



Department for Levelling Up,  
Housing & Communities

# Evaluation of the Changing Futures programme: feasibility study

Annex to the main 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](http://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

<b>1</b>	<b>Cohort sizes and sample sizes</b>	<b>2</b>
1.1	Target cohorts and anticipated cohort sizes	2
1.2	Possible sample sizes for the evaluation	3
1.3	Required sample sizes	6
<b>2</b>	<b>Additional reflections on external comparator groups</b>	<b>9</b>
2.1	Local area roll-out and opportunities for comparator groups	9
2.2	Proposed cross-government and cross-project collaboration to identify external comparator group	9

# 1 Cohort sizes and sample sizes

## 1.1 Target cohorts and anticipated cohort sizes

This section includes reflections on information gathered by MHCLG in conversations with 21 shortlisted areas in relation to their planned target cohorts and roll-out plans. It considers the implications of this information for the evaluation design and implementation.

### Variations in specific make-up of cohort

The people engaged in target cohorts are expected to meet the eligibility criteria set out in the prospectus of experiencing three or more of the five issues included in MHCLG's definition of multiple disadvantage. Some areas have also identified a particular cohort on whom they may want to have an additional focus, including women, BAME people and young adults (18-25). The inclusion of more specific characteristics for the target cohort may affect the comparability of activities and outcomes across areas, but may also provide an additional opportunity to examine differential outcomes for different client groups.

### Anticipated cohort size

Cohort size is likely to vary substantially between areas and differences were linked to the geographical size and population of areas, as well as the ways in which they are defining and selecting their cohort. Areas estimated cohorts of between 50 and over 1,000 people, though most estimates fall between 100 and 600 people<sup>1</sup>.

In the Fulfilling Lives programme, 3,480 clients were engaged in the 12 Fulfilling Lives areas in the first four years of delivery (May 2014 – September 2018)<sup>2</sup>. This is a mean of 290 clients per area though total cohort size varies quite substantially between areas, from 115 in Camden and Islington to 828 in West Yorkshire<sup>3</sup>. These are also cumulative cohorts recruited over a longer period. The largest number of new beneficiaries recruited across the programme in a year was 1,068 from Oct 2014 to Sept 2015, a mean of 89 per area. All but one partnership has engaged fewer direct beneficiaries than originally planned – often substantially fewer.

This suggests that the upper estimates of cohort sizes for Changing Futures may be ambitious, unless areas have already engaged the people in these cohorts, have a particular large population experiencing multiple disadvantage from which to draw them or have more resources to identify and engage people than were available in Fulfilling Lives areas. It may therefore be safer to assume a smaller anticipated cohort size, such as a mean of 175 people per area, or **2,625 people in total across the 15 areas**.

---

<sup>1</sup> Some areas also did not know their expected cohort size.

<sup>2</sup> Lamb, H., Moreton, R., Leonardi, S., Welford, J., O'Donnell, J. and Howe, P. (2019). *What has Fulfilling Lives achieved: method notes*. CFE Research.

<sup>3</sup> This is partly due to the geographic size of area but also the model of support and funding requested; some areas always planned to work intensively with a smaller cohort.

## Timing of recruitment to the cohort

Timing of recruitment to the cohort is likely to be affected by three key issues:

- **How quickly areas can mobilise.** Areas which need to commission and/or set up new interventions and services before engaging clients will need a lead-in time from the point at which funding for 2021-22 is confirmed. In our experience of previous similar programmes this often takes 9-12 months. However, this may be reduced significantly if Changing Futures work builds on existing roles, service structures and commissioning structures. Initial preparation that does not require funding may also be possible in the time since areas submitted their bid and were shortlisted. This may further reduce set-up time.
- **How quickly areas can engage people in the programme.** The extent to which individuals are already known to those designing and delivering local interventions will be important in determining the speed of their engagement in the programme. In the majority of areas, it was not clear exactly how the cohort was to be identified. Where information on this was provided, it was suggested that people would be identified via existing data; through open referrals across a wide range of organisations (criminal justice, voluntary organisations, health and social care) and even self-referrals in one case; and/or through existing multi-agency teams.
- **Whether areas have a maximum caseload or cohort size for receiving support at any one time.** Areas may need to cap the total cohort size for those receiving support at any one time. Assuming that they are successful in engaging clients and supporting them to maintain their engagement, they may therefore reach their maximum cohort size before the end of the two-year period. This may result in waiting lists or alternative support to manage demand.

