

KEEP THEM SAFE

A shared approach
to child wellbeing

Synthesis of Project Evaluations



Prepared by the KTS Evaluation Steering Committee
on behalf of the KTS Senior Officers Group

Keep Them Safe Interim Review, *Synthesis of Project Evaluations* published by:
NSW Department of Premier and Cabinet
GPO Box 5341
Sydney NSW 2001

<http://www.keepthemsafe.nsw.gov.au/>

© State of New South Wales, 2013. This work may be reproduced in whole or in part for study or training purposes subject to the inclusion of acknowledgment to the NSW Government. You must obtain permission if you wish to use this work for other purposes.

Contents

1.	Executive Summary.....	1
1.1	Design and Implementation	1
1.2	Evaluation.....	2
2.	Introduction	4
2.1	Purpose.....	4
2.2	Scope and Alignment.....	5
2.3	Selection	6
2.4	Questions and Limitations	7
2.5	Overview of Individual Initiatives and Evaluations.....	8
3.	Discussion.....	13
3.1	Interim Review Questions	13
3.2	Program Design and Implementation	16
3.3	Program Evaluation	17
3.4	Conclusion	18
4.	Appendix 1 – Analysis of Evaluations	19
4.1	Family Case Management	19
4.2	Family Referral Services.....	24
4.3	Child Wellbeing Units	28
4.4	Whole Family Teams	31
4.5	Alternative Dispute Resolution.....	35
4.6	Family Group Conferencing	40
4.7	Screening and Response Priority Tools	44
4.8	Safe Families.....	51
5.	Appendix 2 – Status of KTS Evaluations	58
6.	Appendix 3 - References	59

Tables

Table 1 – Relationship between Interim Review Questions and Evaluations	5
Table 2 – Progress Piloting Family Case Management	21
Table 3 – Percentage Agreement Between Helpline and CRC Staff by Case Reading Measure	49
Table 4 – Comparative Case Reading Results Across Time	49
Table 5 – Status of KTS Evaluations	58

Abbreviations

ABS	Australian Bureau of Statistics
AEDI	Australian Early Development Index
CSC	FaCS Community Service Centre
CWU	Child Wellbeing Unit
DAGJ	Department of Attorney-General and Justice
DADHC	Department of Aging, Disability and Home Care
DEC	Department of Education and Communities
DPC	Department of Premier and Cabinet
FACS	Department of Family and Community Services
FRS	Family Referral Service(s)
Health	NSW Ministry of Health
KTS	Keep Them Safe
LHD	Local Health District
MoU	Memorandum of Understanding
ROSH	Risk of Significant Harm
ScRPT	Screening and Response Priority Tools
SNF	Sustaining New South Wales Families (a.k.a. Sustained Health Home Visiting)

1. *Executive Summary*

This document presents a synthesis of evaluations and reviews of eight initiatives implemented under Keep Them Safe (KTS), the NSW Government's five-year plan to reform child protection services. All these reviews were conducted relatively early in the implementation of each initiative. As a result, they predominantly focus on the implementation processes itself, and it is in this area that the findings of this synthesis are most comprehensive. Where evaluations were able to examine the impact of these initiatives on outcomes for children, young people and families, these are also discussed here (although caveats regarding robustness of the findings should be heeded).

This synthesis forms part of the KTS Interim Review, and contributes to that broader project in two ways. First, it identifies possible trends regarding project implementation in a range of geographic, agency and intervention contexts. This synthesis does not, however, aim to draw conclusions about KTS as a whole. The evaluations reviewed represent only a small number of the initiatives implemented under KTS. Universal services are not covered, and coverage of secondary and tertiary services is also limited. Additionally, many of the evaluations considered here were completed well before the synthesis was undertaken, which means this document should not be read as reflecting the current status of these projects. The findings of the evaluations relate to the time at which they were reported, and it is to be expected that the agencies responsible will have acted on these findings and made efforts to improve the implementation and impact of their programs.

Second, this synthesis makes a number of observations about program evaluation under KTS. It covers issues such as evaluation design and implementation, and considers early indicators regarding likely achievement of project goals in the longer term.

1.1 *Design and Implementation*

The KTS Action Plan proposed a range of initiatives which were generally evidence-based, but which often constituted new models of service delivery. These new initiatives were implemented in a pressured policy context and under tight timeframes. Despite this complexity, the evaluations considered here suggest that agencies have implemented key components of KTS well.

At the same time, the evaluations suggest that there are some areas in need of attention. The unifying theme among most of these concerns the absence of clarity in the early stages, although this manifests itself in different ways.

- **Clarity of Purpose.** Several evaluations suggested greater clarity might have been achieved in the development phase of many programs. In particular, there was a tendency to distinguish insufficiently between the formal reasons for establishing an initiative (usually because the Wood Report recommended its establishment) and the substantive outcome the initiative was intended to achieve for client groups. This is important because implementation often brings unexpected issues that must be addressed in the process of project delivery, and these are more easily addressed in the light of clearly-stated substantive goals.
- **Clarity of Responsibility.** Many evaluations testify to the challenges which arise from insufficiently clear allocation of responsibility among multiple agencies and stakeholder groups.

- **Clarity concerning Resource Requirements.** Several evaluations found that good collaboration is resource intensive, The staffing commitment and time required for administration associated with joint working often seems to be underestimated during planning.
- **Clarity of Business Information and Reporting.** Several evaluations found that implementation was often not supported by adequate information collection and systems.
- **Systemic issues concerning workforce capacity and service gaps.** Many government agencies and non-government service providers are operating at capacity, and this has hampered the effectiveness of referral services. On a similar note, the rollout of some service delivery initiatives was hampered by difficulties in recruiting sufficient trained staff.

1.1.1 Impact

Although none of these evaluations was primarily designed to examine whether their initiatives were having their intended impact on children, young people or families, many were able to gather some very early indications of possible success. This evidence must be treated as extremely provisional, because it was often gathered at a very early stage in implementation, and is often based on a small number of cases. Nevertheless, by comparing this evidence with the goals of these initiatives and evidence from academic literature and similar projects in other jurisdictions, it is often possible to reach tentative conclusions about whether initiatives have been implemented well have a reasonable chance of meeting their overall goals.

Most of the initiatives reviewed in this synthesis appear to be well-placed to achieve their aims, despite some implementation issues. Qualitative feedback from the small number of clients who could be interviewed generally suggest high levels of satisfaction and a belief that the initiative would prove beneficial. Realising these benefits more broadly is, however, contingent upon key lessons highlighted in the evaluation being heeded, and recommendations implemented. Many of these recommendations relate to the implementation issues noted above.

1.2 Evaluation

This synthesis shows that the NSW commitment to evaluate KTS has been fulfilled, although evaluations have sometimes only been able to reach limited conclusions. It was always likely that evaluation of KTS would be challenging: rigorous, robust, methodologically-sound evaluations of complex initiatives are difficult to conduct – especially when evaluation fieldwork commences while the initiatives are still being rolled out, as was the case for several evaluations considered here. Evaluation is even more challenging for initiatives which, like many under KTS, seek to effect system change or improve outcomes for clients in ways that cannot easily be quantified. Despite these challenges, there are some things agencies could have done to improve the chances of success, and there are some options available for improving the quality of those KTS evaluations which have not yet been conducted.

- **Planning for evaluation was often started too late.** Many KTS programs were implemented without sufficient consideration for how they would eventually be evaluated. Measures of outcomes and outputs were not always agreed, and baseline

data were not always collected before implementation began. It is often difficult to design a sound evaluation in this situation.

- **Evaluation was often concluded too early.** Many of the evaluations considered here were unable to gather sound evidence of improved processes or early impact on outcomes because the initiatives in question were still being set up, or had not dealt with enough clients, or had not done so for long enough to have had much effect. This is not a serious issue with process evaluations, which are primarily concerned with whether implementation has proceeded according to plan. But where evidence of effects or outcomes are sought, then sufficient data to provide statistically meaningful results is required.
- **Data quality is often poor.** Several evaluations have found that administrative data of the kind required to review or evaluate initiatives are either not collected, not reliable, or not readily available to evaluators. Poor data quality is also sometimes the result of evaluators attempting to overcome problems of insufficient data through reliance on techniques which are known to be less robust (e.g. qualitative interviews involving relatively small numbers of participants selected using methods that are acknowledged to be potential sources of bias).
- **Mismatch of expectations.** It appears that there were expectations of evaluations to provide greater levels of evidence than the project conditions or the data could deliver. As a result, the evaluations appear to over-promise but under-deliver.

2. Introduction

This document presents the findings from reviews/evaluations of eight KTS initiatives. It describes the goals of the synthesis, and then discusses project implementation and evaluation under KTS. It draws conclusions regarding project implementation quality, likely project effects (where evidence for this is presented in the relevant evaluations) and lessons regarding evaluation more generally. These initiatives, in the order they are addressed, are:

- Family Case Management
- Family Referral Services
- Child Wellbeing Units
- Whole Family Teams
- Alternative Dispute Resolution – Legal Aid Pilot and Dispute Resolution Conferences
- Alternative Dispute Resolution – Family Group Conferencing
- Screening and Response Priority Tools
- Safe Families/Focus Communities

A summary of these projects and their evaluations is given in section 2.5 below. Detailed analysis is presented in Appendix 1 – Analysis of Evaluations starting on page 19. It should be noted that these evaluations were synthesised based on evaluation documents provided by the line agencies responsible for implementation and evaluation. No independent verification of their contents was undertaken. It should also be noted that the purpose of this document is to synthesise their conclusions. For critical analysis of these claims in light of the other available evidence and the broader context of KTS, readers are advised to consult the reports of the other components of the Interim Review, and in particular the Interim Review Report.¹

2.1 Purpose

This synthesis forms part of the KTS Interim Review,² and contributes to that broader project in two ways. First, it informs the ongoing rollout of KTS by examining a small number of strategically important initiatives which have progressed far enough to be evaluated. It is likely that the challenges shared by the programs examined here are impacting on other initiatives as well, and lessons drawn from them may help in the implementation of others – and, indeed, of future whole-of-government initiatives.

Second, this synthesis helps lay the groundwork for the June 2014 Strategic Impact and Outcomes Evaluation. The evaluations it considers are predominantly ‘process’ evaluations, which is to say they focus on whether each program has been implemented as intended, and whether it is reasonable to expect them to have their intended impacts at some point in future. This provides a basis for future evaluations to consider the impact of implementation on the impact of the project. This synthesis also lays the groundwork by establishing whether the conditions are in place for assessing the effects of KTS. It examines the

¹KTS Evaluation Steering Committee (2013). *Interim Review. Report* ["Interim Review"] Department of Premier and Cabinet, Sydney. Retrieved on 15 May 2013 from http://www.dpc.nsw.gov.au/prem/documents/keep_them_safe/publications_non-automated/Report_of_the_Interim_Review.pdf

²Information on the origins, conduct and governance of the Interim Review can be found in KTS Evaluation Steering Committee (2011). *Keep Them Safe: A shared approach to child wellbeing. Interim Review Plan* ["Interim Review Plan"] Department of Premier and Cabinet, Sydney. Retrieved on 1 May 2012 from http://www.dpc.nsw.gov.au/data/assets/pdf_file/0003/137406/12-04-12_-_Interim_Review_Plan.pdf.

challenges agencies have faced in conducting rigorous evaluations, and the quality of the evidence available, with the goal of enabling remedial action to ensure the final evaluation is successful.

