

Dadansoddi ar gyfer Polisi



Analysis for Policy



Llywodraeth Cymru
Welsh Government

SOCIAL RESEARCH NUMBER:

25/2017

PUBLICATION DATE:

09 MAY 2017

Families First: An evaluability assessment

Mae'r ddogfen yma hefyd ar gael yn Gymraeg.
This document is also available in Welsh.

Title: Families First : An evaluability assessment

Author(s) / Company

Hannah Thomas, Welsh Government
Hayley Collicott, Welsh Government

Views expressed in this report are those of the researcher and not necessarily those of the Welsh Government

For further information please contact:

Welsh Government

Cathays Park

Cardiff

CF10 3NQ

Tel: 0300 025 0508

Email: Hannah.Thomas018@wales.gsi.gov.uk

Table of contents

- 1. Summary2
- 2. Introduction3
- 3. The Families First Programme4
- 4. Programme Rationale5
- 5. Evaluability of Families First.....8
- 6. Currently Possible Approaches16
- 7. Potential Future Approaches.....19
- 8. Conclusions.....21

1. Summary

- 1.1 The aim of this report is to set out the issues and challenges that arise when attempting to conduct a robust evaluation of the Families First programme. The report makes use of an evaluability checklist which systematically explores all aspects of the programme to determine the degree to which it can be evaluated and in what respects.
- 1.2 The completed checklist, combined with the findings from evaluations conducted to date, provides a clear understanding of what can currently be concluded with reasonable certainty about the effects of the Families First programme, and what cannot.
- 1.3 This report makes the following recommendations for any future evaluations of Families First if no changes were to be made;
- Stakeholders attitudes and perceptions of the Families First programme
 - How the programme is implemented and the quality of the implementation
 - Whether the programme design is fit for purpose
- 1.4 Without changes to the data being collected about the families using the Families First services, future evaluations will only be able to replicate previous research. Introduction of additional data collection could lead to more robust evaluations. Such data collection could include;
- Developing new systems for collecting individual level data about service use for the families using Families First, rolled out as standard across all local authorities
 - Establishing any additional services or interventions received by families First recipients and linking this to the individual level data.
- 1.5 This could then allow the following questions about Families First to be considered;
- The level of take up of Families First services by individuals and families
 - Whether the level of engagement with Families First services leads to different outcomes
 - The interactions between the different Families First services
 - The longer term impacts of Families First
 - The demographics of those with higher or lower levels of engagement with Families First services
 - Whether different individual characteristics lead to different outcomes for those who engage with Families First services

1.6 If these changes were to be made, the following may then be possible:

- Analysis of the individual level data to test whether individual characteristics or wider contextual factors affect the impact of services.
- Statistical analysis at an individual level to determine the effects of Families First services
- Where additional services or interventions are being used, whether these have an additional effect on those using Families First services.

2. Introduction

2.1. The aim of this report is to set out the issues and challenges that arise for conducting a robust evaluation of the Families First programme. Several evaluations of the programme have been conducted since 2012 using a variety of methods. Each evaluation has made reference to problems which arise when evaluating the programme; this report aims to set out these issues and challenges, by completing a coherent evaluability assessment of the programme.

2.2. The term evaluability is described as *“The extent to which an activity or project can be evaluated in a reliable and credible fashion”* (OECD-DAC 2010; p.21). Any assessment of the evaluability of a programme will attempt to determine how successful any evaluations will be in providing useful and reliable evidence.

- a) The purpose of this assessment is to determine if future evaluations are feasible, what approaches they could take, and what changes to data collection would be needed to facilitate them.

2.3. The checklist used in the report is taken from The Department for International Development (DFID) and is used to ensure all aspects of the programme are considered systematically. The checklist can be used to determine how effective any evaluation is likely to be, and to identify areas in which the programme can be improved.

2.4. In general, the main questions which should be asked of any programme¹ are:

- Whether an impact assessment is useful?
- Whether it is plausible to expect impacts?
- Whether it is feasible to measure impacts?

Understanding how the programme is being delivered, the context in which it operates, the aims and objectives of the programme, and the rationale for the programme is essential to answer these questions.

2.5. Section 3 sets out the Families First programme and what the guidance states should be offered as part of Families First services. Section 4 describes the programme in more detail and explores the rationale and the logic model behind Families First. Section 5 includes a completed checklist to determine what

¹ http://pdf.usaid.gov/pdf_docs/PNADN200.pdf

aspects of the programme allow for an effective evaluation, and what aspects pose challenges for evaluation. Section 6 then details the approaches which have been taken in previous evaluations, and explores what approaches could be possible if changes were made, and the advantages and costs associated with these changes.

3. The Families First Programme

- 3.1 Families First aims to improve the design and delivery of local authorities' family support services through offering joined-up support, and through co-ordinating the organisations working with families.. The programme promotes the development within local authority areas of effective multi-agency systems and support, with a clear emphasis on prevention and early intervention.
- 3.2 The programme seeks to improve early access to and the delivery of preventative support and is designed to complement mainstream services. Families First is an example of the Welsh Government's 'invest to save' principle, supporting families before their problems become more complex and therefore costly to resolve.
- 3.3 The Families First programme is essentially an 'innovation programme' that requires local authorities to develop their own models of working to address the needs identified in their area. The management and governance of the programme therefore necessitates a balance between addressing requirements that are core to all local authorities and promoting variation in the way the programme is implemented locally. These principles underpinning the programme are that it should be;
 - Family-focused, taking a whole family approach to improving outcomes;
 - Bespoke, tailoring help to individual family circumstances;
 - Integrated, co-ordination of planning and service provision across organisations, ensuring that needs assessment and delivery are jointly managed, with seamless progression for families between different interventions and programmes;
 - Pro-active, seeking early identification and appropriate intervention for families;
 - Intensive, taking an approach and focus which adapts to families' changing circumstances;
 - Local, identifying the needs of local communities and developing service delivery to fit needs;

