITF
- Sood, L. and Owen, A. (2014). A 10-year service evaluation of an assertive community treatment team: trends in hospital bed use. *Journal of Mental Health*, 23(6), 323-327
- Sosenko, F., Bramley, G. and Johnsen, S. (2020) Gender Matters: Gendered patterns of severe and multiple disadvantage in England. Lankelly Chase
- Tazyman, S. and Bowes, L. (2019) *Implementing Randomised Control Trials to evaluate the impact of outreach activity – lessons learned* CFE Research [blog] Available at: <http://cfe.org.uk/2018/11/28/implementing-randomised-controlled-trials-to-evaluate-the-impact-of-outreach-activity-lessons-learned/> [Accessed 17th March 2021]
- Terry, L. and Cardwell, V. (2016) Refreshing perspectives: Exploring the application of peer research with populations facing severe and multiple disadvantage. Revolving Doors Agency, Lankelly Chase

Triangle (2012) Recovery Star, The Outcomes Star for mental health and well-being. Available at: <https://www.outcomesstar.org.uk/using-the-star/see-the-stars/recovery-star-4/> [Accessed 17th March 2021]

Tsemberis, S., McHugo, G., Williams, V., Hanrahan, P. and Stefancic, A. (2007) Measuring Homelessness and Residential Stability: The residential time-line follow-back inventory *Journal of Community Psychology*, 35(1), 29-42

Turley, C., Payne, C. and Webster, S. (2013). Enabling features of Psychologically Informed Planned Environments. Ministry of Justice Analytical Series 2013. London: National Offender Management Service

UK Drug Policy Commission. (2012) A fresh approach to drugs: the final report of the drug policy commission. UK Drug Policy Commission

Vaughn L.M., Whetstone C. and Busch M.D. (2018) Partnering with insiders: A review of peer models across community-engaged research, education and social care. *Health and Social Care in the Community*

Waegemakers Schiff, J. and Lane, A. (2016). *Burnout and PTSD in Workers in the Homeless Sector in Calgary*. Calgary Homeless Foundation

Warren, J., Wistow, J. and Bamba, C. (2014). Applying qualitative comparative analysis (QCA) in public health: a case study of a health improvement service for long-term incapacity benefit recipients. *Journal of Public Health*, 36(1), 126-133

Warwick Medical School (2020) *The Warwick-Edinburgh Mental Wellbeing Scales – WEMWBS*. Available at: <https://warwick.ac.uk/fac/sci/med/research/platform/wemwbs/> [Accessed 17th March 2021]

Watson E. and Meddings S. (editors). (2019). *Peer Support in Mental Health*. London: Red Globe Press

Waugh, A., Clarke, A., Knowles, J. and Rowley, D. (2018) *Health and Homelessness in Scotland* The Scottish Government

White, W. and Cloud, W. (2008) Recovery capital: A primer for addictions professionals. *Counselor*, 9(5), 22-27

Zukoski A. and Luluquisen M. (2002) Participatory Evaluation What is it? Why do it? What are the challenges? *Community based public health: Policy and practice*, 5

Zywiak, W.H., Longabaugh, R. and Wirtz, P.W. (2002). Decomposing the relationships between pre-treatment social network characteristics and alcohol treatment outcome. *Journal of studies on alcohol*, 63(1), 114-121