2.2 Scope and Alignment

As a component of the KTS Interim Review, this Synthesis contributes answers to some of the overarching Interim Review questions. Table 1, below, lists the evaluations included in the synthesis and shows the questions to which they provide some response. The table shows that the evidence allows the Synthesis to speak to whether children's court processes have improved (Interim Review question 5), to changes in the degree of information sharing and collaboration amongst agencies (question 6), to whether new systems are effectively linking families to services (question 7) and to whether services have increased the levels of cultural awareness, cultural competence and partnerships with Aboriginal communities (question 9). It also shows that this synthesis speaks, to some degree, to the questions of whether early intervention services have been enhanced (question 2), whether services are more available and better able to meet the needs of Aboriginal children, young people and families (question 8) and the state of partnerships between government and non-government organisations (question 10).

Table 1 – Relationship between Interim Review Questions and Evaluations

	Family Case Management	Family Referral Services	Child Wellbeing Units	KTS – Whole Family Teams	Alternative Dispute Resolution	Family Group Conferencing	Screening and Response Priority Tools	Safe Families and Focus Communities
1. To what degree is the universal service system stronger and more extensive?								
2. Have early intervention services been enhanced?	•	•	•					
3. Are early intervention services supporting children and parents in the community?	•	•	•					
4. Are workers engaging with the Mandatory Reporter Guide and the new reporting threshold so that the statutory child protection system is more streamlined and focussed on children at greatest risk?			•				•	
5. Have out-of-home care and children's court processes improved?					✓	✓		
6. To what degree is there coordination and information sharing amongst agencies?	✓	✓	✓	•			✓	

7. Do new systems effectively link families to services?	✓	✓	•				✓	
8. Are services more available and better able to meet the needs of Aboriginal children, young people and families?	✓	✓			✓	✓		•
9. Have services increased the levels of cultural awareness, cultural competence and partnerships with Aboriginal communities?	✓	✓			✓	✓		✓
10. What is the current state of partnerships between government and non-government services?	•	•	•			•		
11. Have government and non-government services engaged positively in the process of workforce culture change?	•	•	•		•	•		

Key: ✓ = Directly Relevant; • = Somewhat Relevant

These evaluations also speak to the broader context within which the Interim Review is taking place. They cover five of the six substantive elements of the KTS Action Plan,³ albeit with a bias towards the acute end of the service spectrum. They speak to the Action Plan's concern with better practice and systems (Element 4), early intervention services (Element 2, although only partly), protecting children (Element 3) and addressing the needs of Aboriginal people (Element 5). By extension, these initiatives cover many important elements of the Wood Report's vision for reforming child protection in NSW. Child Wellbeing Units, Family Referral Services, Family Case Management, the Structured Decision Making tools used by the Child Protection Helpline, and Alternative Dispute Resolution were all recommended explicitly in the Wood Report. Whole Family Teams (WFT), Family Group Conferencing and Safe Families/Focus Communities are, for their part, representative of the broad goal of fostering a collaborative, family-centric service system. The only aspect of the service system not covered is the universal system. This is primarily because most of the initiatives the Action Plan identified as "universal" involved enhancement funding to programs which had already demonstrated their worth through prior evaluation.

2.3 Selection

Evaluations have been included in this synthesis primarily on the grounds that they are KTS-specific activities and that they have a completed evaluation. The Action Plan committed the NSW Government to systematically monitoring implementation to ensure actions were delivered and outcomes achieved.⁴ This synthesis is a testament to that ongoing process. While there are a number of completed evaluations, evaluations are scheduled for many other KTS initiatives, with timelines for initiation and completion shaped by the nature of each initiative.

There are some evaluations of KTS-related activities that have not been included here. There are specific reasons for this. A small number of projects such as New Street, Hey Dad! Indigenous Dads, Uncles and Pops, and Mothering at a Distance received enhancement

³Department of Premier and Cabinet (2010). *Keep Them Safe: A shared approach to child wellbeing* ["Action Plan"] Department of Premier and Cabinet, Sydney, p. iv. Retrieved on 1 May 2012 from http://www.dpc.nsw.gov.au/data/assets/pdf_file/0004/57145/Keep_Them_Safe.pdf.

⁴*Action Plan*, pp. 49-51.

funding under KTS on the strength of pre-existing evaluations. Their service models and implementation were established prior to KTS, which supported extensions in their activities. In the case of Brighter Futures, this project was reconfigured during the course of KTS implementation on the basis of evaluations which were already well underway. While these initiatives are relevant to the outcomes of KTS, they are not included here because there is little scope for undertaking significant remedial work during rollout (as they are already mature programs). The impact of these programs is taken into account in other aspects of the Interim Review, above all the Population Outcome Indicators.

2.4 Questions and Limitations

This synthesis contributes answers to these Interim Review questions by answering the following five synthesis questions for each initiative:

1. What did the project set out to do (i.e. what was the initiative supposed to achieve and why)?
2. What was done (i.e. how was the initiative actually implemented, and how closely did implementation follow the original plan)?
3. How was it evaluated?
4. What did the evaluation find (i.e. using the available indicators, what do we know about the initiative itself)?
5. How sound are the findings?

The answers to these questions are based on the evaluation reports provided by the agencies responsible for the individual project and its evaluation. The synthesis has undertaken no further review, and the accuracy of synthesis is contingent upon the accuracy of the underlying documents. The evaluation reports have, however, been read with a critical eye to consider the nature of the evaluation method and the reliability of findings.

It must be emphasised that there are limits to the general applicability of the insights that can be drawn from examination of a small number of cases such as this. These are discussed later in the document. For now, it is sufficient to note that these limitations do not negate the value of this synthesis. The descriptive and analytic detail it is possible to gather from an examination of a smaller number of cases offsets the lack of inherent portability in its conclusions. Moreover, the programs which *are* covered here are significant in their own right and are therefore likely to illuminate fundamental challenges to the implementation and evaluation of KTS, while expanding the scope to include other evaluations would be impractical in any case.

2.5 Overview of Individual Initiatives and Evaluations

This section discusses the eight individual KTS initiatives which this report covers. As noted earlier, the findings predominantly relate to implementation quality, though some early indications of effects are reported where appropriate.

2.5.1 *Family Case Management*

Family Case Management (FCM) provides integrated case management for families who are frequently encountered by child protection and other social service systems. It was initiated as a trial in eight local areas across three NSW regions.

The evaluation was undertaken by ARTD in two stages. The first stage focused on implementation processes and interim results with an intention of informing future rollout of the project. The second stage was oriented towards results, economic analysis, and a definition and sizing analysis.

Implementation was found to be generally consistent with the project plan in two of the three geographic areas, although pressures to implement the program quickly impacted on stability and clarity for service personnel, and engagement of suitable families for the services. Staff commitment to the model and effective case coordinators were integral to the effectiveness of local service delivery. The available evidence, which was limited due to a small evaluation sample, suggests that families experienced the services positively and that some families benefited from this integrated approach. The service model was identified as likely to be cost-effective over the long term.

2.5.2 *Family Referral Services*

Family Referral Services (FRS) link families facing difficulties but who do not meet the threshold for mandatory reporting to Community Services to services in their area. Two different models of FRS were trialled in three different locations. Following the initial trial and identification of the preferred service model, rollout across NSW is underway.

The evaluation of the Family Referral Services was undertaken by Ernst & Young, using a range of methods including service use data, surveys with mandatory reporters and service users, and interviews with FRS staff. An initial, quantitative process evaluation was undertaken to review the first nine months of operation and provide recommendations regarding the optimal service model. The second strategy involved qualitative methods to assist to understand the experiences of service users and the effects of FRS on the service system.

Family Referral Services were found to be achieving the aims of; improving family access to local service systems, especially for vulnerable population groups and Aboriginal clients; and connecting service providers with other services, such as CWU's to local organisations - although it must be emphasised that these findings were based on interviews with a very small number of clients. At the same time, the local service system capacities impacted upon options for FRS referral and handover. Through improving links and informing local service system realignment, FRS services were making positive effects on local service systems. The 'augmented' service model was recommended for future FRS rollout, rather

than the telephone-based model, although flexibility would be required to accommodate local conditions.

2.5.3 Child Wellbeing Units

Child Wellbeing Units (CWUs) advise, support and educate mandatory reporters in human services agencies, and help to decide courses of action when a worker has concerns regarding the safety, welfare and wellbeing of a child or young person. CWUs provide a point for consultation on child protection matters prior to a worker contacting the Community Services Child Protection Helpline, and a source of advice on alternatives where making a report is not called for.

The Nous Group conducted a review of the implementation of the CWUs between April and July 2011. This relied on multiple methods, including a survey of CWU staff and mandatory reporters, a review of operational data, consultations with stakeholders and structured workshops. The purpose of this exercise was to review implementation of the CWUs rather than to assess their impact on the service system or children, young people or their families. This was due to the early timing of the evaluation and the unlikelihood of firm data being available to measure the success. It does however give an indication of the progress of the workforce change. An outcomes evaluation of the CWUs is currently being prepared, and is expected to be conducted in 2013.

Child Wellbeing Units were found to be effective in achieving the aims of supporting their target group, and supporting culture change around shared responsibility for child safety, welfare and wellbeing. The project was generally well implemented, and the evaluators advised that the model should be continued, albeit with particular attention to information systems and stability of staffing to ensure more effective delivery in the longer term.

2.5.4 Whole Family Teams

Keep Them Safe Whole Family Teams work with families where there are presenting drug and alcohol and mental health issues that are impacting on children. There are four pilot WFT services located within Ministry of Health. These services work with families at the tertiary end of child protection spectrum where the family is being case managed by Community Services.

The evaluation of WFT is being undertaken by Urbis, examining both process and outcomes. To date, two evaluative exercises have been completed and reported – Year 1 and Year 2 reports. The first report recommended improvements to internal structures for information sharing and communication, attention to data systems and evidence, and development of an evaluation advisory group. These recommendations were actioned and the positive effects reported in the Year 2 evaluation.

The Year 2 evaluation found that the WFT services were generally stable and becoming firmly established at the local level with effective working relationships and referral processes with Community Services. At the same time, there continue to be pressing issues with data quality and reporting, as well as numerous small differences in service implementation between locations. An important example of these differences is the way in which children are engaged (either directly or indirectly) in each location.

2.5.5 *Alternative Dispute Resolution*

The Legal Aid Pilot (LAP) and Dispute Resolution Conferences (DRCs) offer forms of Alternative Dispute Resolution (ADR)⁵ that provide less formal alternatives to hearings for resolving disputes between families and Community Services that reach the Children's Court. Cases must be referred for ADR by the Magistrate or Children's Registrar responsible for the case. The LAP was based on the model used in Family Court matters. It was undertaken only in Bidura Children's Court in Glebe, and the mediations in this pilot were facilitated by external mediators. DRCs were implemented across NSW and were facilitated by Children's Registrars.

The Australian Institute of Criminology (AIC) undertook the evaluation of these aspects of ADR, as well as the trial of Family Group Conferencing described below. These evaluations focused on both process and outcomes, and involved a literature review, post-conference surveys, observations of conferences, interviews with families and key stakeholders, surveys with legal officers and department staff, and review of quantitative data.

The LAP and the DRCs involved models that had previously been tested and refined; the evaluation found that they had been implemented as intended and adequately resourced. They were shown to be effective in improving participant experience of resolving disputed issues in the Children's Court. They were also shown to improve processes and practice standards for facilitators and mediators improved during the course of the trial. There was high level support for the trials, and effective information flow and promotional material assisted with stakeholder engagement. It was recommended that, if both forms of ADR were to continue to be delivered in future, clear guidelines that allow an assessment of the suitability of matters for either program should be developed..

2.5.6 *Family Group Conferencing*

Family Group Conferencing (FGC) is the third component of ADR implemented under KTS. It is discussed separately here because it is undertaken prior to Children's Court proceedings under the aegis of Community Services. It aims to achieve a family-developed plan to address issues that Community Services' staff have identified as impacting negatively on the children or young people involved. These plans are not legally binding. It was trialled in two FACS regions – one city (Metro-Central) and one rural (Northern).