- 3.4 Families First partnerships are given substantial flexibility in how they deliver the programme, but, local delivery is expected to incorporate a range of common features:
- Strategic planning, based on local audits of need and current provision;
 - Joint Assessment Family Framework (JAFF) development and implementation;
 - Team Around the Family (TAF) development and implementation;
 - Strategically commissioned projects;
 - A disability element; and
 - Use of learning sets.
- 3.5 Families First succeeds the Cymorth grant which Welsh authorities received from 2003 to support children and young people. Families First was rolled out in April 2012 across all 22 Local Authorities in Wales, following a pioneer phase which tested a range of delivery models across five consortia.
- 3.6 Four national intended outcomes have been identified for Families First, which are that;
- Working age people in low income families gain, and progress within employment;
 - Children, young people and families, in, or at risk of, poverty, achieve their potential;
 - Children, young people and families are healthy, safe and enjoy well-being;
 - Families are confident, nurturing and resilient.

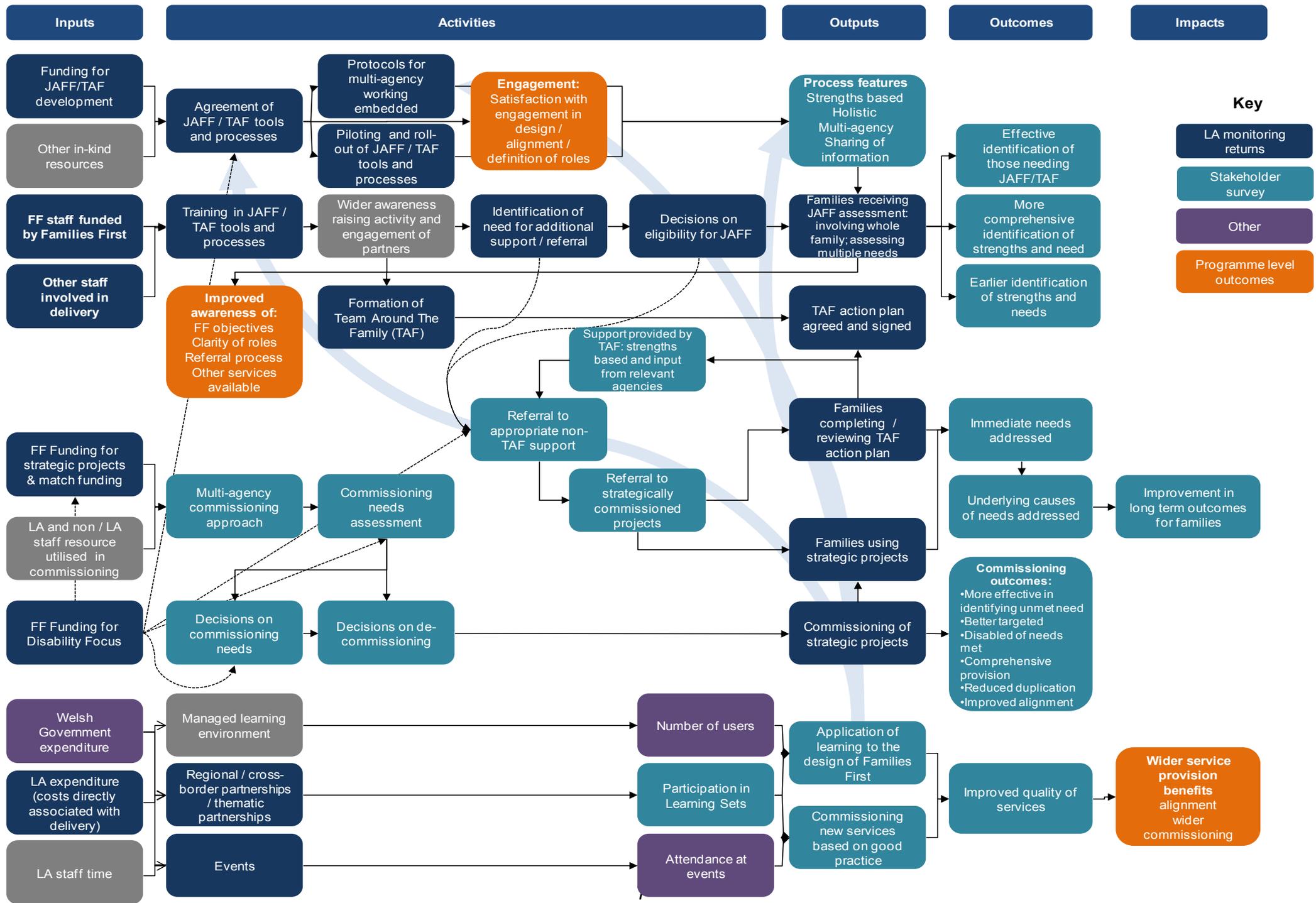
A number of population indicators have been chosen which should help gauge whether Families First interventions are on track to achieve their intended outcomes.

4. Programme Rationale

- 4.1. In evaluating a policy or programme, it is critical to develop a full understanding of the theory behind it. This means clearly defining the steps involved in delivering the policy, the context in which it is operating, what resources are required, what is actually undertaken as part of the policy and what changes are expected as a result. One way in which this understanding can be developed and described is through a logic model.
- 4.2. The Magenta Book states '*Logic models describe the relationship between an intervention's inputs, activities, outputs, outcomes, and impacts*' (pg. 22).
- Inputs are the quantifiable resources that are required to deliver the policy/programme. This can include funding, staff, buildings or physical materials, such as books or leaflets.
 - Activities are the actual tasks involved in delivering the policy. For example, this could be providing training courses, distributing information or building a road.