As these elements are unknown and are likely to vary across local areas, it is difficult to confidently estimate recruitment rates to the programme.

## 1.2 Possible sample sizes for the evaluation

There are a number of issues which affect the likely sample size which will result from the target cohort sizes. These include:

- **How the cohort is being defined.** For example, in areas where the target cohort is being defined as those people in receipt of a specific service or intervention it is theoretically possible to gain consent to collate and share data in relation to all the people in the cohort. If areas are taking a more place-based approach to cohort definition (e.g. considering all people who are homeless or street attached in a particular geographical area to be within the cohort) this may reduce the proportion of people from whom consent to share data can be gained, since some people might not be engaging with any interventions. Both of these approaches to cohort definition have

been taken in areas involved in the MEAM Approach network, though the first approach was more common.

- **The rate of recruitment to the cohort.** For clients to be included in the sample, ideally they would start receiving interventions or support by the end of 2021-22, to allow at least a 12-month data collection period following the start of support. As discussed in section 1.1.3, the rate of recruitment is difficult to determine in advance. Assuming a steady recruitment rate to the cohort, we could estimate that approximately **1,300 clients** will be recruited by the end of 2021-22 and theoretically eligible for inclusion in the sample.
- **Success in gaining consent.** The proportion of clients for whom consent to collate and share data is gained will impact on sample size. In the first four years of Fulfilling Lives, consent was gained from 84% of the total of 3,480 clients known to have engaged with the programme across the 12 Fulfilling Lives areas (2,913 of 3,480 clients)<sup>4</sup>. In the first three years of the MEAM Approach evaluation, on the other hand, evaluators received data on 45% of clients supported by 21 network areas (579 of 1,277 clients)<sup>5</sup>. This difference likely results – at least in part – from the fact that greater resource was available in the Fulfilling Lives areas to engage clients in the evaluation. Resources in Changing Futures will be more akin to those in Fulfilling Lives so an optimistic estimate of the proportion of clients who will provide consent may be reasonable. If we assume an 80% success rate in gaining consent, this would result in a total sample of **1,050 clients**.
- **Attrition.** Attrition levels may well be fairly high. For example, in the first four years of Fulfilling Lives 2,913 clients had provided consent for their data to be shared but only 1,665 of these people remained on the programme for 4 quarters and were therefore eligible for inclusion in some of the analyses of change over time outlined in Figure 1<sup>6</sup>.
- **The ability of areas to collate and share data.** There are two key elements to this:
- **Setting up appropriate data sharing agreements and processes.** In our experience of the Fulfilling Lives and MEAM Approach evaluations, this can be time-consuming and success in gaining data from all relevant partners is patchy. This may mean that data for clients who join the cohort early (prior to data sharing arrangements being in place) needs to be collated retrospectively. It will likely also impact on data quality (see below).
- **Having appropriate systems and adequate local resource to collate data.** Providing local areas with a standardised data collection tool can ensure that they have an appropriate system in place to collate data. However, this generally needs to be separate to core local case management and data reporting systems (because these differ across local areas and data cannot always be extracted from them in an

---

<sup>4</sup> Lamb, H., Moreton, R., Leonardi, S., Welford, J., O'Donnell, J. and Howe, P. (2019). *What has Fulfilling Lives achieved: method notes*. CFE Research.

<sup>5</sup> Cordis Bright (2019). *Year 3 evaluation: technical appendix*. MEAM.

<sup>6</sup> Lamb, H., Moreton, R., Leonardi, S., Welford, J., O'Donnell, J. and Howe, P. (2019). *What has Fulfilling Lives achieved: method notes*. CFE Research.

appropriate format for the evaluation)<sup>7</sup>. This often means that data needs to be manually matched and inputted by a local staff member, requiring substantial time. It is therefore particularly important that local areas are supported to establish clear responsibilities and resource for data collection.