The Australian Institute of Criminology evaluated the FGC project using a range of qualitative and quantitative methods similar to those used for the evaluation of the other aspects of ADR. For the evaluation of FGC the preparation of case summaries was added.

The evaluation of Family Group Conferencing found that while the model was robust and the facilitators skilled and available, there were only a limited number of conferences conducted during the trial period. This appeared to be due to restrictive criteria, limited promotion of the program and lack of internal support from key staff. For the conferences that were conducted and evaluated, all parties (family members, professional stakeholders, facilitators) reported satisfaction with the conference process. Successful conferences were the result of skilled facilitators and thorough preparation. One of the key outputs – Family Plans – were developed in almost all conferences, and from the limited available data there

⁵ In this report, ADR is used to refer to the Legal Aid pilot, Dispute Resolution Conferences, and Family Group Conferencing.

were some indicators that actions from these plans were implemented. The evaluators found that follow-up from conferences and monitoring and review of Family Plans needs to be improved.

2.5.7 *Screening and Response Priority Tools*

Screening and Response Priority Tools (SCRPT) were implemented at the Community Services Child Protection Helpline to assist caseworkers to make more consistent and accurate decisions regarding cases that should be 'screened in' for a child protection response or 'screened out' with potential for follow-up at the local level. The Response Priority tool, as the name suggests, assists caseworkers to allocate specific response times to the reports that they have screened in. These tools were purchased from the Children's Research Centre (CRC) in the United States, but extensively modified for local conditions.

The evaluation of these tools was undertaken using two approaches. Firstly, Community Services reviewed child protection reporting data as well as staff satisfaction with the tools. Secondly, the CRC undertook structured reviews to monitor implementation quality, and undertook other quality review strategies across a two year period.

The evaluation found that there had been a clear reduction in the volume of reports received by Community Services with the greatest decrease among the reports that required the least rapid response. This was a key aim of raising the reporting threshold to Risk of Significant Harm, and implementation of the Structured Decision Making tools of which SCRPT form a part. Helpline staff implementing the SDM Screening Tool and Response Priority Tool found them to be beneficial and supportive of their professional practice. At the first review of implementation quality there was a satisfactory degree of consistency in the identification of issues by staff using the Screening Tool, but improvement needed to be made in using the Response Priority Tool. By the fourth round of the review processes the level of consistency between Community Services' staff use of the tools and the CRC's use was at the benchmark levels of 90 to 95% across most review categories.

2.5.8 *Safe Families*

Safe Families⁶ is an element of the *NSW Interagency Plan to Tackle Child Sexual Assault In Aboriginal Communities 2006-2011*, and was supported under KTS. Safe Families included a range of case management and community development activities to address vulnerabilities in families as well as building strong and safe communities. These were provided by teams of co-located agency partners – Health, FACS, Police and Aboriginal Affairs.

The evaluation of this initiative (as well as Focus Communities) was undertaken by Kristine Battye Consulting. The evaluators used a combination of methods including a literature review, focus groups, interviews and generating quantitative data. These methods were appropriate for researching with Aboriginal people.

The evaluation found that the project plan was consistent with existing literature, and that Safe Families could have been a culturally appropriate and effective public health strategy. It also found that there was some evidence of formal case management and other engagement activities at a local level. But it identified numerous implementation issues

⁶ Safe Families is linked to the *Focus Communities* initiative, but as the latter did not receive KTS funding it has not been included in this review.

which impacted upon the potential to achieve the majority of the goals. Lack of data regarding the identified indicators meant that few benefits of the program could be measured. While belief in the need for a collaborative approach to the issue was uniformly reported by staff implementing the project, only localised incidents of positive interagency relationships were identified. These appeared to be supported by positive leadership and collaborative approach to the issue.

3. Discussion

This section draws some general observations from an analysis of individual program reviews (see Section 4, below). It focuses on three issues: the answers which these evaluations offer to the Interim Review questions; the successes and challenges they share, and which may also be affecting other KTS initiatives; and what these evaluations tell us about progress in implementing the government's commitment to comprehensive evaluation. Given the diagnostic goal of the Interim Review, this section deliberately devotes more attention to the challenges than the successes.

It must be emphasised, once again, that this synthesis is based on an examination of too few initiatives to be able to draw conclusions about KTS as a whole. It is not possible, for example, to make any comment on progress in the universal services space, and any comments about relative performance in the secondary and tertiary spaces can only be made with a great deal of circumspection (even though within each of these spaces, there appear to have been both successes, and clear lessons for future implementation and evaluation). Moreover, the evaluations analysed here were mostly completed well before this document was drafted. As a result, it is possible that many of the shortcomings and challenges identified here have already been the focus of remedial action.

3.1 Interim Review Questions

Section 2.2 above indicated that this synthesis would seek to contribute to the Interim Review's analysis of seven Interim Review questions, and that it might also be able to contribute indirectly to three others. Each of these questions is addressed in turn below.

3.1.1 *Have early intervention services been enhanced?*

The initiatives in this synthesis which related to early intervention services were Family Case Management and Family Referral Services. The very existence of these two new initiatives provided more capacity in the early intervention space. These were both identified as being implemented effectively, though each had areas that could be improved. The evaluators of both service models found reason to support their further rollout.

3.1.2 *Are early intervention services supporting children and parents in the community?*

The evaluations of both Family Case Management and Family Referral Services provided initial evidence that these services were having positive effects for some families and for some service user groups. This must be taken as an indicative finding due to the early stage at which the evaluations were conducted, the low numbers of service users with which the evaluations could engage, and the lack of evidence of effect over time.

3.1.3 *Are workers engaging with the Mandatory Reporter Guide and the new reporting threshold so that the statutory child protection system is more streamlined and focused on children at greatest risk?*

The ScRPT evaluation did not directly address engagement with the new threshold or the Mandatory Reporter Guide (MRG), but data contained in associated early reports show that immediately subsequent to the introduction of the threshold, the MRG and new systems such as CWUs, the volume of reports to the Helpline dropped as intended. These data did not, though, allow an investigation of the attitudes of mandatory reporters, or of the role of

the MRG or other Structured Decision Making (SDM) tools. Other evaluation components address this question more directly.

3.1.4 *Have out-of-home care and children's court processes improved?*

Children's Court processes, including for those children and young people who are in the Care of the Minister and subject to a review under Section 90 of the *Children and Young Persons (Care and Protection) Act, 1998*, were the specific domain of the ADR processes evaluated by Australian Institute of Criminology (AIC). The three components of ADR that were evaluated and included in this synthesis – Legal Aid Pilot (LAP), Dispute Resolution Conferences (DRC), and Family Group Conferencing (FGC) – were experienced positively by the parties involved. In addition, initial indicators of success – resolution or narrowing of issues in dispute, or development of a Family Plan – were identified as positive initial achievements of the conferences. These were made possible by sufficient resourcing for the duration of the pilot, and having skilled facilitators (including Children's Registrars) to conduct the meeting processes. Again, there were aspects that could be improved, with the post-conference activities and the meeting review responsibilities being particularly noted for FGC. An important point regarding the ADR project overall is that caseworkers, managers and legal representatives needed to contribute more to the cultural shift that is required if ADR is going to consistently achieve beneficial outcomes for families.

3.1.5 *To what degree is there coordination and information sharing amongst agencies?*

Several aspects of KTS seek to improve coordination and information sharing amongst agencies. The main initiatives considered here are the Child Wellbeing Units (CWUs) and Family Case Management (FCM), and to a lesser degree Family Referral Services. Other important initiatives not considered here include legislative reform to remove restrictions on information sharing (the so-called "Chapter 16A provisions", which are considered in other components of the Interim Review).⁷ In the case of the CWUs, the systematic sharing of information across agencies via the WellNet database was identified as helping CWU staff to provide more coordinated and informed advice to mandatory reporters.

Among the direct service delivery initiatives, FCM's emphasis on multi-agency collaboration to address the needs of family members and reduce issues impacting on children and young people was also identified as contributing to positive outcomes. Here, and also in the case of Whole Family Teams, the role of skilled and passionate local coordinators and the development of positive working relationships among different stakeholders were identified as crucial contributors. In the case of WFTs, the positive relationship between the Health-based WFT and Community Services has improved quality and accuracy of referrals, and stability in the working relationship.

⁷See, for example, *Interim Review*; KPMG (2013). *Keep Them Safe Location Based Evaluation. Report* ["Location Based Evaluation"] Department of Premier and Cabinet, Sydney. Retrieved on 15 May 2013 from http://www.dpc.nsw.gov.au/prem/documents/keep_them_safe/publications_non-automated/Location_Based_Evaluation.pdf; KPMG (2012). *Keep Them Safe Workforce Survey. Final Report* ["Workforce Survey"] Department of Premier and Cabinet, Sydney. Retrieved on 4 December 2012 from http://www.dpc.nsw.gov.au/data/assets/pdf_file/0006/146670/KTS_WS_-_Final_Report_no_appendices.pdf

3.1.6 Do new systems effectively link families to services?

Of the project evaluations included in this synthesis the question of effectiveness with which families are linked to services is addressed most directly by FRS. The three pilot services were found to have received over 1,100 referrals from various different sources and then made more than 2,000 referrals out to local services. While the evaluation did not have the capacity to follow families and assess the uptake or outcomes of the service referrals, that the FRS was evaluated as identifying support needs of families and linking those families to support services is a positive indicator for achieving the higher level goal.

The second evaluated project which has an important role in supporting the links between families and services is the CWU. Where a mandatory reporter contacts the CWU and it is identified that the matter being discussed does not meet the Risk of Significant Harm (ROSH) threshold, then the CWU can assist the local worker to identify suitable support services. The CWU evaluation was not designed to assess the success of this strategy and as a result, this question cannot be answered here.

Finally, a finding from the Family Case Management evaluations was that one of the benefits for families of FCM intervention was to remove barriers to service access.

3.1.7 Are services more available and better able to meet the needs of Aboriginal children, young people and families?

The evaluations of various projects in this review have considered the degree to which the project made a difference for Aboriginal families. These projects were FCM, FRS, ADR and Safe Families. In the FRS evaluation, Indigenous families appeared to respond particularly positively to the service model. This was also the case for ADR where that form of dispute resolution and/or decision making was experienced as more culturally appropriate. An integral part of that was the capacity for ADR to engage the broader family group and include them in decision making and providing support.

Safe Families was the only Aboriginal-specific project reviewed in this synthesis. It was the project found to have the poorest implementation and the lowest apparent traction in attempting to achieve its aims of case management and community development. While there were indicators of activity within local communities, the establishment of foundational structures supporting the project rollout, such as Local Aboriginal Reference Groups, had not been completed. It was noted that organisational structures, more than local practice, lead to this situation.

3.1.8 Have services increased the levels of cultural awareness, cultural competence and partnerships with Aboriginal communities?

The evaluations of ADR and FRS suggest that, in these respects at least, KTS is contributing to improved cultural awareness, competence and partnerships in responding to Indigenous clients of the child protection system. The FGC evaluation notes that the facilitators had a specific focus on ensuring cultural appropriateness of the conferences, based on the family's wishes.

3.1.9 *What is the current state of partnerships between government and non-government services?*

There is relatively little evidence in this evaluation synthesis to address this question in any depth. The main exception is FCM, a joint pilot between government and non-government agencies where case management is undertaken by the agency or organisation with the most significant ongoing contact with the family. A comparison between the three trial sites suggested that commitment on the part of staff on the ground, and particularly the existence of local positions to foster local-level collaboration, can dramatically improve the effectiveness of these partnerships. Partnerships are also relevant to the FRS, which were shown to have had significant influences on the service system through identifying service gaps and family needs. The quality of relationships with external services was, however, not assessed in the evaluation.