- Outputs are the direct results of the activities and what the recipient does as a response to the activities. So in the example of providing a training course, the output could be people attending the course.
 - Outcomes are results of the outputs and will occur some time after the activity. Continuing with the example of a training course, the outcome would ideally be that the recipients' skills or knowledge are increased.
 - Impacts are the long-term, wider effects of a policy/programme, and are usually considered the ideal goal. They are often considered on a population level and so would describe how an activity can lead to lasting changes in the group involved in the activity. On a national level this may be incomes are increased, or health improves.
- 4.3. A logic model can also be thought of as a means of portraying a theory of change. A theory of change attempts to explore the assumptions around a programme and on what basis it claims to be able to bring about the outcomes and impacts that are stated in the logic model². It can be broadly thought of as a critical appraisal of a logic model, by which the links between the inputs, activities, outputs, outcomes and impacts are explored and the assumptions are stated and challenged.
- 4.4. Inherent in all social policies are assumptions about how the policy will engage with the target recipients. These can range from the size of the expected recipient population to how the activities will influence the recipients outcomes. It is these assumptions that are tested in policy evaluations.
- 4.5. In addition to the underlying assumptions, a theory of change also sets out the context in which the policy or programme is operating. This is vitally important, as the context can have a big influence on the effect of an intervention. The context in this case refers to the political, social and economic systems in place, the situation of the target beneficiaries prior to implementation and who is involved in the both the implementation of the programme and the problem being addressed. It should also include the intended mode of implementation, as 'fidelity' to this can have a big effect on how successful an intervention is.

² http://r4d.dfid.gov.uk/pdf/outputs/mis_spc/DFID_ToC_Review_VogelV7.pdf



5. Evaluability of Families First

- 5.1. This checklist has been taken from pp. 19-23 of: Davies, R., 2013. *Planning Evaluability Assessments: A Synthesis of the Literature with Recommendations*. Report of a study commissioned by the DFID.
- 5.2. The checklist identifies the following dimensions of evaluability;
- Evaluability “in principle”, given the nature of the project theory of change
 - Evaluability “in practice”, given the availability of relevant data and the capacity of management systems able to provide it.
 - The utility and practicality of an evaluation, given the views and availability of relevant stakeholders
- 5.3. This aims to act as a systematic process which provides an accountable means of ensuring coverage of all the relevant issues. The expected outputs concern not only the evaluability of the project but also the practicality and utility of an evaluation. It is hoped this will inform future evaluations and research and potentially suggest how the programme could be changed to allow for more effective evaluations.
- 5.4. Table 1 shows the completed evaluability checklist for the Families First programme.

Table 1. Completed evaluability checklist for Families First		
1. Project Design		Notes – (positives and negatives)
Clarity?	Are the long-term impact and outcomes clearly identified and are the proposed steps towards achieving these clearly defined?	<ul style="list-style-type: none"> • Aims are defined to reduce some Adverse Childhood Experiences (ACEs), including verbal/physical/sexual abuse, exposure to domestic violence, exposure to substance/alcohol misuse, parental separation and poor parental mental health. Families First provides multi-agency early intervention / preventative family centred support. • Families First guidance is clear about the aims of the programme and identifies clear steps to achieving these aims in terms of JAFF / TAF and commissioned projects • While there has been some variation in the main aims of the programme, in particular stepping away in 2016 from delivering activities which solely focus on employment, the main components and approach have remained the same. <ul style="list-style-type: none"> • intended impacts are not clearly defined and that, even if they were, data collection/monitoring/evaluation isn't in place to allow for assessment of whether they are being met. • Due to the bespoke nature of the programme there is no uniform model of delivery across Wales, JAFF and TAF use and models vary between Local Authorities. • ACEs are useful as summary but don't cover everything which Families First aims to support.

<p>Relevant?</p>	<p>Is the project objective clearly relevant to the needs of the target group, as identified by any form of situation analysis, baseline study, or other evidence and argument? Is the intended beneficiary group clearly identified?</p>	<ul style="list-style-type: none"> • Interventions delivered to the families are bespoke and based on individual family needs • The JAFF includes an early assessment to identify need and most appropriate ways to address this for the family • Services are commissioned based on identified local need <hr/> <ul style="list-style-type: none"> • There is no definition of the needs families should present with to receive support from the programme. This would be determined at a local level • Difficult to assess the relationship between interventions and outcomes – limited to qualitative information based on family circumstance. • No target group as eligibility criteria are based on needs identified locally – programme is “open to all, dependent on need”. • No clear comparison group to look at those who use Families First services and those who don’t.
<p>Plausible?</p>	<p>Is there a continuous causal chain, connecting the intervening agency with the final impact of concern?</p> <p>Is it likely that the project objective could be achieved, given the planned interventions, within the project lifespan? Is there evidence from elsewhere that it could be achieved?</p>	<ul style="list-style-type: none"> • Programme design is based on robust evidence of the benefits of early intervention, holistic, whole family support. • Guidance supports the use of evidence based interventions. • Services are commissioned based on identified local need with the Families First aims in mind. <hr/> <ul style="list-style-type: none"> • No specific project lifespan – available for as long as families need it. Impacts are not likely to be seen immediately but may appear after interventions have ceased. • Difficult to quantify relationship between interventions and impacts at a national level. Each local authority commission projects based on their local area need. • Some of the comparable evidence to Families First used in the pioneering phases is international and may not be generalizable to the Welsh context. • Some stakeholder perceptions are that the aspirations of Families First in improving population outcomes are high.
<p>Validity and reliability?</p>	<p>Are there <i>valid</i> indicators for each expected event (output, outcome and impact levels)? I.e. will they capture what is expected to happen? Are they <i>reliable</i> indicators? I.e. will observations by different observers finding the same thing?</p>	<ul style="list-style-type: none"> • Four national population outcomes identified at design of programme through RBA and population level indicators developed to establish whether Families First is on track to achieving the population outcomes. • Core set of indicators should allow for national assessment of progress made as a result of the programme • Local authorities report on these indicators to Welsh Government using the same reporting method which defines the population indicator and provides examples to help local authorities provide this information <hr/> <ul style="list-style-type: none"> • There are some discrepancies between local authorities in filling in the reporting information • Unable to measure output at the individual level, thus unable to determine whether any change could be explicitly attributable to Families First. • Local authorities have been encouraged to develop their own local-level indicators which renders it more difficult to achieve a Wales-wide understanding of performance and achievements of Families First.