- Both of these issues may be mitigated by the central collation of data against any indicators for which data is available centrally (though this obviously presents different resourcing challenges for MHCLG and partners).
- **Data quality.** For the Fulfilling Lives and MEAM Approach evaluations, the eligible samples for different analyses were substantially smaller than the total sample of clients for which evaluators had received data. This was partly as a result of missing data or low data quality, and partly as a result of the specific eligibility criteria for each analysis. It is likely that the Changing Futures evaluation will experience similar data quality challenges. Figure 1 provides examples of eligible sample sizes for the Fulfilling Lives and MEAM Approach evaluations. These are not directly comparable as the evaluations were at different stages and took different analytical approaches. They are provided as an indication of the sample sizes which might be achievable in the Changing Futures evaluation. If Changing Futures achieved similar data quality this might result in eligible samples for different analyses of **between 105 clients and 315 clients**. However, it may be that additional resource and a central government role in mandating data collection can increase these sample sizes.

---

<sup>7</sup> In Fulfilling Lives, Homeless Link issued a version of their Inform tool that aligned with the Fulfilling Lives data collection framework, which was helpful. Even so, data input, checking, cleaning, and extraction are time consuming tasks and most Fulfilling Lives partnerships ultimately employed dedicated staff to support this.

**Figure 1: Eligible sample sizes for most recent published analyses in Fulfilling Lives and MEAM Approach evaluations**

	Fulfilling Lives (after four years) <sup>8</sup>		MEAM (after 3 years) <sup>9</sup>	
Focus of analysis	Number	% of total sample	Number	% of total sample
Total sample	2,913	100%	579	100%
Change over time in Homelessness Outcomes Star scores	829 for baseline to 12 month	28%	158 for baseline to last/most recent quarter of support	12%
	390 baseline to 18 month	13%		
Change over time in NDTA scores	955 for baseline to 12 month	33%	159 for baseline to last/most recent quarter of support	12%
	602 for baseline to 18 month	21%		
Change over time in accommodation status/use	488 for quarter 1 to quarter 4 to quarter 8	17%	244 for first to last/most recent quarter of support	19%
Change over time in service use	323 for quarter 1 to quarter 4 to quarter 8	11%	321 - 349 for first to last/most recent quarter of support for different types of service use	25% - 27% for different types of service use

### 1.3 Required sample sizes

#### **Required sample size for comparing means in the same people before and after receiving a Changing Futures intervention (Level A)**

It is possible to apply a ‘rule of thumb’ approach to estimating the minimum sample size that would be needed to determine whether Changing Futures has had an impact on individual level outcomes, for example personal wellbeing, which might be measured by the longer or shorter form Warwick Edinburgh Mental Wellbeing Scale (WEMWBS). This method relies on a range of assumptions, which would need to be discussed and agreed at the start of the evaluation.

<sup>8</sup> Source: Lamb, H., Moreton, R., Leonardi, S., Welford, J., O'Donnell, J. and Howe, P. (2019). *What has Fulfilling Lives achieved: method notes*. CFE Research.

<sup>9</sup> Source: Cordis Bright (2019). *Year 3 evaluation: technical appendix*. MEAM.



The usual steps in determining what sample size would be needed to see how effective the programme has been, would be<sup>10</sup>:

1. Decide on the smallest difference we want to be able to detect, for example in the case of wellbeing the smallest change in WEMWBS scores at time 1 and time 2.
2. Convert the minimum difference to be detected to standard deviation units by dividing it by the standard deviation. Ascertaining the standard deviation may be possible by referring to previous studies, which might show the range of change that would be expected. In the absence of such information, Conroy (2016) recommends assuming that most values will fall within +2 and -2 standard deviations from the average, giving an approximate range of 4 and standard deviation of (4 divided by 4) = 1.
3. Use the table below (available in Conroy, 2016) to get an estimate of the sample size needed to detect a difference of the desired size if such a difference exists.