3.2 Program Design and Implementation

In addition to illuminating these specific aspects of KTS, the evaluations considered here raise several systemic challenges to the design and implementation of large, complex suites of initiatives like KTS. The unifying theme among most of these concerns **clarity in the early stages of planning**, and manifests itself in different ways.

- Many of the evaluations testify to the need for initiating briefs to clearly define the substantive purpose and scope of the initiative, and to the importance of distinguishing this from the formal rationale of implementing recommendations handed down by an authority such as the Wood Commission.
- Many also testify to the challenges which arise from the absence of clear role demarcation – both between multiple agencies, and between government and different stakeholder groups. It is possible to address this through staff training and stakeholder communication prior to and during implementation, but this requires a good deal of planning and forethought.
- This touches on a broader issue of particular relevance to KTS: good collaboration is resource intensive, and the time required for administrative work associated with joint working seems to be underestimated. Extra funding is necessary if it is to be achieved. Moreover, there is also a need for clarity around terminology: the words collaboration, partnership and coordination are used in overlapping but not always entirely consistent ways in the documents considered here, making it difficult to determine whether they are discussing comparable phenomena.
- There is also a need for clarity around administrative data collection systems, which were found to be insufficient for use by staff in the field, or for subsequent evaluation and review. The implications of these systems for evaluations are discussed below. It is worth noting that several evaluations found evidence that implementation was not supported by adequate data collection systems (e.g. the Children's Court and WellNet).

In addition to the need for conceptual clarity, several evaluations raised systemic issues concerning workforce capacity and service gaps. Many government agencies and non-government service providers are operating at capacity. This has compromised the implementation of new service delivery initiatives such as Safe Families, and has also

hampered the effectiveness of referral services (which cannot work well if the services to which they would like to refer families are at capacity or do not exist in the first place).

3.3 Program Evaluation

The main point to be drawn from the evaluations examined here is that rigorous, robust, methodologically-sound evaluations of ‘young’ and ‘still developing’ initiatives are difficult to conduct. A second conclusion which is suggested but not conclusively supported by these documents is that impact evaluations are difficult for programs seeking to achieve outcomes that cannot easily be quantified. Each of these challenges is inherent in evaluation, but has been exacerbated by factors which could be mitigated: the impact of sequencing and duration of implementation and evaluation; and data quality.

Many of the evaluations considered here encountered difficulties because agencies privileged implementation over evaluation. This is understandable, particularly in the context of high-profile reforms such as KTS which address problems on which the need for action is a matter of widespread consensus. Unfortunately, sound evaluations are most likely when their constituent components are undertaken at the right time and in the right order. Specifically, this means:

- **Evaluation planning should commence before implementation.** Ideally, evaluation methods should be selected and baseline data should be collected early, since it is often difficult to retrofit a method and collect baseline data after implementation has begun. Despite the emphasis in the *Action Plan* on evidence-based policy and ongoing monitoring, this proved to be a particular challenge for many programs because of the complexity of the reform and the ambitious timeframes for implementation to which the Government committed itself. Even so, the benefits of sound planning were borne out in the evaluations considered here: the relative success of the Family Referral Services and Alternative Dispute Resolution evaluations were, in part, due to this (although even these encountered significant difficulties).
- **Evaluation research should not be undertaken too early,** since programs usually take time to set up and to have measurable impacts. The evaluations within this cohort that sought to gauge project outcomes (such as FCM and ADR) were unable to produce conclusive evidence because the initiatives had not dealt with enough clients, nor had there been enough time between the intervention and the evaluation to show effects over time. The length of time allowed for the evaluation process affected the ability to obtain meaningful results. Attempting an evaluation too early often has methodological consequences, too, such as difficulties in recruiting participants and control groups. Family Case Management is a case in point, although Safe Families also encountered difficulties recruiting clients for the Stage 1 evaluation. A related issue is that **the terms on which evaluations are conducted should be clear and explicit.** There is a fundamental difference between process reviews, which seek to establish whether an initiative is being implemented as originally planned (or, to put it another way, whether it is reasonable to expect the initiative to deliver results) and an outcomes evaluation (which seeks to establish whether the initiative is actually delivering those results). Both are far easier to conduct where initiatives are designed with evaluation in mind: where desired outcomes, measures and program logics are clearly defined prior to implementation,

and where they guide both implementation and evaluation. The absence of a clear results logic probably contributed to the challenges faced by the Family Case Management evaluation.

The problems of data quality encountered during the evaluations considered here have at least three sources. The first has already been discussed: the tendency to conduct reviews too early means that data are often not available. Second, administrative data of the kind required to review or evaluate initiatives are often regrettably either not collected, not reliable, not easily available or (in some cases) not made available to evaluators. The third flows from the fact that evaluators, in attempting to overcome these barriers, are often forced to adopt research techniques which are known to be less robust. This leads to a widespread over-reliance on qualitative interviews, often involving relatively small numbers of participants selected using methods that are acknowledged to be potential sources of bias. Unfortunately, in the absence of reliable quantitative data, it is usually quite difficult to reach any firm conclusions about the extent of these limitations. It would be particularly useful if evaluations of human service programs were able to rely on more robust time-series data, whether collected in bespoke fashion for individual evaluations or available through administrative systems.

3.4 Conclusion

This synthesis has drawn on the information available from evaluations of eight KTS-specific initiatives to understand the early stages of those projects individually, and to highlight what those evaluations can tell us about KTS implementation and evaluation in general. The primary lessons are that KTS initiatives are generally well conceived and implemented though the time pressures did impact on quality in some areas. Evaluations have been able to draw useful conclusions regarding programs and been generally sufficient to meet most of their aims, though there is room for improvement for future evaluative activities.

The *KTS Action Plan* committed to evaluating the individual components as well as the reforms as a whole. The majority of the evaluations in this synthesis were developed and contracted by agencies holding the lead in implementation of the specific project. In more recent times the KTS Evaluation Steering Committee has taken an oversight role to advise and coordinate evaluation activities and ensure integration of project evaluations across KTS. This improvement should help to build consistency across evaluations, integration across initiatives and to ensure high level expertise to guide evaluations.

4. Appendix 1 – Analysis of Evaluations

The following section of the synthesis considers evaluations of three secondary services – the Family Case Management project, the Family Referral Services Program, and the Child Wellbeing Units, which aim to improve outcomes for vulnerable and at risk children, young people, families and communities. It then discusses four initiatives in the tertiary (statutory) child protection space – Legal Aid Pilot and Dispute Resolution Conferences (which were evaluated together and reported here under one heading), Family Group Conferencing and Community Services' Screening and Response Priority Tools. It concludes by discussing an early intervention program specifically seeking to improve outcomes for Aboriginal communities – Safe Families.

4.1 Family Case Management

The Family Case Management (FCM) initiative provides integrated case management for families who are frequently encountered by the broader child protection system. It was adopted in direct response to recommendation 10.7 of the Wood Report, which called on key agencies to identify their most frequent clients, and provide these families with an integrated response involving the participation of relevant non-government organisations. The Inquiry also recommended the development of mechanisms to identify these families and to enable families to exit with suitable supports in place.⁸ FCM forms part of the *Action Plan's* focus on providing better protection for children at risk. It also forms part of the *Action Plan's* focus on changing practice and systems:

Effective referral systems are needed to connect children and families with the right services (universal, early intervention and statutory child protection services). These services need to be properly coordinated by government and non-government agencies, and agencies need to be able to share information so that children and families do not fall through the cracks. A centrepiece of the reforms is a new model for the intake and referral of child protection concerns.⁹

4.1.1 What did the project set out to do?

The goal of FCM was to reduce risk of significant harm to children and young people in “frequently encountered” families (i.e. those which are in regular contact with a range of agencies or services, and in which at least one child is at risk of harm). FCM sought to improve family functioning by more promptly identifying families with needs that required a multi-agency response, and ensuring service providers worked jointly to identify those needs and provide for them.¹⁰ This, in turn, sought to reduce the burden imposed by frequently-encountered families on the statutory system.¹¹

⁸Wood (2008). *Report* [“Wood Report”], Vol. 1 State of NSW through the Special Commission of Inquiry into Child Protection Services in NSW, Sydney. Retrieved on 1 May 2012 from http://www.dpc.nsw.gov.au/data/assets/pdf_file/0010/33796/Volume_1_-_Special_Commission_of_Inquiry_into_Child_Protection_Services_in_New_South_Wales.pdf Recommendation 10.7.

⁹*Action Plan*, p. iii.

¹⁰Department of Premier and Cabinet (2010). *Family Case Management Evaluation Design and Execution. Request for Proposal* [“FCM RFP”]. Unpublished document held on file by Department of Premier and Cabinet, Sydney.

¹¹ARTD Consultants (2012). *Family Case Management. Final Report* [“FCM Final Evaluation”]. Unpublished document held on file by Department of Premier and Cabinet, Sydney, p. xi. Community Services had previously estimated that 7,500 families accounted for 48% of child protection reports in 2009 (*Action Plan*, p. 14).

FCM was implemented in two stages: an initial pilot followed by state-wide roll-out. The pilot (Stage 1) was to last for 18 months, and was to be limited to three locations. State-wide roll-out (Stage 2) would be guided by the findings of an independent evaluation of Stage 1.

The pilot model emphasised pervasive collaboration at the local level. Teams of local managers from agencies and NGOs would work together to identify 'frequently encountered families', and develop and deliver integrated case plans. Each region was to identify 30 families to participate in the pilot. Case management was also collaborative, and applied to all relevant members of identified family groups, not just the children or young people considered to be at risk. It was intended that families would remain in FCM for 12 months, at which point any ongoing actions or needs would be handed over to individual agencies for action. Brokerage funds of \$5100 per family were made available, with local management groups responsible for determining how they were to be used (whether it be purchasing a specific service for a family to pooling the money to provide a specialised service for all families).

The pilot also had several strategic purposes. The *Action Plan* noted that the characteristics and needs of frequently encountered families in a number of locations would need to be identified, to better understand what was required for a more coordinated service-delivery approach by government and non-government agencies.¹² In addition, the pilot was intended to prepare for state-wide rollout by driving improvements to joint working practices (across government and between government and NGOs) to make integrated case management easier. This was to allow front line staff to focus on service delivery rather than negotiating procedural, systems or process barriers.¹³

During the pilot, a FCM Coordinator was to support local project teams in each region. It was intended that coordinators would provide administrative and organisational support, as well as some case management input if needed. They were also to play an important role in identifying barriers and resolving issues to improve joint working practices. During the set-up of Stage 1, the Department of Premier and Cabinet (DPC) was to provide overall project management and coordinate the evaluation of FCM.

4.1.2 What was done?

The FCM pilot started in early 2010 and was established at eight sites across three NSW regions: South East NSW (with FCM teams in Bega, Goulburn, Queanbeyan), South Western Sydney (Fairfield, Greenacre and Green Valley/Miller) and Western NSW (Orange and Leeton/Narrandera). The Western NSW pilot focussed on Aboriginal families in particular. The evaluation of the pilot concluded in early 2012, and FCM has since been rolled out in Tamworth and Wollongong.

Implementation was shared between multiple agencies and service providers at various levels of government. High-level project management and evaluation was handled by DPC. Within each region, implementation was led by a nominated 'host' agency: Housing NSW in South East NSW, Ageing Disability and Home Care (ADHC) in South Western Sydney, and Mission Australia was engaged to auspice the project in Western NSW. In all three regions the following agencies participated in the local management groups: CS; Ministry of

¹²*Action Plan*, p. 15.