		<ul style="list-style-type: none"> RBA represents a useful tool in providing managerial evidence for accountability but it does not provide a substitute for evaluation on a local nor national level. Population-level change in the indicators is likely to be subject to many factors beyond the intervention. It is not possible to isolate its contribution with any certainty.
Testable?	Is it possible to identify which linkages in the causal chain will be most critical to the success of the project, and thus should be the focus of evaluation questions?	<ul style="list-style-type: none"> Information is collected on performance measures, families with additional needs relating to disability and referrals into Families First for families with a JAFF / TAF
		<ul style="list-style-type: none"> All interventions are based on family need meaning provision is different for each family Lack of individual level data means it is not possible to test the 'post-treatment' state of each recipient of the Families First programme. Lack of individual level data means it is not possible to test any linkages between services.
Contextualised?	Have assumptions about the roles of other actors outside the project been made explicit? (both enablers and constrainers) Are there plausible plans to monitor these in any practicable way?	<ul style="list-style-type: none"> Local authorities are responsible for delivery of the programme in accordance with programme guidance. Their role is delivering the programme is clearly set out and understood. Welsh Government receives copies of the Local Authorities Families First plans which can be used to monitor local authority Families First services, these are updated throughout the year where changes are made
		<ul style="list-style-type: none"> Localised delivery means delivery of the programme varies across Wales as services are commissioned in response to identified local needs. No comprehensive data collection of the differences in local delivery. Some changes have occurred within the current lifespan of the programme, in that a move away from any projects specifically focused on employment only was brought in 2016.
Consistent?	Is there consistency in the way the Theory of Change is described across various project multiple documents (Design, M&E plans, work plans, progress reports, etc.)	<ul style="list-style-type: none"> Understanding of the programme's aims is consistent, with the emphasis on whole family approach. A single regional and portable JAFF was developed for use in each local authority. Theory of change developed for Families First as part of the evaluation
		<ul style="list-style-type: none"> No global theory of change produced at inception, project development of pioneer stages Interventions change per local authority depending on the local needs identified Interventions per family are different due to bespoke nature of the programme
Complexity?	Are there expected to be multiple interactions between different project components? [complicating attribution of causes and identification of effects] How clearly defined are the expected interactions?	<ul style="list-style-type: none"> Families First promotes holistic multi-agency working and multiple interactions are to be expected. The nature of these interactions varies depending on how the programme is delivered at a local level. The effect of these interactions are unknown as the nature of the programme means take up of services varies. The effects of the intervention can't be isolated from other influences and factors, therefore attribution is an issue for Families First. The extent of variation in delivery and variation in governance structures suggests that the delivery of Families First will be more successful in some

		local areas than others. There is also some uncertainty if all models will deliver the intended outcomes.
Agreement?	To what extent are different stakeholders holding different views about the project objectives and how they will be achieved? How visible are the views of stakeholders who might be expected to have different views?	<ul style="list-style-type: none"> Stakeholders have a consistent understanding of the aims of the programme. High-level buy-in from stakeholders and the programme is highly valued. Local authorities welcome the development of TAF and JAFF models being fit for purpose locally, appropriate to local delivery structures and build on existing strengths.
		<ul style="list-style-type: none"> Non-prescriptive programme guidance and local needs based delivery has resulted in significant variations in programme delivery.
2. Information Availability		
Is a complete set of documents available?	...relative to what could have been expected? E.g. Project proposal, Progress Reports, Evaluations / impact assessments, Commissioned studies	<ul style="list-style-type: none"> Complete documentation is available on the programme including <ul style="list-style-type: none"> Three previous evaluation reports Programme guidance Case study evidence from previous evaluations
		<ul style="list-style-type: none"> Evaluation and evidence only relevant until September 2015 Revised guidance currently being drafted EIA currently being drafted.
Do baseline measures exist?	<p>If baseline data is not yet available, are there specific plans for when baseline data would be collected and how feasible are these?</p> <p>If baseline data exists in the form of survey data, is the raw data available, or just selected currently relevant items? Is the sampling process clear? Are the survey instruments available?</p> <p>If baseline data is in the form of national or subnational statistics, how disaggregated is the data? Are time series data available, for pre-project years?</p>	<ul style="list-style-type: none"> JAFF has been suggested in previous evaluations to be a tool to establish a baseline against which to measure progress and capture 'distance travelled'. Initial analysis of population indicators agreed with local authorities in year one of evaluation acted as a baseline against which the four programme outcomes could be measured against in later evaluations.
		<ul style="list-style-type: none"> Analysis of trend data at the population level does not provide much insight into the mechanisms that are driving a more theoretically informed approach to evaluating the services No baseline survey of families was conducted at implementation and it might be possible to do this retrospectively for those with a JAFF / TAF but not for all those using Families First provided services.
Is there data on a control group?	Is it clear how the control group compares to the intervention group? Is the raw data available or just summary statistics? Are the members of the control group identifiable and potentially contactable? How frequently has data been collected on the status of the control group?	<ul style="list-style-type: none"> Control group not easily identifiable for those who do not have a JAFF / TAF One of the main difficulties for evaluation is the 'self-selection' into the programme. Adequate control groups are probably impossible, given the need to control for this, alongside the way the programme was designed and implemented. No data collected on individual basis to track Families First service use. Local Authorities used their own systems for measure baseline and family progress which are not comparable across all 22 authorities.

