June 2023 | **Guide**

Domestic abuse programme evaluation

LEARNING FROM CASE STUDIES

Foundations
What Works Centre for Children & Families

foundations.org.uk

INTRODUCTION

Domestic abuse affects up to one in five children during their childhood and is associated with a higher likelihood of negative outcomes for children. High-quality evidence in this sector in the UK is currently limited. This is in part due to specific challenges relating to evaluation.

These challenges with evaluation can be addressed with careful long-term support and guidance, working in partnership with the domestic abuse sector. Foundations, the national What Works Centre for Children & Families, has undertaken preliminary work around Domestic Abuse evaluation and commissioned a rapid systematic review on domestic abuse interventions for children about whom there are safeguarding concerns.

This evaluation guide, illustrated with case studies from the literature, is part of a suite of outputs from this work aiming to inform future evaluations of domestic abuse programmes. It aims to provide insights into good-quality evaluation, and how common challenges in this field have been successfully tackled in previous evaluations. This guide can inform future evaluations in the UK.

What does this evaluation guide do?

This guide provides insights into opportunities to address some of the key barriers to evaluation in the domestic abuse sector through illustrative case studies from the more extensive UK and international literature. The guide considers enablers to evaluation across a variety of interventions including therapeutic and preventative programmes.

Who is this evaluation guide for?

The intended audiences for this output are those in the Domestic Abuse sector who are interested in evaluating programmes for children and families. Interested audiences may include evaluators, academics and local areas interested in undertaking their own evaluations.

How to use this evaluation guide

This guide considers the key building blocks for high-quality impact evaluations in the sector. It builds on the legacy EIF organisation "10 steps for evaluation success" and should be used in conjunction with this guide.[1]
The guide considers issues related to:

- **Participants** – by participants we mean those who participate in an evaluation, also known as beneficiaries or service users
- **Intervention** – by intervention we mean the programme or service/support offer that is being evaluated

[1] See https://www.eif.org.uk/resource/10-steps-for-evaluation-success

- **Comparator** – by comparator we mean a comparison or control not receiving the intervention, which a robust evaluation relies on to establish the relative added value of the intervention being evaluated
- **Outcomes** – by outcomes we mean the changes that the intervention aims to make for beneficiaries.

The illustrative case studies have been selected from a non-systematic review of the international literature of robust impact evaluations of domestic abuse programmes.

Participants

When considering evaluation, participants are a key consideration. The number of participants taking part in the intervention – commonly referred to as "sample size" in the research field – is particularly key to consider. A sufficient sample size in impact evaluation is required to provide a robust assessment of whether the programme has an impact and the size of that impact. It is difficult to draw confident conclusions from evaluations which are under-powered as they have a small sample size.

Sample size can be a key barrier to robust impact evaluations in the domestic abuse sector, particularly for services that deliver therapeutic offers due to their intensity. This can be a difficulty when children of different ages are beneficiaries of the service and when the service is tailored to their age and/or developmental stage.

Limitations to sample size can also come about due to issues with recruitment, retention and/or attrition (ie. drop out) and this is something that can be addressed by evaluators. In relation to domestic abuse there are a variety of concerns that may impact participation in programmes and/or programme evaluations, including not identifying as a perpetrator or survivor of abuse, fears around personal safety, and the varying levels of risk that they experience. For those experiencing domestic abuse, key triggers for attrition include prior to starting the programme and periods of transition, for example moves between refuges, safe or supported accommodation, and home.

Difficulties with sample size can also arise due to an interest in exploring impact for subgroups. For example, you may wish to explore impacts for families from different minority ethnic groups, to ensure that your programme is inclusive and accessible. To explore whether the programme is working for different groups it is necessary to have evaluations that have adequate sample size to detect these differences.

RECRUITMENT, RETENTION AND ATTRITION

Case Study – Prenatal care

A randomised trial undertaken in a prenatal care setting by Kiely and colleagues (2010) in the US provides insights into recruiting a minority ethnic population to an evaluation of an intervention which aimed to reduce risk factors including intimate partner violence during pregnancy and the postpartum period.[2] They had success in recruiting a large sample of African American women through a two-step screening process. Women who met the demographic inclusion criteria and gave their consent to participate in screening, were invited to a first-stage screening for risk factors (including domestic abuse) which was delivered through a self-completed, computer-assisted process which was simple and confidential. Those who screened positive were then invited to a second stage which involved an interview. Those who were eligible following interview were then invited to participate in the evaluation. The authors proposed that by providing women with a number of exit points during screening, women who consented to participate were more likely to stay in the study.