¹³*FCM RFP*, pp. 4-5.

Health; DEC; Housing NSW; ADHC; Aboriginal Affairs; Juvenile Justice; the NSW Police Force; and Corrective Services.

Due to delays caused by the tendering process in Western NSW, implementation commenced earlier in South West Sydney and South East NSW. Coordinators commenced work in South East NSW and South West Sydney in December 2009. Whilst Mission Australia commenced work in March 2010 in Western NSW, a Coordinator was not recruited until August 2010. Progress in each region to December 2010 (12 months into the project), is depicted in Table 2 below.

Table 2 – Progress Piloting Family Case Management

<i>Milestone</i>	<i>South West Sydney</i>	<i>South East NSW</i>	<i>Western NSW</i>
Local Management Group meetings commenced	February 2010	February 2010	May – June 2010
Families identified	February – September 2010	February – April 2010	June 2010 – ongoing
Number of families consented by December 2010¹⁴	33	30	11
Number of families active at December 2010	27	30	9

By April 2011 (the end date of Stage 1 evaluation), no clients had completed their 12 months in the project. 44 families had received brokerage funding, of whom 85% required less than \$3,000. By March 2012, 62 families had participated, and 45 had participated for sufficiently long to be considered for the outcomes evaluation (90 days over a period of twelve months or longer).¹⁵

4.1.3 How was it evaluated?

ARTD Consultants were engaged by DPC to conduct a two-stage evaluation of the FCM project. The first stage was a process and interim results evaluation.¹⁶ The Interim Evaluation was undertaken to inform decisions about the broader roll-out of FCM and identify whether any changes to the model were needed. It was based on administrative data collected by FCM Coordinators (such as meeting attendance records and brokerage tracking sheets), family data collected by case managers (including the North Carolina Family Assessment Scale-General (NCFAS-G) results and service use records), and data collected by ARTD (including a staff survey and interviews with FCM coordinators).

The second stage combined a results evaluation, an economic evaluation, and a definition and sizing analysis.¹⁷ This second stage of the evaluation was primarily designed to answer the question ‘what outcomes were achieved for families, in what contexts and through what mechanisms?’. It used quantitative data from a number of sources to compare families’

¹⁴This count includes withdrawn families.

¹⁵FCM Final Evaluation, p. 10.

¹⁶ARTD Consultants (2012). *Stage 1 Family Case Management. Interim Evaluation Report* ["FCM Interim Evaluation"]. Unpublished document held on file by Department of Premier and Cabinet, Sydney

¹⁷FCM Final Evaluation

service contact in the 12 months prior to participation in FCM with their contact in the 12 months since participating in FCM. Data on family functioning, which was assessed using the North Carolina Family Assessment Scale – General (n = 18), was combined with data on patterns of service use (n = 45) provided by FACS -CS, Police, DEC and Health, FACS –Housing NSW.¹⁸ These were supplemented with qualitative data from interviews with service users (n = 11) and primary contacts (n = 12). The economic evaluation was conducted by way of a cost-benefit analysis, and sought to answer the question of whether the incremental money spent on FCM in 2010/11 and the outcomes achieved appeared to justify wider roll-out of the program. The synthesis does not examine the definition and sizing analysis, as it was essentially a separate research exercise designed to scope and help define the FCM target group for subsequent rollout.¹⁹

The final evaluation only examined the FCM project in South East NSW and South Western Sydney. At the time the evaluation commenced, the FCM project in Western NSW had not engaged the planned number of families (30). The small sample size meant that, for ethical reasons, the region could not be included in the evaluation. In January 2011 ARTD prepared a separate report on the implementation of FCM in Western NSW at the request of the project host – DPC.

4.1.4 *What did the evaluation(s) find?*

The Interim Report found that the FCM was largely implemented consistently with the project plan. The only significant deviations from the project plan were that Stage 1 of the project was extended beyond 18 months, and that state-wide roll-out has not yet occurred.

The Interim Report also found that the integrated style of case management of the kind provided by FCM could indeed provide a more ‘joined-up’ service than could often be achieved by any one agency working individually. It identified four factors as particularly important contributors to this success. First, the FCM Coordinators played a critical role as sources of information and coordination, and the evaluators suggested that they would be a key component of the model for the roll-out to future sites. Second, staff commitment to the aims and approach of FCM also facilitated implementation by ensuring good participation in the project. Third, the evaluators found that assigning project governance to DPC was seen to be important by other participants because DPC is able to facilitate coordination across human service, justice, health and education agencies. Finally, the evaluators also noted that brokerage funding was one of the key mechanisms used to help individualise case management plans. They also reported, though, that from the available \$5,100 per family the average amount of brokerage funding used was \$3,000. In light of this underspend they suggested funding allocated per family could be reduced.

The Interim Report was also designed to identify factors that hindered successful implementation of the project and that would require attention during the state-wide rollout. Consistent with this, the evaluators identified several issues flowing from the fact that the pilot was implemented within an urgent policy context which made it difficult to implement in a comprehensively planned manner.²⁰ First were problems caused by a lack of clarity with respect to role demarcation: the evaluators found that management groups

¹⁸Housing NSW data was subsequently excluded from the analysis. See discussion below.

¹⁹FCM Final Evaluation, p. 45.

²⁰FCM Final Evaluation, pp. 53-54

were too involved in case planning, and there was insufficient clarity about roles and responsibilities, particularly for case managers. There was also high staff turn-over as agencies progressively worked out which staff should be involved in the project and in what role. Second were problems with stakeholder engagement: in Western NSW, the limited involvement in the project by Aboriginal community stakeholders impeded the process of engaging families. More generally, the process for identifying families proved unexpectedly time-consuming; the evaluators found that a family's readiness for change was a key factor in determining which families to work with, and recommended that a client readiness tool be used to identify families for participation. Third, urgency impacted on the quality of training: the evaluators found that FCM was introduced without organised training for staff to support its delivery, and many staff reported that they felt ill-equipped to take on case management responsibilities. Finally, urgency caused administrative issues in accessing brokering funds, leading to significant delays in processing claims and often to detrimental effects on clients.

The Outcomes Evaluation found that FCM worked well for some families, but that it was not methodologically possible to attribute these outcomes solely to the pilot (for reasons discussed below). The evaluators reported that the analysis of family functioning (using North Carolina Family Assessment Scale – General assessment scores) and interviews with family members provided evidence suggesting that, for some families, the project is making a real and positive contribution to the wellbeing of family members and overall family functioning. On an individual level, service users who were interviewed provided particularly positive accounts of the benefits of FCM, most of which involved the project removing fairly basic barriers to individual functioning and linking families to services. As a group, the Evaluation found that better outcomes were associated with families which had been involved in the pilot for longer, which had fewer children and a sole primary caregiver, and who appeared to have less serious issues requiring coordinated management to begin with (i.e. a lower level of 'observed' negative involvements with government agencies on commencement).

The Outcomes Evaluation found that FCM could also be expected to represent a net economic benefit to government as well, but not before the ninth year of operation. It presented administrative data to show that participants experienced fewer reports to Community Services and fewer problems at school, but slightly more hospital emergency department presentations and family members being listed as a person of interest (POI) to Police. However, the evaluators noted there were a number of methodological limitations to the analysis, which made it difficult to draw conclusions about service use findings.

4.1.5 How sound are the findings?

The multiple methods proposed for the two stages of the evaluation were designed to provide a rounded set of data for analysing the effectiveness of the FMC program. The qualitative and quantitative measures were broad in scope, though perhaps unrealistic due to the complexity of accessing some of the specific agency data sought. Similarly, the intention to undertake a quasi-experimental study through use of a comparison group and sample of 90 families reflected a mismatch between the evaluators' hopes and the level of program maturity. For the final evaluation plan it was agreed that, given the pilot context for the project, a formative rather than experimental approach to the evaluation would be adopted.

Data quality was second contributor to the inconclusive nature of the findings. These flowed, in part, from the decision to conduct the evaluation at the same time as the project was being rolled-out. In addition, some of the data that was collected by service providers was of insufficient quality to use for analysis. NSW Housing data, for example, could not be used in analysis for this reason,²¹ which meant that changes in service use for public housing tenants could not be included in the evaluation despite apparently suggesting encouraging results.

The combination of poor data quality and small sample sizes meant that the reliability of many quantitative measures was compromised and the capacity to undertake statistical analysis was limited. It was not possible to cross-check statistical results against interview data, because constraints imposed by the ethics committee prevented families who were interviewed from being identified in the family functioning or administrative data. Nor was it possible to control for the possibility that the families who agreed to participate in the interviews may have been those who were most positive about their experience of FCM, for exactly the same reason.

4.1.6 Conclusion

The Interim Evaluation and the Final Evaluation suggest FCM offers a potentially effective model for collaborative case management as a core business practice in dealing with frequently encountered families. The significant impact that local implementation structures and coordinators have upon effective implementation and outcomes of the project is an area that should be closely monitored and supported. The evaluators also provided recommendation for improvements to the FCM model and associated processes when implementing FCM in other regions. Due to practical challenges encountered during the conduct of the Final Evaluation, however, the evidence that FCM will deliver significant, quantifiable outcomes, or that it will deliver sufficient economic value to justify its costs, is indicative rather than conclusive.

4.2 Family Referral Services

Family Referral Services (FRS) link vulnerable children, young people and families who are below the threshold for statutory child protection intervention with appropriate support services in their local area.²² They were piloted and rolled out in response to recommendations 10.1 and 10.2 of the Wood Report,²³ which called for the establishment of what it called “Regional Intake and Referral Services”. These were to work with other initiatives such as the Child Wellbeing Units to improve access to services for children and families who fall below the threshold for statutory intervention but would benefit from specific services.²⁴ Like FCM, FRS form part of the KTS Action Plan’s emphasis on changing practice and systems,²⁵ albeit with an emphasis on strengthening early intervention and community-based services rather than on improving the statutory child protection system.

²¹FCM Final Evaluation, footnote 9

²²Action Plan, p. 24.

²³Wood Report, pp. vi-vii

²⁴<http://www.health.nsw.gov.au/initiatives/kts/frs.asp>

²⁵Action Plan, p. 24

4.2.1 *What did the project set out to do?*

The Wood Commission's goal for Regional Intake and Referral Services was to prevent families who fall below the threshold for statutory intervention from moving into the statutory system. They were to do this by identifying secondary services families required, and referring them to appropriate providers to ensure they did not "fall through the cracks".

Justice Wood made several specific recommendations about the form the Regional Intake and Referral Services should take. They were to be located within an NGO in each region, which would determine the nature of the services required and refer a family to the appropriate service provider for services such as case management, home visiting, intensive family support brokerage, quality child care, housing and/or parenting education. They were to include one or more child protection caseworkers seconded from Community Services, with the overall number of staff dependent on anticipated demand in the particular region. If this child protection caseworker formed the view that a child or young person may be at risk of significant harm, they would perform a history check on Community Services' KiDS database and, if necessary, refer to the matter to the Helpline. Finally, there was to be at least one Regional Intake and Referral Service in each Community Services Region.

The Government renamed Regional Intake and Referral Services, preferring instead Family Referral Services, and proposed to trial them in three areas (one metropolitan and two regional/rural). The purpose of this trial was twofold: first, to determine appropriate services for the child/family, building on rather than supplanting agency action, expertise and capacity; and second, to drive improved links between local government and non-government services, and provide advice to agencies to support better realignment of local services. The trial services used different service delivery models – telephone only or 'augmented services' with capacity for time-limited direct support to connect families with ongoing services – and different operating contexts – rural, city and suburban.