<p>Is data being collected for all the indicators?</p>	<p>Is it with sufficient frequency? Is there significant missing data? Are the measures being used reliable i.e. Is measurement error likely to be a problem?</p>	<ul style="list-style-type: none"> • Performance measures are collected about families completing a JAFF / signing a TAF action plan • There are national statistics produced on the indicators used within Families First • Previous evaluations developed a Family Outcome tool (FOT) to measure both 'soft' and 'harder' outcomes. <hr/> <ul style="list-style-type: none"> • No information collected on families which are using commissioned project services through family first. • Local-level indicators established through results based accountability. Focus on population outcomes, or evidence for accountability, does not by itself produce evidence of 'what works'. • Families define their own goals to work towards during the life of the TAF and families have a significant input into determining the 'scores' they record on Distance Travelled Tools. The Family Outcomes Tool only captures data for families who are helped by TAF only, and not those who only benefit from Families First projects. • No data collected on individual basis to track Families First service use. • The FOT data is no longer collected as the evaluation which used this has concluded.
<p>Is critical data available?</p>	<p>Are the intended and actual beneficiaries identifiable? Is there a record of who was involved in what project activities and when?</p>	<ul style="list-style-type: none"> • There is some data collected on performance measures for those completing a JAFF / signing a TAF but this is not linkable to any other data source. • Data is collected on commissioned projects but this varies per Local Authority and by project. • Previous evaluations used data captured by local authorities (using Distance Travelled Tools) to provide an overall assessment of what proportion of families experiencing Families First have seen improved outcomes. <hr/> <ul style="list-style-type: none"> • No available data on what Families First services have been used by individuals. • Attribution issues as the effects of the intervention can not be isolated from other influences and factors. This is especially difficult for Families First as there are multiple initiatives and multi-agency interventions which aim to deliver complex solutions to 'joined up' problems. • Families First delivered at a local level is linked to the delivery of other whole-family interventions such as Flying Start and Integrated Family Support Service, as a result at a local level this varies significantly by authority.
<p>Is gender disaggregated data available?</p>	<p>In the baseline? For each of the indicators during project intervention? In the control group? In any mid-term or process review?</p>	<ul style="list-style-type: none"> • Performance measures are captured for those completing a JAFF / TAF, so there is some data available <hr/> <ul style="list-style-type: none"> • The lack of data means no disaggregation is available at the individual level

If reviews or evaluations have been carried out...	Are the reports available? Are the authors contactable? Is the raw data available? Is the sampling process clear? Are the survey instruments available?	<ul style="list-style-type: none"> • Previous evaluations available online on the Welsh Government website with the author details. • Details on the survey instruments used in previous evaluations is clear. • Staff changes mean lead authors may not be easily available.
Do existing M&E systems have the capacity to deliver?	Where data is not yet available, do existing staff and systems have the capacity to do so in the future? Are responsibilities, sources and periodicities defined and appropriate? Is the budget adequate?	<ul style="list-style-type: none"> • No current capacity issues from the local authority perspective on the information they are currently recording. • Data is not available at the individual level and for all the entitlements in which the children and family can engage with. • There are concerns that there is insufficient resource (both staff and budget) to improve monitoring systems.
3. Institutional Context		
Accessibility to and availability of stakeholders?	Are there physical security risks? Will weather be a constraint? Are staff and key stakeholders likely to be present, or absent on leave or secondment? Can reported availability be relied upon?	<ul style="list-style-type: none"> • Little or no risk involved in contacting stakeholders. • All stakeholders are easily contactable. • Those families with a JAFF/ TAF are easily contactable • Possible risks involved with locating and identifying all service users (if required). To do this would require local authority intervention.
Resources available to do the evaluation?	Time available in total and in country? Timing within the schedule of all other activities? Funding available for the relevant team and duration? People with the necessary skills available at this point?	<ul style="list-style-type: none"> • It is accepted that evaluation of the refocus Families First is currently undergoing will be required and therefore is likely. • Not known whether there is funding for an evaluation although it is accepted evaluation of the refocused programme will be necessary. • Previous evaluations have demonstrated a limited capacity for individuals with Welsh speaking skills by contractors.
Is the timing right?	Is there an opportunity for an evaluation to have an influence? Has the project accumulated enough implementation experience to enable useful lessons to be extracted? If the evaluation was planned in advance, is the evaluation still relevant?	<ul style="list-style-type: none"> • Evaluation evidence has been used to assist in refocussing the programme and new material would be used to inform ongoing work around refining and improving programme delivery. • Previous evaluations have provided information on the implementation, delivery and attempted impact analysis. • As Families First is currently undergoing a refocus it would be a beneficial time to consider evaluation to introduce any new data collection processes • Families First is currently going through a re-focus, removing any projects with a sole focus on employment and is currently in a transitional period. • Local variations in delivery reduce the influence of any findings as they may not always be applicable.