They further reduced attrition through the use of a set of retention strategies, including financial incentives, confirming/updating contact information at each interaction, and attention to cultural competence by all those involved in the evaluation.

The thoughtful approach to safeguarding may also have enhanced recruitment and retention in this evaluation. Staff did not want to increase the risk of abuse through participation in the study or exposing a pregnancy, and therefore carefully considered methods of communication – for example asking women if it would be safe to leave a voicemail message.

What does this mean for recruitment and retention for future evaluations?

Future evaluations could benefit from:

- Simple screening procedures to promote recruitment
- Providing opportunities for potential participants to change their minds during the recruitment process
- A suite of retention strategies focusing on facilitating continued participation through (financial and non-financial) incentives and reducing barriers such as ensuring up to date information
- Training staff involved in the evaluation in cultural sensitivity
- Careful consideration of safety in relation to participation

[2] See https://pubmed.ncbi.nlm.nih.gov/20093899/

- Piloting and testing different recruitment and retention methods in a pilot impact evaluation – guidance for undertaking pilot impact evaluations can be found on the EIF Evaluation Hub.[3]

ONLINE RECRUITMENT, RETENTION AND ATTRITION

Case Study – I-DECIDE

The evaluation of I-DECIDE – a web-based healthy relationship tool and safety decision aid for women experiencing domestic violence – provides insights into online recruitment, retention and attrition.[4]

The usual method of recruitment to the I-DECIDE tool is online and so the evaluation made use of this through a mixture of targeted women's health or domestic abuse related websites, social media platforms and general online advertising. This provided an opportunity to recruit those women who may not previously have identified themselves as experiencing domestic abuse and were not accessing other support services.

The evaluation team also considered retention strategies, which were particularly important due to the limited contact with the evaluation team. First and foremost, warm and friendly communications were used to enhance feelings of safety and support among participants, including in regular "touch base" emails. Women also received financial incentives for participation.

Further, in the specific context of domestic abuse, the team were also cognisant of the need to keep participants safe online during participation in the programme and evaluation. Women were therefore provided with browser-specific instructions such as clearing their search history and using private sessions. They were supported to set up new email accounts if they had concerns that the perpetrator could access their emails. The evaluation itself was described as a "Women's Health Study" and the I-DECIDE webpage as "Women's Wellbeing Project" and neither referenced domestic abuse in any way on public-facing pages. Protected content could only be accessed through a username and password. The websites also had a "quick exit" on every page, meaning a single click would result in the page closing and redirection to a generic webpage/search engine.

What does this mean for recruitment and retention for future evaluations?

Future evaluations where the programme is delivered digitally/virtually could benefit from:

[3] See https://evaluationhub.eif.org.uk/pilot-impact-studies/
[4] See https://www.thelancet.com/journals/lanpub/article/PIIS2468-2667(19)30079-9/fulltext

- Advertising through a broad range of online websites and social media but without mention of the term “domestic abuse”, to recruit women who may not yet have identified that they are experiencing domestic abuse and are not accessing services
- Creating online communications that enhance participant feelings of safety and security when there is no in-person contact
- Careful consideration given to the online safety of online participation
- Testing recruitment and retention methods ahead of a full impact evaluation, for example via a pilot impact evaluation – further guidance on this is provided in the EIF Evaluation Hub.[5]

Intervention

Another important aspect of an evaluation is the programme itself, sometimes called the “intervention”. Some programmes in the domestic abuse sector are what can be described as complex interventions. By “complex” we mean that they are often multi-component programmes tailored to the bespoke needs of families who are often experiencing a range of risk factors. We recognise that often local offers are informed and responsive to the context of the local population and local multi-agency systems.

Perhaps as a consequence of this complexity, domestic abuse programmes do not always have a well-articulated theory of change to act as the basis for evaluation. There was also a notable absence of manualisation of many programmes in the literature, reflecting the level of tailoring to population and context. However, having an explicit theory of change, capturing core components and critical details about the intervention such as why the intervention is needed, who it is for, how it will work, and what the intended outcomes are, is a cornerstone of evaluation. In addition, a standardised manual is important, setting out how an intervention is delivered, to ensure what families are receiving is consistent, and in line with what is intended by the theory of change. Standard manuals can include advice on flexibility and tailoring of the intervention to different participants.