4.2.2 *What was done?*

The NSW Government committed \$38 million to a staged implementation of FRS over five years from 2009. This included \$3.6 million allocated to the piloting of FRS in three regions during the first 12 months. The Ministry of Health was nominated as the lead agency for the implementation of FRS.

The FRS pilot began in May 2010 in Mt Druitt, Newcastle and Dubbo, and to better determine appropriate service models, each took slightly a different form. Relationships Australia piloted a telephone-only FRS in Mt Druitt, designed to provide information and/or referrals for children, young people and/or their families to specific local and available support services. The Benevolent Society piloted an "augmented" FRS in Newcastle, which included both a telephone service and the capacity to conduct more active referrals and supplement access to support services with the use of brokerage funding. In Dubbo, UnitingCare Children, Young People and Families piloted an augmented FRS with an Aboriginal focus.

Over the nine month evaluation period a total of 1,115 contacts (referrals) were received, with government agencies being the primary source. A total of 2,067 referrals to other service providers were made by the FRS across all sites, with most referrals made to suppliers of support services within the local areas. The majority of clients were referred to

parent/family support services (26%), followed by welfare services (23%), health services (16%) and housing services (14%).

4.2.3 *How was it evaluated?*

The Ministry of Health commissioned two separate evaluation exercises for the FRS pilot. Ernst & Young were engaged in April 2009 to undertake a process evaluation covering the period from May 2010 (program start date) to January 2011.²⁶ In order to inform the state-wide implementation of the FRS Program, Ernst & Young were also asked to examine the effectiveness and efficiency of the FRS over the first nine months of operation and provide recommendations regarding the optimal service model. To the extent that the evaluation of the pilot stage of FRS was also intended to inform future state-wide implementation of the FRS, the evaluation can also be considered formative in nature. This process evaluation was based on quantitative data on service utilisation, service equity and reach, and service uptake and operational costs.

In addition, the Ministry also engaged ARTD Consultants to conduct a service system survey, qualitative research with FRS staff and a client satisfaction survey. Baseline and follow-up surveys and interviews were conducted, and were used to inform the quantitative analysis undertaken in the process evaluation.

4.2.4 *What did the evaluation(s) find?*

The evaluators found that the FRS pilot sites were making a difference to clients in vulnerable families by identifying appropriate support services, improving links between services, and informing realignment of support services at the local level. They also provided evidence of increased effectiveness, over time, in facilitating vulnerable population access to support. FRS were also shown to be servicing a larger proportion of Aboriginal clients, compared to their catchment area, suggesting they met the goal of providing culturally appropriate referral pathways for Aboriginal families.

The evaluators found that the augmented service model itself was a contributor to success, in that it appeared to perform better than a telephone-only service model in facilitating the access of vulnerable populations to support services. The availability of augmented services was particularly important for building trust and engagement with Aboriginal children/young people and families. Similarly, the telephone model proved to be a barrier to service in some circumstances: for example, the Mt Druitt FRS had some difficulty establishing contact with families a second time, which presented logistical issues for some clients and impacted on client access.

The evaluators found, however, that the FRS model rolled out state-wide should be sufficiently flexible to allow each site to tailor the services provided to suit the specific needs of their clients. This is partly because participation in local service provision networks enabled the FRS to source new services that were small and unique and not widely known in the area, and partly because targeted brokerage to Aboriginal agencies had proved to be an effective way to increase Aboriginal engagement. Location near other services and flexibility

²⁶Unless otherwise specified, information for this section of the synthesis has been extracted from Ernst & Young (2011). *FRS Evaluation. Final Report* ["FRS Evaluation"]. Unpublished document held on file by NSW Health, Sydney.

would also help facilitate Chapter 16A information sharing – about which the evaluation specifically identified a need for further training within the service sector.

The evaluators recommended that this local flexibility should be complemented by central support: they found that NSW Health played a key role in coordinating the pilots, facilitating the sharing of experiences between sites, promoting the FRS pilot sites and facilitating partnerships with other key government and non-government organisations. The evaluators also confirmed that there was a need for one or more child protection workers to be placed inside the FRS and provided with access to the Community Services' KiDS database in order to help identify whether a child has a history of being at significant risk. This was explicitly recommended by the Wood Report, but not trialled as part of the FRS pilot.

The evaluators also found that two sets of external factors influenced the ability of the FRS to achieve some of their key goals. First was the capacity of the existing service system. Many services in the three pilot areas were already at full capacity and had long waiting lists. The FRS, through the evaluation, identified gaps in available services including homelessness, services for parents with children suffering a serious or terminal illness, services prepared to work with parents who have a mental illness, and disability respite services for children under 6 years. As a result, the 48-hour referral target proved difficult to meet, with the delays being caused by waiting lists, intake procedures and stringent eligibility criteria (which made it difficult to find a place for some clients with complex or borderline needs). The second external factor was degree of engagement with the target population itself. The original forecast for the potential demand for FRS was well in excess of the actual demand. This was perhaps due partly to lack of awareness, and on this basis the evaluators recommended more effective advertising to increase the community knowledge of the service.

Associated with broadening community knowledge regarding the FRS the evaluators identified significant economies of scale in the provision of FRS. The unit costs of service provision decrease significantly as the number of clients increase, suggesting there are efficiencies to be derived from the roll-out of the FRS to those regional areas that have large numbers of potential clients. This, coupled with the fact that the number of complex cases was much higher than expected, suggests that accurately forecasting service demand will be crucial to successful state-wide rollout.

4.2.5 How sound are the findings?

Overall, the evaluation was well-planned and well-executed, and produced reliable evidence that FRS are meeting their objectives by providing linkages for vulnerable children and families to appropriate health, welfare and family support services. It is important to note that the purpose was to examine service activity and inform further rollout, not to evaluate outcomes arising from client's use of the services to which they were referred. Thus, further evaluation is needed to determine the impact of FRS on client outcomes.

One significant contributor to the success of the evaluation was the fact that it was planned before roll-out of the initiative began. This admittedly caused some difficulty during the start-up period, when FRS processes and procedures were still being developed and continuing to evolve. In terms of assessing whether FRS are meeting their objectives, an evaluation of six months of operation is a short time for a new service to have a widespread impact in its sector. Data quality was variable across all sites, particularly in quarter 1, but

because the evaluation was well-planned, an analysis of quarter 1 data was used to improve data collection in quarters 2 and 3.

Early planning could not overcome every problem, however: from a strict methodological point of view, it is difficult to draw definitive conclusions about the relative performance of the two service delivery modes due to the significant differences in client populations and service delivery models being operated by the three FRS pilot sites, and relatively low survey response rates. But the evidence that was gathered provided enough information to support the recommendations that were made.

4.2.6 Conclusions

The implementation of FRS across the three trial sites was found to be effective and generally as planned, though there were lessons learned and improvements made along the way. The FRS were effective in reaching and supporting families, though service demand (at the early stage) was much lower than projected, suggesting a need for further community engagement and advertising. At the same time, the number of families with complex needs was higher than expected.

Based on an analysis of the factors which had facilitated and impeded implementation, evaluators found that the greatest net benefits were likely to be realised in future by rolling out the augmented FRS model to those regions where relatively high proportions of the population are Aboriginal or in receipt of welfare payments; and where those populations are relatively centralised around a major population hub.

Following the completion of the evaluation and undertaking a needs mapping process, FRS have been progressively rolled out state-wide. The original three services were renamed the Mt Druitt, Hunter Central Coast and Western FRS, and their catchment areas were expanded. The Illawarra and New England North West FRS were established in 2011, and as of July 2012 the tendering process to establish the further services was underway.

4.3 Child Wellbeing Units

Child Wellbeing Units (CWUs) have several overlapping functions. They advise, support and educate their respective agencies' mandatory reporters to help them determine whether a child or young person is subject to suspected risk of significant harm (ROSH). In cases where children are at risk of significant harm, they ensure that these matters are reported to the Community Services' Child Protection Helpline. In other cases, they identify potential responses by the agency or other services to assist the child or young person or family. Over time, they aim to drive better alignment and coordination of agency service systems.²⁷ CWUs were established in direct response to recommendation 10.1 of the Wood Report, and are a central element of the systemic reforms in the Action Plan.

4.3.1 What did the project set out to do?

The CWUs are a core feature of Justice Wood's vision for child protection in NSW, and are intended to support the new model for the intake and referral of child protection concerns, and the new statutory reporting threshold. They contribute to better outcomes for vulnerable and at-risk children and young people by making the service system more

²⁷ FRS Evaluation.

effective. Mandatory reporters in relevant government agencies who believe a child or young person is at risk of significant harm (ROSH) can still report directly to the Child Protection Helpline. CWUs constitute a first point of contact for those mandatory reporters who require assistance to determine the level of risk, and help to identify appropriate local action or referral options for cases which do not meet the new threshold for reporting to Community Services. At a systemic level, CWUs reshape agency responses to child protection by providing training, support and advice to staff and working with other agency units to look at more effective delivery of services to children who are deemed to not be at risk of significant harm.

4.3.2 *What was done?*

CWUs were established in January 2010 in the NSW Police Force, the Department of Education and Communities, the Department of Family and Community Services, and the NSW Ministry of Health.²⁸ They are supported by a dedicated database, WellNet, which was also created in 2010 to allow information to be centrally recorded and accessed across the CWUs. A key function of WellNet is to store information on previous contacts, allowing CWUs to take this information into account when appraising the level of risk to a child. This was a direct response to the Wood Report's concern with cumulative harm. The CWUs are staffed by managers and officers with expertise in child protection intervention. Two Community Services child protection caseworkers have been seconded to CWUs to facilitate interagency collaboration and information exchange between the CWUs and Community Services.

4.3.3 *How was it evaluated?*

In April 2011, DPC engaged the Nous Group to conduct a review of the implementation of CWUs.²⁹ The review was conducted from April to June 2011. The objectives of the review were to report on the strengths and weaknesses of implementation to date, and identify any changes to existing structures, staffing, processes or systems required to ensure the CWUs' effectiveness in improving child protection and wellbeing services in NSW. The review was based on data from CWU Staff Surveys (n= 74), an online Mandatory Reporter Survey (n= 2537), a review of operational data, stakeholder consultations (n=26) and structured workshops with CWU staff.

4.3.4 *What did the evaluation(s) find?*

The evaluators found that implementation of the CWUs had been largely successful, particularly given the tight timeframes in which it was undertaken. They found that CWUs contributed to the overall reduction in reports to the Child Protection Helpline (although due to data limitations discussed below, they recommended further analysis of the specific contribution of CWUs to this reduction be undertaken as part of the broader evaluation of KTS). They also found that CWUs were integral to promoting a shared responsibility for child protection and the changes in work practices that underpin the KTS reforms. Finally, they found that CWUs are well-placed to identify service gaps.

²⁸ *Action Plan*, pp. 23-24.

²⁹ Unless otherwise specified, information for this section of the synthesis has been taken from Nous Group (2011). *Keep Them Safe: Review of NSW Child Wellbeing Units. Final Report* ["CWU Review"]. Unpublished document held on file by Department of Premier and Cabinet, Sydney.

The evaluators found that the flexibility of the CWU service model has enabled the CWUs to meet the needs of their individual agencies by balancing common elements and shared objectives. Overarching CWU Operating Guidelines provide a level of consistency, complemented by agency-specific operating procedures. The evaluators concluded that there would be merit in maintaining the existing model as it is appropriate for the different agencies, meets the needs of different mandatory reporters, and has resulted in increased awareness and understanding of the shared responsibility amongst mandatory reporters. Similarly, the out-posted case worker positions were also identified as a strength of the model. The evaluators noted that CWU staff consultations and a separate Community Services review of the service had revealed that the arrangement was an important facilitator of effective information exchange and inter-agency collaboration.