Coordination requirements?	How many other donors, government departments, or NGOs need to be or want to be involved? What forms of coordination are possible and/or required?	<ul style="list-style-type: none"> • Stakeholders are highly invested in the programme and have actively participated in previous evaluations. It is reasonable to expect their continued future involvement. • Ownership of the delivery of Families First Services, including JAFF/TAF and commissioned projects rests with the Local Authorities
4. Demands		
Who wants an evaluation?	Have the primary users been clearly identified? Can they be involved in defining the evaluation? Will they participate in an evaluation process?	<ul style="list-style-type: none"> • Internal Welsh Government staff <ul style="list-style-type: none"> ○ Involved in defining evaluation approaches • Welsh Government Ministers <ul style="list-style-type: none"> ○ Not involved in the evaluation process but will give approval to funds and approach • Local authorities <ul style="list-style-type: none"> ○ Likely to be willing to be involved in defining the evaluation approach. • Public service providers <ul style="list-style-type: none"> ○ Likely to be interested in findings of any evaluation and may be involved in the process • Non-governmental Organisations <ul style="list-style-type: none"> ○ Involved in project delivery in local authorities as part of the commissioned projects under Families First ○ Likely to be involved in any evaluation
What do stakeholders want to know?	What evaluation questions are of interest to whom? Are these realistic, given the project design and likely data availability? Can they be prioritised? How do people want to see the results used? Is this realistic?	<ul style="list-style-type: none"> • Internal Welsh Government staff/Welsh Ministers: <ul style="list-style-type: none"> ○ The effects/impact of the programme ○ Is it delivering against expected outcomes? ○ Can it be improved / refined? • Internal Welsh Government staff/local authorities: <ul style="list-style-type: none"> ○ If specific elements are working. ○ What is most successful? ○ What models are most successful? ○ What is least successful? ○ How the refocus of Families First has affected delivery?
		<ul style="list-style-type: none"> • Challenges given the lack of available and consistent data. Not always quantifiable. • The heavy reliance on qualitative data limits the quality of any evaluation of the programme's effects
What sort of evaluation process do stakeholders want?	What designs do stakeholders express interest in? Could these work, given the evaluation of the questions of interest and likely information availability, and resources available?	<ul style="list-style-type: none"> • Previous evaluations have been highly valued, although many qualitative in nature. • Qualitative data is feasible as does not require control groups and can give comprehensive information. • There is an appetite to demonstrate the impacts of the programme through an impact evaluation.
		<ul style="list-style-type: none"> • Qualitative data presents some challenges in drawing definitive conclusions about progress. • Outcome evaluation would require quantitative data. These designs with relation to Families First are less practical due to lack of available data and the inconsistency of programme delivery and lack of control group makes comparable data difficult to collect.

What ethical issues exist?	Are they known or knowable? Are they likely to be manageable? What constraints will they impose?	<ul style="list-style-type: none"> Ethical issues relating to families first (vulnerable groups, data transfer, data storage) all create constraints, but it is possible to overcome them through correct procedure and adherence to ethical guidelines and data protection legislation.
		<ul style="list-style-type: none"> There are a number of ethical issues that arise: <ul style="list-style-type: none"> Vulnerable groups – many of the beneficiaries are young children or vulnerable adults, such as being exposed to domestic abuse Data transfer – the large number of delivery organisations makes transfer of sensitive data a regular occurrence. Data storage – in order to evaluate Families First identifiable data will need to be collected and stored
What are the risks?	Will stakeholders be able to manage negative findings? Have previous evaluation experiences prejudiced stakeholder's likely participation?	<ul style="list-style-type: none"> Stakeholders have responded to previous evaluations in a constructive manner, and have used the evidence to make changes to the programme.
		<ul style="list-style-type: none"> Stakeholders have a very positive view of the programme and where evaluation outcomes don't reflect the perceptions this can be discouraging.

- 5.5 The outputs from the checklist are an assessment of the evaluability of Families First and the practicality and utility of an evaluation.. For Families First a number of issues relating to the information available and design of the programme make evaluating the programme difficult.
- 5.6 Some of the key positives to draw out include the programme design being based upon robust evidence and the role evaluation evidence has played in shaping the programme and making changes to it.
- 5.7 One of the key areas making the programme difficult to evaluate relates to the data available, in particular the inability to confirm what Families First services have been used by individuals. This means that evaluating the outcomes and impacts of the programme is not currently possible without additional data collection.
- 5.8 Additionally, the focus on population outcomes does not by itself produce evidence of the effects of the intervention, or how it might have had any effect. At best, movements in population indicators might suggest that something may have had some effect, almost always in combination with other factors (known and unknown), perhaps giving a hint of the possible size of any contribution - but such indicators cannot establish anything about how an intervention may have worked.
- 5.9 The absence of a realistic counterfactual group against which to compare those receiving Families First services presents a major barrier to robust evaluation. This

problem is enhanced because the implementation and goals of Families First vary by Local Authority.

- 5.10 It is difficult to assess whether Families First works. Trying to answer any questions about impacts and family outcomes of the Families First programme is a challenge, as currently it cannot be determined for certain what services families have used.

6. Currently Possible Approaches

- 6.1 A national evaluation of Families First was published five years ago. The waves of the evaluation are detailed below, summarising the main advantages of the work and its limitations

Methods	Advantages	Limitations
Families First: Pioneer Stage Review (2012) http://gov.wales/docs/dhss/publications/120924familiesen.pdf		
<ul style="list-style-type: none"> 89 face-to-face / telephone interviews / focus group discussions. 	<ul style="list-style-type: none"> Rich qualitative data on the implementation and transition from Cymorth to Families First in the pioneer areas Identifies benefits to the learning sets embedded within Families First 	<ul style="list-style-type: none"> Too early to provide accurate data regarding the level of need for families first services.
Evaluation of Families First: Year 1 Report (2013) http://gov.wales/docs/caecd/research/131219-national-evaluation-families-first-year-1-report-en.pdf		
<ul style="list-style-type: none"> In depth qualitative interviews and consultations with stakeholders Desk research including reviews of documentary evidence of local authority action plans, policy documents and population statistics 	<ul style="list-style-type: none"> Provides an assessment of the rationale for the programme and review of the progress made by Local Authorities over the first year of delivery Provides a baseline understanding of programme implementation Findings from year 1 cross referenced with a range of primary data collected in years 2 and 3 of evaluation 	<ul style="list-style-type: none"> Due to timing of evaluation with programmes implementation it was too early to assess the effectiveness of the different models being implemented across local authorities. Focuses on the perceptions of the design, implementation Does not address questions around impact and effectiveness
Evaluation of Families First: Year 2 Report (2014) http://gov.wales/docs/caecd/research/2014/140612-evaluation-families-first-year-2-en.pdf		
<ul style="list-style-type: none"> Stakeholder Survey Wave 1 Local authority and family case studies Desk research of population statistics 	<ul style="list-style-type: none"> Provides an assessment of early evidence of impacts and outcomes Identifies good practise 	<ul style="list-style-type: none"> Does not address questions around impact and effectiveness Case study families are more likely to be much more engaged with Families First than a typical family. Qualitative approach to programme is needed as no suitable comparison of local areas not running the programme is available