Due to the variation in programmes offered across a spectrum of domestic abuse, the setting of delivery varies greatly. Many families impacted by domestic abuse engage with community-based services, but these can experience issues relating to capacity, demand and supply. Other potential settings include primary care, schools and mental health services.

[5] See https://evaluationhub.eif.org.uk/pilot-impact-studies/

MANUALISATION

Case Study – MOSAIC

The randomised trial in primary care of the MOSAIC programme of non-professional mentor support to reduce intimate partner violence[6] was underpinned by an intervention manual. The manual sets out the rationale of the intervention and how it is implemented and delivered.[7] The provision of the programme manual allowed for non-professional mentors to be selected and trained in a consistent manner and to deliver the offer in a standardised manner.

What does this mean for future evaluation of interventions?

Future intervention evaluations could benefit from:

- Setting out the rationale for the programme and its implementation in a theory of change, logic model and intervention blueprint/manual. The EIF Evaluation Hub offers advice on theory of change[8] and logic models and blueprints.[9]
- Conducting implementation and process evaluations alongside impact evaluations to understand issues such as relations between how programmes are implemented and their outcomes and the role of context – the EIF Evaluation Hub provides guidance on conducting these at an early stage.[10]

PRENATAL CARE SETTING

Case Study – Prenatal care

Prenatal care was the setting for this evaluation of an intervention to reduce recurrence of intimate partner violence during pregnancy and postpartum.[11] Routine prenatal care appointments afforded the opportunity to deliver the intervention and to recruit women to the evaluation. Sessions were conducted privately within clinics, or proximate to them, immediately before or after usual prenatal care. There was no formal assessment of acceptability of this setting, but we can draw some preliminary conclusions from the relatively high rates of recruitment and retention. However, it is important to highlight that this evaluation was conducted in the US which has a different prenatal care system and therefore we need to be careful about what conclusions we can draw.

[6] See https://bmcpublichealth.biomedcentral.com/articles/10.1186/1471-2458-11-178
[7] See https://www.latrobe.edu.au/__data/assets/pdf_file/0009/217485/MosaicManualv2FINAL.pdf
[8] See https://evaluationhub.eif.org.uk/theory-of-change/
[9] See https://evaluationhub.eif.org.uk/logic-models-blueprints/
[10] See https://evaluationhub.eif.org.uk/implementation-and-process-studies/
[11] See https://pubmed.ncbi.nlm.nih.gov/20093899/

What does this mean for the setting of future evaluations?

Where interventions are aiming to reduce domestic abuse during pregnancy and the postpartum period, it could be useful for evaluators to consider:

- Delivering the intervention in a setting where women are already receiving prenatal care
- Co-ordinating intervention sessions with routine visits
- Conducting an implementation and process evaluation to explore issues relating to delivery that may need to be resolved before a full impact evaluation is considered – there is guidance on this in the EIF Evaluation Hub.[12]

COMMUNITY SETTING

Case Study – Community treatment of PTSD

An evaluation of post-traumatic stress disorder in children exposed to domestic abuse was undertaken in a community setting.[13] To facilitate evaluation in this context a number of factors were taken into consideration in the design of the intervention and trial methodology. In terms of the intervention the "dose" (the number of sessions delivered to families) had to be adapted, reducing from the usual 12 to 8 to fit with an existing offer. In terms of the evaluation, the procedures were simplified taking a pragmatic approach, for example there was minimal exclusion criteria meaning that as many families were eligible for participation as possible, including those in current contact with the perpetrator. This ensured that the study was more inclusive and representative.

What does this mean for the setting of future evaluations?

When designing evaluations in community settings careful consideration should be given to:

- Adaptation of intervention programmes to real-world settings, for example fit with existing offers
- Simplification of evaluation procedures to facilitate pragmatic evaluation
- The impact of evaluation inclusion and exclusion criteria on inclusivity and representativeness
- Conducting an implementation and process evaluation to explore key unknowns relating to delivery in this setting.