Apart from the strengths of the service model, the evaluation suggested that information systems, training and leadership contributed to the successful establishment of CWUs. The information on previous contacts with individual children and young people stored on WellNet was confirmed as enabling mandatory reporters, Community Services and CWUs to make more informed decisions about each child or young person. The successful integration of WellNet with the Police IT system (COPS) was also identified as an achievement. Surveys indicated that thorough initial communication with and training of mandatory reporters and CWU staff meant that people felt they had the skills necessary to perform their roles. These surveys also identified clear leadership and strong operational management by CWU Directors as contributors to the integration of CWUs in their agencies' broader child protection response.

Perhaps somewhat paradoxically, information systems and staffing issues were also identified as barriers to implementation. Despite the overall success of the WellNet database, the evaluators noted its limited capability for data analysis and reporting. CWUs are obliged to manually analyse data that is exported from WellNet in a raw form and this hinders their capacity to assess their performance and anticipate future demand. The staff survey and stakeholder interviews revealed, for their part, that staff instability and turnover a major challenge. Contributing factors appeared to be reliance on temporary positions and the lack of career development and progression opportunities. Disparities in Assessment Officer staffing classifications and commensurate salaries were also identified as a cause of dissatisfaction for staff in the Police CWU.³⁰

On the basis of these findings, the evaluators identified a number of opportunities to improve the ongoing effectiveness, efficiency and strategic contribution of the CWUs. Key recommendations relate to improving WellNet's analytical and reporting capability and exploring the possibility of integrating WellNet with agency systems. The evaluators suggested that CWUs should play an important role in contributing to the systematic capturing of service information. In order to do so, technical aspects of WellNet should be explored which could support the identification of service gaps and capturing of service information. The evaluators also found there were opportunities to provide more relevant and timely feedback to mandatory reporters and improve cross-agency collaboration. Finally, they recommended changing the terms of reference and operations of the CWU

³⁰Police CWU Assessment Officers, are employed at the Clerk 5/6 level whereas all other CWUs employ Assessment Officers at the Clerk 7/8 level.

Director' Forum to promote a strategic focus, and proposed that the Directors' Forum should establish KPIs that are explicitly linked to the objectives of KTS.

4.3.5 *How sound are the findings?*

The Nous review was conducted during the implementation of CWUs, and was intended to identify ways of strengthening future functioning. It was conducted with sufficient rigour to meet these immediate aims, but as the evaluators acknowledged their conclusions must be taken as subject to further confirmation due to the small sample sizes of some data and reliability of others. To take one example, few FaCS staff responded to the staff survey (n=6), and response rates were low across the CWUs more generally. There was also insufficient reliable data about the workload of CWU staff, including the relative proportion of workload across each of the three key roles. This impacted on the ability of the evaluators to assess productivity, funding levels and projected demand.

Given these limitations and the formative goals of the review, the evaluators recommended that a formal evaluation of the CWUs against the objectives of the KTS Plan be conducted after June 2012. If the review recommendations regarding the development of KPIs and improvements to WellNet were pursued, they also suggested assessing whether CWUs are meeting their objectives. Similarly, they noted that it might be possible for a future economic appraisal to be conducted if reliable client and workload data were collected.

4.3.6 *Conclusion*

The evaluators concluded that there would be merit in maintaining the existing model as it is appropriate for the different agencies, meets the needs of different mandatory reporters, and has resulted in increased awareness and understanding of the shared responsibility amongst mandatory reporters.

As a strategy that integrates with the legislative change of raising the reporting threshold to suspected risk of significant harm, CWUs appear to be contributing to the reduction of reports to Community Services and may be leading to greater service access and localised support for vulnerable families. But, as the CWUs are one component in a range of initiatives designed to support mandatory reporters while reducing the burden upon Community Services, it would be important to incorporate the CWUs into a review of the collective effects of all associated projects. These include the Mandatory Reporter Guide (MRG), and Community Services Screening and Response Priority Tools (SCRPT). An outcomes evaluation would be required to assess the degree to which CWUs have supported an increased degree of local support to families in need.

4.4 **Whole Family Teams**

The Keep Them Safe Whole Family Team (WFT) Pilot is one of a suite of initiatives being implemented in the statutory child protection space under the *Action Plan*. It is particularly focussed on the needs of whole families where parents have mental health and/or drug and alcohol problems³¹.

The Pilot was established to trial a new, integrated model for delivering tertiary specialist health services to families where drug & alcohol and/or mental health problems and child

³¹ *Action Plan*, p. 13.

protection concerns exist. Whole Family Teams were not a recommendation of the Wood Report(s), neither were they a specific commitment under the *Action Plan*. Whole Family Teams were developed in direct response to the consistently recognised and inter-linked issues of carer mental health and drug and alcohol abuse that impacts upon children.

4.4.1 *What did the project set out to do?*

The WFT intervention is a tailored, child centric, service model which seeks to improve parenting, reduce substance misuse, and the impact of parental mental health problems in order to reduce risks to a child. The specific aims of the Pilot are to increase safety for children, improve family functioning, decrease drug and alcohol related harm, and improve mental health.

As indicated by the name 'Whole Family Teams', the model of service delivery targets the needs of the family as a whole, in order to create a safer environment for children. WFTs provide integrated interventions tailored to the needs of the family group, including mental health, drug & alcohol, parenting interventions, and referral to other services as required, in order to improve the circumstances for a child at risk.

The model was designed to accept priority referrals from Community Services in instances where a report of concern has been made about a child and an assessment has found that a child is at risk of significant harm. Referrals from Local Health Districts (LHDs) can also be accepted if they meet criteria and do not restrict or jeopardise priority access for Community Services.

The initiative is designed so that mental health and drug & alcohol clinicians are co-located and work together with families to provide integrated specialist care. Community Services are major partners and the model states that casework managers should remain actively involved with the family, communicate regularly with the KTS-WFT and attend case conferences.

The intention is that interventions will be delivered across a 6 month period and each WFT will provide services for approximately 11,091 hours of client related provider time per year - drug and alcohol 4,767 hours and mental health 6,324 hours.

4.4.2 *What was done?*

The WFT project is managed by MH-Kids, Mental Health & Drug & Alcohol Office (MHDAO). Four pilot sites, located in Nowra, Lismore, Gosford, and Newcastle, have been funded for 5 years to June 2014. Establishment of the Pilot services began in 2009 and teams began taking referrals from Community Service Centres from December 2009.

Governance for the project was established early and provides comprehensive direction for each aspect of implementation. A state-wide Steering Committee oversees the project and developed a model of care in 2010, which has provided a framework for all Pilot sites. The model outlines a statement of service and guiding principles, but has allowed flexibility in determining locally responsive models of care.

Additional structures have been added to the model in response to the first phase of evaluation. These include an Implementation Group to ensure ongoing collaborative work in planning, coordination and implementation at a local level, a Technical Data Working Group

(TDWG) to advise and assist on data systems, collection methods and reporting processes, and an Evaluation Advisory Group (EAG).

4.4.3 *How was it evaluated?*

Urbis was commissioned by NSW Health to conduct a four-year evaluation of the WFT Pilot, finishing June 2014. The evaluation, currently at its midpoint, will analyse the Pilot in terms of effectiveness, appropriateness, efficiency and process.

The evaluation for the first two years has largely focussed on qualitative data to analyse the process measures of Pilot establishment and early operations. In the first year, the evaluation team visited each site and undertook interviews with key personnel to explore establishment challenges, early lessons, the development of clinical and organisational processes to support the operations of the WFTs, and distinctions between the four Pilot sites. In the second year, the evaluators again visited each site twice and interviewed key participants, including WFT staff, senior management within NSW Health and Community Services, and other services relevant to the local operations of each service.

Some data collection and reporting issues have been recently addressed and early data, which have been received but not yet analysed from LHDs via the state Health Information Exchange and local stand-alone database reports, reflect positively on client outcomes. In order to further support evidence of the effectiveness of WFT interventions and the outcomes for families, the methodology has been revised to include sixteen case studies.

4.4.4 *What did the evaluation(s) find?*

The Year 1 evaluation recommended improvements to the WFT in three areas. These concerned leadership and communication (which was actioned through the establishment of an Implementation Group designed to increase communication and provide a channel for escalating issues), improvements to “data systems and evidence”³² (with a consequent action of changing to use the North Carolina Family Assessment Scale as the initial capacity assessment tool and for subsequent use as an indicator of change), and improvements to evaluation oversight (with the establishment of the Evaluation Advisory Group).

The Year 2 evaluation found that the WFT model is bedding down in local areas with aspects such as collaboration with Community Services becoming stronger and more effective. One result of this is that appropriate referrals from CS (and other nominated services when capacity allows) are being received by WFTs. As a result, there are few referred clients who are assessed as not suitable for the service, and few who withdraw from the WFT service. Despite these successes, there are some systemic issues such as CS internal capacity to assess and allocate families before referring them to WFT, which impacts on referral rates and smooth transfer. Additionally, there is room for improvement at the point of clients engaging with other service providers such that progress made with the WFT is sustained.

The evaluators recommended that greater attention be paid to those aspects of the service model which are being implemented differently in different locations. The most significant example is the decision to begin or continue to provide a service to families whose children have been placed in care, even if only temporarily, is addressed differently in different sites,

³²Urbis (2012). *Keep Them Safe – Whole Family Teams: Year 2 Evaluation Final Report* ["Whole Family Teams Year 2 Report"]. Unpublished document held on file by NSW Health, Sydney.

and should be resolved for all services. A linked issue is the need to develop a shared definition of 'success'. This is both a practical issue of data collection, and a philosophical issue given the complex issues the WFTs are trying to address. Other differences between locations include the degree to which WFT staff engage directly with the family's issues or whether the WFT acts as case coordinator to engage a range of other services to meet the child or family needs, the degree to which children are provided with clinical services through the team or assistance comes through improvements in the circumstances of their parents.

Being a pilot service, gathering data on the activities and the clients of the WFTs is vital, but was not occurring to the standard required at the point of the Year 2 evaluation and the evaluators reported this as the "biggest concern and challenge for the Pilot sites".³³ Issues such as inconsistent understandings of terminology, incompatibility between the WFT database and a Local Health District (LHD) software upgrade, and not undertaking data entry, had impacted upon the data quality and availability. As a result, reports that were required were not able to be generated.

The evaluators found that Year 2 of the WFT has been characterised by less change at both state or site levels, and this has reduced concerns over governance. Whole Family Team staffing has been generally stable though some ongoing recruitment has been required across all sites, including centrally. In addition, training resources were considered adequate, though they need to be aligned with 'core competencies' for WFT work. These competencies were not defined at the time of the evaluation.

Finally, there is anecdotal evidence that some families have benefited from the WFT intervention, with improvements in mental health or in substance use and changes in parenting behaviour. At the same time, objective evidence, which verifies improvements in mental health status or drug & alcohol use of families who are receiving WFT services, is limited.

4.4.5 How sound are the findings?

As previously noted, the evaluation methodology for the first two years has largely focussed on implementation processes, achieved by generating qualitative data through interviews undertaken during site visits. While this provides opportunity to understand how implementation is experienced at the various sites, and provides a description of work practices, it is limited in the fullness of the picture of implementation that can be provided. Other data such as implementation-related documents, basic client numbers compared with anticipated client levels, etc. could provide greater detail about the initiative.

A great deal of attention has gone into ensuring the mechanisms for the next phase of evaluation are in place. Specifically, the data system was almost operational at the point of the Year 2 evaluation and should provide useful material for future evaluative actions. Additionally, time was spent in 2012 seeking ethics approval for undertaking interviews with service users. This did take longer than anticipated but should provide useful material for understanding the impacts of the service, from a service user perspective.