Evaluation of Families First: Year 3 Report (2015)

<http://gov.wales/docs/caecd/research/2015/150930-evaluation-families-first-year-3-en.pdf>

<ul style="list-style-type: none"> • Stakeholder survey and consultations • Distance travelled tools developed by evaluators – ‘Family Outcomes Tool [FOT]’ 	<ul style="list-style-type: none"> • Reviews Rationale for programme • Assessing process change • Provides a form of impact analysis on service users and population • Uses a Family Outcome tool to aggregate data captured by local authorities to provide an overall assessment of what proportion of families experiencing Families First have seen improved outcomes. • Many tools used in the FOT based on academically validated scales 	<ul style="list-style-type: none"> • Perceptions gathered from staff to assess the impact of Families First on families • Does not provide statistical data and unable to provide evidence of impact. • Due to the way the data was aggregated some local authorities did not contribute data towards particular domains because their distance travelled tools did not measure the relevant indicators for that domain.
---	---	---

6.2 The design of the programme has facilitated evaluation which have been predominantly qualitative in nature, precluding higher quality quantitative assessments

6.3 Families First changed the emphasis of local authorities’ monitoring away from measuring outputs to measures that consider the quality of delivery, in order to drive continuous improvements in services; and to consider family outcomes against agreed objectives, in order to achieve better, measurable outcomes for families. Families First adopted a set of monitoring approaches which were also intended to help in measuring progress:

- Results Based Accountability (RBA): An approach through which data is collected against pre-defined results and outcomes. This approach has been embedded in the programme action plan template which local authorities were required to complete as a requirement of funding and update in 2012.
- Report cards: A tool for capturing performance information (often updated quarterly) relating to commissioned projects in particular.
- Distance Travelled Tools (DTT): A framework for capturing family indicators and outcomes over the course of programme delivery and allowing for progress to be captured and updated.

6.4 Whilst RBA represents a useful tool in providing managerial evidence for accountability, it does not provide a substitute for evaluation on either a local or national level. Additionally, while Distance Travelled Tools are effective in giving families and practitioners an indication of the progress made by individual families, they are less valuable as a source of consistent monitoring data across all families and local authorities because the way they are applied is necessarily somewhat subjective..

6.5 The data available and the design of the Families First programme mean that some evaluation approaches are not possible, including Randomised Control Trials

(RCT), which are considered to be the most robust way to determine causation, which in this case would be the effects of the Families First services on recipient families.

6.6 In considering the methods employed in previous evaluations and those not possible, some other approaches could be considered for future evaluations. These are detailed below, with their applicability to Families First as it currently exists.

(a) Regression discontinuity analysis

6.7 This method of analysis requires a programme to include the whole population but to have continuous eligibility criteria, i.e. a cut off point on a scale, such as age or deprivation index. It then compares the outcomes of those who are only just eligible (i.e. just below the cut off) with those that are only just outside the eligibility criteria.

6.8 It is assumed that these two groups are similar in their characteristics, and so any difference between them in terms of their outcomes is due to the programme. This creates an estimate for the counterfactual by comparing very similar, but not identical groups. This method could be applied to Families First by comparing the outcomes of families who weren't considered eligible for Families First support to those who have received the support using the rainbow scale³

6.9 However, due to the nature of Families First those who just qualify for the Families First services are placed between 'prevention' and 'protection' and developing an arbitrary cut off to separate families into groups would be a difficult exercise.

(b) Matching Techniques

6.10 Matching techniques rely on using data held on individual programme participants and from those outside the programme. This identifies those included in the programme that have similar characteristics to those outside of the programme. It is assumed that if they have similar characteristics, any differences in their outcomes are due to engagement with the programme.

6.11 This approach can show whether there is a difference between groups, but not its causes. Only if there is very close matching between the groups can it be reasonably inferred that the programme may have caused the difference.

6.12 A large scale survey could collect the necessary data for application of this method. However, it would rely on those individuals using Families First services to be aware that the services they are using belong to the programme. As the programme aims to offer joined-up seamless support it is not always clear to families that the services they are using belong to Families First, unless they are completing a JAFF or a TAF plan.

(c) Difference in difference analysis

6.13 This method attempts to determine impact by comparing the change observed in the treatment group with that observed in a suitable comparison group. This method assumes that the differences, not related to the intervention, between the treatment and comparison groups remain the same over time. This provides an estimate for what would have happened in the absence of the intervention.