Comparator

When exploring the impact of an intervention, such as a domestic abuse programme, statistical comparisons are made to groups (or time periods) which have not been exposed to the intervention. This comparison group (or time period) is commonly known as a counterfactual (or a comparator, or control) and is a proxy for what is expected to have happened in the absence of the

[12] See https://evaluationhub.eif.org.uk/implementation-and-process-studies/
[13] See https://jamanetwork.com/journals/jamapediatrics/fullarticle/384142

intervention. For robust impact evaluations the comparator is critical to be able to establish the relative added value of the intervention.

Impact evaluations can adopt experimental or non-experimental designs. An important distinction between the two types of evaluation is how individuals (or groups) are allocated to the intervention or to the comparator. In experimental evaluations (such as randomised control trials, RCTs) they are allocated at random, and in non-experimental (such as quasi experimental designs, QEDs) the allocation is not random.

Experimental studies, where there is randomisation, are more robust but are not always feasible, ethical or acceptable, and this may be particularly true in the domestic abuse sector. In non-experimental studies, where there is no randomisation, it can, however, be difficult to identify an appropriate comparison group.

Even where there is randomisation, if the allocation is not concealed (that is people know who is receiving the intervention and who is not) then this can introduce bias into the results. Concealing the allocation to intervention or comparator (known as blinding) to programme recipients, practitioners and evaluators can provide a fairer test of programme impact. However, achieving blinding in real life can be difficult for evaluators, and may be unfeasible for recipients
Another, real-life difficulty for evaluators, is an issue known as "contamination". This is where a family allocated to comparator group receives an intervention programme (or part of the intervention) when they should not. This can happen for a range of reasons, including simple word of mouth between practitioners or participants, but there are steps evaluators can take to reduce the risk of contamination.

RANDOMISATION AND CONTAMINATION

Case Study – MOSAIC

The MOSAIC trial, which explored non-professional mentor support to reduce intimate partner violence, randomised at the level of the cluster – in this case the primary care clinics or teams – rather than at the level of the individual.[14] This meant that they could avoid contamination – where those in the comparator group receive or are exposed to the intervention unintentionally. Having sites that offered the intervention and others that didn't reduced the chance that patients at the comparator clinics would experience the intervention. This approach was considered to be successful as only one woman in the comparator group received the intervention.

What can we learn for future evaluations?

Future evaluations can consider the following in relation to comparators:

- If the evaluation will include randomisation, will this be conducted at an individual level or at a cluster level

[14] See https://bmcpublichealth.biomedcentral.com/articles/10.1186/1471-2458-11-178

- Cluster randomisation can reduce the likelihood of contamination, although may bring other challenges – for example increasing the sample size required – and therefore needs careful consideration
- Data should be collected to understand whether contamination has occurred and to what extent
- Whether to conduct a pilot impact evaluation to explore whether contamination is an issue ahead of a full impact evaluation – guidance on pilot impact evaluations is available in the EIF Evaluation Hub.[15]

QED AND WAITING LIST CONTROL

Case Study – DART

The DART evaluation, undertaken by NSPCC, employed a quasi-experimental approach.[16] In order to have a comparator group of families not receiving the intervention, they created what is known as a "waiting list control". This meant that some families were added to a waiting list for the intervention and received it after the intervention and evaluation had been completed with a first cohort of families. This ensured that no families were left without support, but there was often a wait time before families were able to access support.

What can we learn for future evaluations?

When planning and conducting future evaluations in the domestic abuse sectors, evaluators need to carefully consider several issues:

- Whether a waiting list control would be acceptable as a comparator in an evaluation to families and to practitioners delivering interventions – this could be explored by co-developing the evaluation design
- Whether a waiting list control for families would be ethical, balancing the need for timely interventions – however, there is also a need to consider the implications of delivering services to vulnerable children that have not been subject to rigorous evaluation and may be associated with unintended consequences
- The length of the intervention and the time it may take for anticipated outcomes to be observed, and implications of this on the desired follow-up period following the intervention and the waiting list duration – for instance, short follow-up periods may not provide data on the longer-term impact of services that are usually required by evaluations; shorter follow-up periods are more common when conducting pilot impact evaluations,[17] but full impact evaluations often favour longer-term outcomes.[18]

[15] See https://evaluationhub.eif.org.uk/pilot-impact-studies/
[16] See https://learning.nspcc.org.uk/media/2356/impact-evaluation-scale-up-domestic-abuse-recovering-together.pdf
[17] See https://evaluationhub.eif.org.uk/pilot-impact-studies/
[18] See https://evaluationhub.eif.org.uk/measuring-impact/