**Figure 2: Sample size calculator**

Desired difference (in SD units)	Sample size required at 90% confidence level	Sample size required at 95% confidence level	Percentage of people who will change in the expected direction
2	3	4	98%
1.5	5	6	93%
1.4	6	7	92%
1.3	7	8	90%
1.25	7	9	89%
1.2	8	10	88%
1.1	9	11	86%
1	11	13	84%
0.9	13	17	82%
0.8	17	21	79%
0.75	19	24	77%
0.7	22	27	76%
0.6	30	37	73%
0.5	43	52	69%
0.4	66	82	66%

<sup>10</sup> Conroy, R. (2016). *The RCSI Sample Size Handbook 2016*, Royal College of Surgeons of Ireland

0.3	117	145	62%
0.25	169	208	60%
0.2	263	325	58%

Source: Conroy (2016)

With personal outcomes such as wellbeing or self-esteem it may be difficult to decide what constitutes a significant or acceptable difference, so it might be better to ask for what proportion of those taking part in the programme would we expect or want to see an improvement over time. For example, if the Changing Futures programme decides it would like to see an improvement in (S)WEMWBS scores in at least two thirds of the people receiving an intervention, then scores would have to increase by 0.4 standard deviation units (using the column for 66% in the table above). The minimum sample size needed would be 66 people at the 90% confidence level and 82 at the preferred 95% confidence level. Based on these assumptions, to be able to compare across funded areas, *each area* would need to collect completed and usable measures at both time 1 and time 2 for at least 82 people taking part in the programme. Similarly, if we wanted to be able to compare changes in scores for different groups within the service user population, we would need a larger sample. For example if we wanted to compare changes for women and men at an area level with 95% confidence, we would need a sample of at least 82 women and 82 men from each area.

### **Sample sizes and detecting the impact of Changing Futures on local area population level indicators (Level B)**

The main difficulty with using change in local population level indicators as a means of understanding the impact of Changing Futures is that the number of participants in Changing Futures will be very small relative to the size of the population for whom relevant data, for example hospital admissions, A&E attendances and arrests, is reported. For example, each area may be able to collect data on non-elective hospital admissions for between 120 and 360 programme participants. The mid-2019 average population size of a CCG footprint area was reported to be 416,940 <sup>11</sup>, and the approximate average number of non-elective admissions per CCG footprint area was 36,000 per year in 2019, pre-Covid-19. Similarly, the average number of A&E attendances per acute trust footprint was 188,000 in 2019. In contrast, the average number of people sleeping rough in each of the 83 areas with the highest levels of rough sleeping at the start of the Rough Sleepers Initiative was 43. With such small numbers of people, it was therefore possible to use area-level data to identify change and understand the impact of the programme. Detecting the impact of Changing Futures on most of the area-level indicators relevant to the hoped-for change would not be feasible, unless it were possible to obtain administrative data relating specifically to a sub-population experiencing multiple disadvantage.

---

<sup>11</sup> [Population estimates by output areas, electoral, health and other geographies, England and Wales - Office for National Statistics \(ons.gov.uk\)](https://ons.gov.uk/population/population-estimates-by-output-areas-electoral-health-and-other-geographies-england-and-wales)

## 2 Additional reflections on external comparator groups

### 2.1 Local area roll-out and opportunities for comparator groups

For the most part, areas were not far enough along with their planning and design work to be able to give much information on their proposed roll-out during calls with MHCLG in March 2021. Two areas referenced the existence of waiting lists (although it is unclear what exactly these refer to) and a third area suggested they might establish a waiting list. These might provide organic opportunities for creating comparator groups, but if only three areas are set up in this way this is unlikely to yield enough numbers for a whole programme comparator group, and would also present some issues in that only a proportion of areas would be represented in the comparator group.

In one multi-authority area, delivery at an individual level will happen in less than half of the local authorities. In another area, the work may be focused on a particular part of the local authority. This may also offer opportunities for identifying a comparator group within these areas.