³ <http://gov.wales/docs/dsjlg/publications/cyp/111219-families-first-guide-en.pdf> - page 5

- 6.14 This method cannot be applied to give reliable estimate of the impact of Families First, as the differences in how Families First is delivered across local authority vary dramatically, and so it cannot be expected that the differences between them would remain constant. Also, the lack adequate 'baseline' data means that it is impossible to establish a 'pre-treatment' position.
- 6.15 Many of these methods have been considered in previous evaluations, and without changes to the data being collected about the families using the Families First services, future evaluations will only be able to replicate previous research. The questions which remain unanswerable without changes to the available data include;
- The level of take up of Families First services by individuals and families
 - How the level of engagement with Families First services influences outcomes
 - What the interactions are between the different Families First services
 - What the longer term impacts of Families First are
 - What the demographics of those with higher or lower levels of engagement with Families First services are.
 - Whether different individual characteristics are associated with different outcomes for those who engage with Families First services

7. Potential Future Approaches

- 7.1 The evaluability checklist clearly identifies three main challenges for evaluating Families First: lack of data about engagement with Families First services at the individual level, lack of a realistic counterfactual and the variation in programme delivery between Local Authorities. Without substantial changes to the programme or monitoring and data collection, there is nothing that can be done to address the issues presented by the latter two challenges. However, it is possible to start collecting additional data to support a more robust evaluation.
- 7.2 If data on individual use of Families First services were collected this would allow evaluations to answer questions on the effects of the programme. The data should include the service use on an individual level, along with demographic variables such as; age, family size, parent's employment status / qualifications and household size.
- 7.3 In addition to the use of Families First services, it would be useful to gather information to establish the individuals' use of other interventions such as Flying Start or Communities First. This would then provide an opportunity to better understand whether any changes observed are the result of Families First alone or more likely to be found where individuals have engaged with other interventions alongside Families First, or with no analogous interventions.
- 7.4 While individual data would allow for a range of research questions to be answered, collecting the data would not be easy. It is highly likely that it would require

substantial resources from Welsh Government and local authorities. Furthermore, gathering additional data from the individuals using the services of the programme will not overcome the issue of having no realistic counterfactual group.

- 7.5 Collecting this individual level data could be achieved either through a survey of the beneficiaries of the Families First programme or through building in additional data collection systems to the data already collected by local authorities from all participants of the programme. The gains and risks of employing either option are considered below.

1. Sample Survey

Gains

- Understanding of Families First engagement at a national level
- Snap-shot of numbers of participants for each of the services
- Potential to determine interactions between services with a large enough sample
- Longitudinal survey could track development and outcomes over time
- Cross-sectional survey able to compare outcomes for groups at different stages in the programme

Risks

- Need sample of sufficient size to provide reliable analysis
- Suitable comparison group needs to be identified
- Would need to be repeated to assess new different or future cohorts
- Poor response rate and/or attrition of participants⁴ (if longitudinal)
- Possible response bias from those most willing to engage with Families First
- May not be suitable to each authority due to variations in programme roll out

2. Individual level data collection

Gains

- Able to assess potential effects of Families First on an individual basis
- Detailed data on engagement with services of the Families First programme
- Up-to-date and ongoing information about exact numbers of participants for each of the entitlements
- Ability to determine interactions between services
- Linkable to other datasets which would allow for analysis of wider outcomes
- Can track participants through the programme and compare those at different stages of the programme

⁴ Longitudinal surveys attempt to survey the same participants over a range of time points, often several years. However, some of these participants may choose to drop out of the research before the survey is concluded. The attrition rate is the proportion of those participants that do not complete the whole survey.

Risks

- Data Protection and sensitivity of items collected. Consent to share the data would need to be gained
- Data management and storage tools need to be considered and developed
- The process of collation would need to be consistent across Local Authorities
- Validation of the data and practicality of the actual collection
- Data collected by non-researchers, possibly leading to systematic biases
- Some local areas may lack the systems necessary to collect individual level data, at least in the short term.

7.6 It may be a challenge to implement the necessary systems for individual data collection, and use of a survey may not provide the necessary data. Additionally, any new data collection system would need to consider how each local authority is implementing the Families First programme and allow for a standardised means of collecting data that each local authority could use.

7.7 There are substantial gains to be made and risks if the data continues to not be collected, without knowing the true effects of the programme, it is not known if it is good value for money. In order to determine where possible improvements can be made, there needs to be full comprehension of the effects of the services through gathering individual level data.

8. Conclusions

8.1. The aim of this report was to determine if:

- It is plausible to expect impacts
- It is feasible to measure impacts
- Whether an impact assessment would be useful

8.2. Families First was developed and continues to develop in line with the best available evidence, which suggests it is plausible to expect beneficial effects for families. However, it is unclear what these impacts necessarily should be, due to the lack of data of those using the services provided as part of Families First.

8.3. Due to this lack of data, previous evaluations have adopted predominately qualitative methods and have been subject to constraints in terms of design of the programme and the delivery of the programme in each local authority.

8.4. Additionally, the way in which Families First was rolled out across Wales means a counterfactual group can not be currently identified. There is no easy remedy for this, due to the unique roll out of Families First in each local authority and the bespoke family tailored delivery of Families First services to each family. However, additional data collection should propose the opportunity to potentially identify comparison groups.

8.5. In order to demonstrate impacts of the programme and for any outcome evaluation to be useful, additional collection of data is required at an individual level. Without this information, currently, it is possible to determine some limited outcomes.

It is possible to reach findings about:

- Stakeholders attitudes and perceptions of the Families First programme
- How the programme is implemented and the quality of the implementation
- Whether the programme design is fit for purpose

It is not currently possible to assess:

- The impact of Families First services on individuals and families
- The specific effects of Families First services, or their interactions with other interventions
- The actual take-up of services by individual families, and any groups that have higher levels of take-up.

Recommendations

8.6. This report makes the following recommendations for any future evaluations of the programme in its current form:

- To continue to use qualitative research to understand the perceptions of the families using Families First services
- To identify a sample of families/parents/children who have engaged with Families First services and track these families longitudinally

The possible improvements to the programme that could lead to more robust evaluations:

- Develop new systems for collecting individual level data about service use for the Families using Families First, rolled out as standard across all local authorities
- Establish where families are receiving additional services or interventions and link this to the individual level data from Families First

If these improvements were to be made, the following activities may then be possible:

- Analyse the individual level data to test whether individual characteristics or wider contextual factors affect impact of services.
- Identify, where additional services or interventions are being used, whether these have an additional effect on those using the Families First services.