“ACTIVE” CONTROL GROUPS

Case study – Community treatment of PTSD

In an evaluation of the treatment of post-traumatic stress disorder in children exposed to domestic abuse the comparison group received what can be described as an “active control”.[19] By active we mean that they received an alternative programme of support (this can also be service-as-usual) which is similar in the way that it is delivered, for example the number of sessions received. This allows us to understand whether it is the content of the sessions rather than just the number of interactions that families have with a professional that is important and brings added value.
The intervention group received trauma-focused cognitive behavioural therapy and the control received standard cognitive behavioural therapy. In both the intervention and the “active” control group offer, children and parents each received individual 45-minute sessions for eight consecutive weeks (or until programme completion). Children and parents saw the same therapist. The only difference between the delivery of the intervention and the active control was that, in the trauma-focused intervention, in parts of two sessions children and parents were seen together. Otherwise only the content of the sessions was different and therefore it is easier to draw conclusions about whether it is the intervention content that is making a difference to outcomes.
This active control condition meant that children who were victims of domestic abuse were not left without support. However, it did mean that it could be more difficult to demonstrate an effect of the trauma-focused cognitive behavioural therapy, given that it is compared to another form of support rather than no support at all.

What can we learn for future evaluations?

In future evaluations the following can be given consideration:

- What the comparator group of families receive during an evaluation and whether they should receive an “active” programme (which is similar in terms of delivery mode and dose) – this is closely linked to considerations as to whether no support (or delayed support) is felt to be appropriate for families that may be allocated to the comparator group
- Whether “active” control groups are acceptable to families and to practitioners delivering interventions
- Whether an “active” control group would create challenges in understanding the impact of the intervention of interest
- What opportunities this provides for comparing a current intervention with a current intervention plus an additional/new intervention component.

OUTCOMES

The primary purpose of an impact evaluation is to investigate whether the intervention delivered has led to a change in potential outcomes targeted by the intervention (as highlighted in the

[19] See https://jamanetwork.com/journals/jamapediatrics/fullarticle/384142

intervention's theory of change). As a result, identifying key outcomes and accurately measuring changes in those outcomes over the duration of the intervention are key considerations for impact evaluations.

Due to the range of different types of programmes for children and families in relation to domestic abuse, there are a range of potential outcomes. When considering outcomes, a useful resource has been developed by academics from University College London (UCL), who have undertaken work to understand the types of outcomes that are important for children exposed to domestic abuse.[20] The team worked with those who use, deliver or commission services to agree which were the most important outcomes to measure. In future evaluations there would be an aim for these core outcomes to be reported as a minimum where the child is the focus of the programme even if it is also delivered to other members of the family.

There are also a wide range of potential tools that can be used to measure outcomes. The key factors when considering an evaluation are the use of measurement tools that are valid, reliable and appropriate for use with the specific population. The Child Outcomes Research Consortium (CORC) provides more information about validity, reliability and appropriateness of measures.[21] There is also a need for measures that are acceptable to the population who you are working with. It is also important to consider the use of objective (e.g. independent assessments) versus more subjective measures (e.g. self-reports). CORC is a useful resource when considering which outcome measures to use and provides information on outcome measures and their properties which are important for evaluation (such as their validity and reliability).[22]

VALIDATED OUTCOME MEASURES

Case Study – DART

In the NSPCC evaluation of the DART programme[23] the Strengths and Difficulties Questionnaire (SDQ) was included as an outcome measure. The SDQ is a widely used measure of children's emotional and behavioural difficulties and their prosocial behaviours. The SDQ entry in the CORC library states that at present there is some promising evidence relating to the properties of reliability and validity.[24] However, there are some issues remaining around the use of this measure. It can be useful, for example, to gather scores from multiple informants to ensure robust insights from this type of more subjective measure. The DART evaluation attempted to do this by collecting both parent and child report SDQ outcomes, and there were some differences in changes between

[20] See https://www.ucl.ac.uk/children-policy-research/projects/core-outcome-sets-family-and-child-focused-interventions
[21] See https://www.corc.uk.net/media/1487/corc-approach-v3.pdf
[22] See https://www.corc.uk.net/outcome-experience-measures/
[23] See https://learning.nspcc.org.uk/media/2356/impact-evaluation-scale-up-domestic-abuse-recovering-together.pdf
[24] See https://www.corc.uk.net/outcome-experience-measures/strengths-and-difficulties-questionnaire-sdq/

parents and child reported scores, with mothers rating a greater reduction in child difficulties than children themselves.