### 2.2 Proposed cross-government and cross-project collaboration to identify external comparator group

Since the initial feasibility study report, there has been a proposal for cross-government and cross-project collaboration to enable the identification of an external comparator group for evaluations of interventions for people with multiple disadvantage. This draws on an approach used in the Housing First evaluation, and the learning from this. It involves qualitative research with people experiencing multiple disadvantage in areas not included in the Changing Futures programme, which would be commissioned by MHCLG and partners, conducted by a contractor and facilitated by services in the selected local authority areas. Individual-level outcomes measurement tools could also be incorporated into this direct research with people experiencing multiple disadvantage.

Our reflections on this proposed approach are:

- The collaboration is intended to enable the identification of an external comparator group for a range of different programmes relating to multiple disadvantage. This would require careful selection of areas to ensure that they are not currently participating or likely to participate in any of the relevant programmes. Potentially, this will leave areas whose comparability to the areas involved in programmes is limited. In particular, it is more likely that areas with larger recognised populations of people experiencing multiple disadvantage or greater known discrepancies between the needs of this population and available services will have applied to participate in programmes. Therefore the range of potential comparator areas may have smaller recognised populations experiencing multiple disadvantage or smaller populations with known unmet needs.

- Equally, careful thought will need to be given to the types of service that can facilitate contact with people experiencing multiple disadvantage to ensure that these services are working with comparable client groups (i.e. those with multiple disadvantage rather than one predominant need) and are not offering support that is comparable to the interventions being delivered under Changing Futures (i.e. not specialist multiple disadvantage interventions). This might be less straightforward than it would be if looking for a comparator group of clients experiencing one predominant need.
- It appears that within the Housing First evaluation this approach was successful in gathering baseline information about a suitable external comparator group. This could potentially be replicated within the proposed collaboration and result in the identification of an external comparator group, which might then be matched to the Changing Futures cohort using, for example, propensity score matching.
- The learning from the Housing First evaluation suggests that direct research at Time 2 with people experiencing multiple disadvantage proved very challenging. This raises questions about whether follow-up data drawn from direct contact with people experiencing multiple disadvantage could feasibly be collated. This would rule out the inclusion of comparison at two (or more) time points of individual-level data from outcomes measurement tools which need to be completed with the individual in question. It would also rule out the collection of follow-up qualitative data from individuals on their engagement with and experiences of support and services, or on changes in their circumstances and support networks, which might be needed to understand and qualify change in indicators for which data can be collected at Time 2.
- One example of data which could still be collated at Time 2 is administrative data, particularly if this were centrally-collated or if specific resource were provided to the local areas in the comparator group in order to collate it. Therefore this collaboration could be particularly useful if the Changing Futures evaluation (or other evaluations which might include this comparator group) intends to base outcomes measurement on administrative data. For this to enable robust comparisons, it will be important to find ways to obtain data about individuals' engagement and changes in circumstances/support networks.
- Equally, it may be possible to encourage the services which have facilitated identification of individuals in the comparator group to provide Time 2 data for measures which do not require joint completion with the client. However, this is likely to be resource-intensive for local areas so may need to be incentivised or centrally-supported in some way.
- Some consultation with area-level stakeholders would also be important to understand any changes to local systems and services which might have taken place over the period, and impacted on the way support is delivered or how people might experience it.

Overall, the approach is likely to be resource-intensive. Based on the experience of Housing First, it may prove to be a useful way to identify a suitable comparator group and to collate baseline data on this group. However, it is not guaranteed that it will result in

data at Time 2 which is comparable to the available data for the intervention group and which can be adequately qualified by information about individuals' engagement and experiences between baseline and Time 2. We would therefore not necessarily recommend its use if the sole purpose is to gather comparator data for Changing Futures. However, if MHCLG and partners would like to explore new ways of collaborating to generate comparator data then this could present an important opportunity to do so. This may provide important learning for future similar collaborations and – if successful in obtaining adequate Time 2 data – may represent an alternative mechanism for delivering counterfactual approaches for client groups for whom traditional approaches to external comparison are less feasible.