What can we learn for future evaluations?

Future evaluations will work towards including the core outcome set developed by UCL as a minimum where child outcomes are the focus of the intervention and at present can consider:

- The use of valid and reliable outcome measures
- The use of multiple informants where possible
- Conducting a study to pilot for outcomes ahead of a full impact evaluation – guidance is provided on the EIF Evaluation Hub.[25]

Exploring the use of (more) objective measures

In the evaluation of a community treatment for children exposed to domestic abuse and experiencing post-traumatic stress disorder[26] the evaluators included a more objective outcome measure: the Kiddie-Sads-Present and Lifetime Version (K-SADS-PL) diagnostic tool. This is a semi-structured interview which can draw together information from multiple informants (parent and child) and is determined by the clinical judgement of the trained raters. In this evaluation, two project coordinators blinded to allocation completed the assessments, and checks were conducted to ensure reliability of their judgements.

What does this mean for future evaluations?

Future evaluations may benefit from considering:

- The value of the use of measures with increased objectivity balanced with the increased time taken to collect this type of data
- The value of undertaking a pilot for outcomes study before proceeding to a full impact evaluation – guidance on conducting pilot impact studies is provided in the EIF Evaluation Hub.[27]

ROUTINELY COLLECTED DATA

Case Study – Prenatal care

In a US evaluation of an intervention to reduce recurrence of prenatal and postpartum domestic abuse, evaluators included objective medical records as a measure of the impact of the evaluation on infant outcomes.[28] Intimate partner violence is a known risk factor for pregnancy complications such as low birth weight, preterm delivery and neonatal death. Data on infant outcomes were extracted from medical records on delivery as part of the evaluation.

[25] See https://evaluationhub.eif.org.uk/pilot-impact-studies/
[26] See https://jamanetwork.com/journals/jamapediatrics/fullarticle/384142
[27] See https://evaluationhub.eif.org.uk/pilot-impact-studies/
[28] See https://pubmed.ncbi.nlm.nih.gov/20093899/

This study was conducted in the US; in the UK there may be challenges to collecting routine medical records in domestic abuse evaluations, although there may be opportunities for linkage with routine administrative datasets. There may also be other objective measures of outcomes that could be explored; however, information sharing and linking around sensitive issues like domestic abuse requires careful consideration to ensure acceptability and safety.

What does this mean for outcomes in future evaluations?

Future evaluations could consider:

- If and how routinely collected data could be safely included in programme outcome assessments
- Sensitivities around the use of routinely collected data in evaluations or domestic abuse programmes and seek input from experts by experience
- The need for consent to linkage to any administrative datasets.

UNDERSTANDING WHAT CHANGE IS MEANINGFUL

Case study – Meaningful change

There is also a need for evaluators to consider what is meaningful change. A US evaluation of a domestic abuse enhanced home visiting programme found a statistically significant reduction in frequency of experience of domestic violence across time between the two groups on their outcome measure – however, this difference was small.[29] Evaluations in the domestic abuse sector may want to consider what is meaningful change for survivors and their families.

What does this mean for future evaluations?

Evaluators are encouraged to consider:

- What change would be meaningful for families and what would be achievable
- Working with survivors and practitioners to understand what meaningful change would look like.

SUMMARY AND NEXT STEPS

This guide is a high-level summary of some of the key challenges with evaluation of domestic abuse programmes, and how these could be overcome. We welcome further feedback and debate on this topic – please email info@foundations.org.uk to contribute, using 'Domestic abuse evaluation case studies' as the subject.

[29] See https://pubmed.ncbi.nlm.nih.gov/27206047/

Foundations intends to conduct and fund further evaluations of domestic abuse programmes in England – both findings and further learning on how to conduct high-quality evaluation will be published. This will begin with a Domestic Abuse Evaluation Accelerator Fund, which will fund pilot impact evaluations. These pilot studies will test robust impact evaluations at a smaller scale initially, to test for the acceptability and viability of different methods of data collection and use of outcome measures.