³³Whole Family Teams Year 2 Report, p. 23.

4.4.6 Conclusions

Three years into the Pilot, recommendations from the previous evaluation have been implemented, collaborative structures and processes appear to be largely stabilised, and working relationships between the WFTs and CS appear to be maturing. Constraints remain in the ability of CS to refer to WFTs due to Community Services' own workload and the requirement to allocate a CS case worker to each family before they can be referred to the WFT. At the same time, it is recognised that CS have invested significantly in their own staff resources in order to develop referral and liaison processes to assist the WFTs.

It is generally agreed that WFT staff have been well provided with resources and training to undertake their work and staffing has been generally stable. Three teams continue to recruit, and it is anticipated that recruitment may be more difficult in the remaining years of the Pilot due to the time-limited nature of the positions. Issues such as defining case loads remain an area of ongoing work.

The single most significant underlying tension within the Pilot continues to be the extent of focus on the child. An important contribution which the Pilot may make to the sector will be to define the most beneficial way in which a multi-disciplinary MH, D&A, and parenting service can engage with children who are at risk.

Data recording and reporting is an area that requires immediate attention both for local use by WFTs and for the next steps of the evaluation. Presently these issues are more at the user input stage and are amenable to simple interventions such as development of a data dictionary and suitable guidelines for staff.

4.5 Alternative Dispute Resolution

The term "Alternative Dispute Resolution" (ADR) refers here to several initiatives piloted in response to recommendations 12.1 and 13.12 of the Wood Report. The Report found that, whilst the *Children and Young Persons (Care and Protection) Act 1998* (the "Care Act") included allowed for the use of alternative dispute resolution (ADR) both prior to and during care and protection proceedings, in practice these provisions were not used sufficiently to make a difference. ADR is an important element of the Action Plan's focus on providing better protection for children at risk, particularly by "simplifying proceedings... to ensure that considered decisions are made to ensure the best outcomes for children" and by reforming "court processes so they are fairer and more user-friendly for children and their families."³⁴ In fulfilment of this commitment, an expert Working Party recommended a state-wide pilot of Dispute Resolution Conferences (DRCs) in the Children's Court, and a pilot of external care and protection mediation for care matters in the Bidura Children's Court in Glebe (the "Legal Aid Pilot", LAP). A third component – Family Group Conferencing (FGC) – is reviewed separately, below.

The two ADR mechanisms reviewed here are available during children's court proceedings (Family Group Conferencing is available prior to proceedings commencing). These two mechanisms have similar aims: to provide the parties with an opportunity to agree on the action that should be taken in the best interests of the child and, where an agreement cannot be reached, narrow the scope and length of the court hearing; to produce child

³⁴ *Action Plan*, pp. iii, 10, 17.

protection decisions that are better informed and more responsive; to foster collaborative rather than adversarial relationships between Community Services and families.

4.5.1 *What did the project set out to do?*

The ADR mechanisms trialled in this initiative seek better outcomes for children, young people and their families by involving them in processes leading to decisions in court proceedings, and thereby ensuring they are accepted by all parties and more likely to be implemented. Traditional hearings in the care and protection jurisdiction are adversarial and conflict-driven by nature, and so do not facilitate positive working relationships between family members and child protection services. ADR processes are, by contrast, designed to foster collaborative rather than adversarial relationships, which is particularly desirable where those involved have to work together after the initial dispute has been resolved. When used in the context of the Family Court, they assist those involved in family breakdown to communicate better with one another, and to reach informed decisions about their children. They also provide the parties with an opportunity to agree on the action that should be taken in the best interests of the child and, where an agreement cannot be reached, to narrow the scope and length of the court hearing.

Strictly speaking, the *Action Plan* only committed the government to seeking expert advice on possible models for ADR and on how to implement whichever models seemed preferable. The implicit intent of the commitment was, however, to act on this expert advice, and if possible to gradually introduce ADR in the care and protection jurisdiction in NSW – although no timeframes were given for this.

4.5.2 *What was done?*

An Expert Working Party was established, comprising representatives from the ADR Directorate of the Department of Attorney General and Justice (DAGJ), the Children's Court, Legal Aid, Community Services, the NSW Law Society and Bar Association AbSec and the Crown Solicitors and the academic community. This group reviewed available evidence regarding, and proposed four options for ADR in NSW. These applied to several different points in statutory child protection processes, and were:

- Family Group Conferencing (discussed below)
- Dispute Resolution Conferences (discussed here)
- Legal Aid Pilot (discussed here)
- Care Circles (a program already in existence in Nowra, which was proposed for review and potential adoption elsewhere in NSW)³⁵.

The recommendations of the Expert Working Party were accepted and each of the models implemented. Aspects of the project were overseen by the ADR Directorate within DAGJ, though implementation involved key partners such as NSW Children's Court, Legal Aid and Community Services..

The evaluation of the Dispute Resolution Conference trial was conducted between February 2011 and December 2011, during which time 1096 individual conferences were held. The

³⁵ The Aboriginal Care Circle Pilot was evaluated in June 2010 by Cultural and Indigenous Research Centre Australia on behalf of the then NSW Department of Human Services, community Services and the then NSW Department of Justice and Attorney General. This evaluation is not discussed here.

conferences themselves were scheduled to last two hours each, and were convened by a legally-qualified Children's Registrar who had undergone specialised training in ADR. Care matters could be referred to a DRC at any stage in the court process after a care application had been filed and the relevant parties notified, and followed a conciliation (i.e. advisory) model. DRCs take place in the relevant Children's Court building.

The evaluation of the Legal Aid Pilot was conducted between September 2010 and December 2011, and involved 100 care matters referred from the Bidura Children's Court. The Legal Aid Pilot was based on the Legal Aid Family Dispute Resolution Service, which seeks to resolve family law disputes without recourse to the court system. Mediations held as part of the Legal Aid Pilot were convened by an external mediator, followed a facilitative model, and were scheduled to run for 3 hours. Matters could only be referred to the Legal Aid Pilot after it had been established that the child was in need of care and protection or after the granting of leave ("section 90 applications"). The mediations were conducted in the Legal Aid NSW head office in central Sydney, and eighty-four were held during the evaluation period.

4.5.3 How was it evaluated?

The Department of Attorney General and Justice engaged the Australian Institute of Criminology (AIC) to undertake a process and outcome evaluation of DRC model and the LAP. The purpose of the review was to assess the implementation and effectiveness of ADR in the care and protection jurisdiction. The AIC developed a program logic model and evaluation framework that aligned with the KTS Implementation Plan.³⁶

The AIC evaluation comprised a literature review, participant surveys (with families, legal representatives, CS caseworkers, CS Manager Casework), observations of conferences, semi-structured face-to-face interviews with parents and family members, semi-structured face-to-face and telephone interviews and focus groups with key stakeholders, a qualitative survey of Legal Aid lawyers, Aboriginal Legal Service lawyers, CS caseworkers, CS Manager Casework, CS lawyers, Children's Registrars and mediators, and data collected by Children's Registrars and mediators and from court files. The evaluation was commenced in March 2011 and was concluded in June 2012.

4.5.4 What did the evaluation(s) find?

The evaluators found that a concerted effort had been made to ensure that both pilots were consistent with good practice principles for court-referred ADR, and that the pilots had largely been implemented as intended. They also noted that the standard of ADR appeared to have improved over time as the level of experience, knowledge and skills among those parties involved in the conferences (e.g. registrars and mediators) increased.

Effective implementation was supported in four other ways. Firstly, the pilots were well resourced by providing enough conference/mediation convenors to allow proper planning and preparation for each conference. Secondly, engagement with the Trial and Pilot was supported by key stakeholders, an effective communication strategy, and training for

³⁶Urbis (2011). *Implementation Plan for Evaluation of Keep Them Safe* ["Implementation Plan"] Department of Premier and Cabinet, Sydney. Retrieved on 1 May 2012 from http://www.dpc.nsw.gov.au/data/assets/pdf_file/0019/125146/Urbis_Final_KTS_Implementation_Plan_-_publication_version.pdf.

agency staff. Third, the models had a degree of flexibility to meet the distinct needs of different families. Finally, there were processes in place to support continuous improvement in delivery and to address implementation challenges.

The evaluators also identified four kinds of barriers to implementation within each pilot – logistics, reluctance on part of professionals, roles and responsibilities, unfamiliarity with legal proceedings. Firstly, some stakeholders identified logistic issues such as the 2 hours allocated to Dispute Resolution Conferences not being sufficient to address the issues. Others noted that some conference rooms were not big enough to comfortably accommodate the number of participants. Second, some professionals appeared reluctant to engage with the ADR processes and were perceived as approaching the conferences with set or adversarial positions, and appearing unwilling to work with families. The ways that CS personnel were perceived by families to have approached the conference was the strongest predictor of overall family satisfaction with the conferencing process. This measure was low early in the trial though did improve over time. Third, there were two issues with roles and responsibilities. Firstly, working relationships with families appeared to be impeded when Community Services' conference participation was lead by managers or legal representatives, as opposed to caseworkers. Secondly, Indigenous mediators within the Legal Aid Pilot were indentified at times as providing advice to families which was outside the role of the mediator. This reflects the complicated position of a mediator who brings specific cultural knowledge to the process, and has been the subject of further work by Legal Aid since the evaluation. A fourth barrier to implementation was the fact that a number of participants in both kinds of ADR did not have legal representation. As a result, these participants had difficulty participating in the proceedings despite attempts by other professionals who were present to involve them.

Despite these barriers, the evaluation found that these components of ADR were experienced positively by the majority of participants. More specifically, around 80% of participants found the conferences useful, even in the cases when they were not happy with the outcomes. It was also felt that working relationships between families and CS were improved in around half of the conferences held during the trial period.

Conferences were also found to have had a positive impact on the issues, through resolution or reduction of the issues in dispute, in approximately 80% of cases. This does indicate that for 20% of cases there was no change. For cases with at least one Indigenous participant, the degree of reduction or resolution of the issues was slightly higher than for non-Indigenous families. Indigenous families also generally experienced the conferences positively, felt that the right people had been invited, and felt that the conference processes were more appropriate than Preliminary Conferences and Children's Court hearings. Finally, the evaluators also identified two issues which would need to be addressed in future. First, they found that the current model whereby the Bidura Children's Court refers care matters to external mediation and all other Children's Courts refer matters to a DRC is not sustainable. The evaluation recommended that court-referred ADR should continue to be supported and that a decision needs to be made about the expansion of the Legal Aid Pilot to other Children's Court locations, noting that whatever model is adopted, certain conditions should be met. The observation was made that this would require a decision on how the two forms of ADR can work together. Second, they noted that although in many cases DRCs or the Legal Aid Pilots were able to resolve issues and disputes about contact between the child and parents, a large proportion of contact issues were not resolved by

ADR. This highlights the need for an appropriate review mechanism for resolving contact disputes when ADR is unsuccessful.

The evaluators presented some evidence to support the claim that ADR is cost-effective, but concluded it was not strong enough to justify ADR solely on the basis of the savings it generates. In both pilots, the average total fees paid for actual court time appears to have been lower for clients involved in matters that were referred to ADR, which would suggest that the length of time that practitioners (and clients) spend in court was less. The introduction of DRCs also appears to have contributed to a reduction in the proportion of matters that resulted in a hearing. But, other costs associated with the process mean that it is not possible to confidently assert that these ADR strategies reduced overall costs. The justification for ADR lies less in its direct savings than in the fact that it is relatively cost efficient in delivering benefits to parents and families that the traditional court process cannot deliver.

4.5.5 How sound are the findings?

The ADR evaluation was thorough, and the evaluator's conclusions were reasonable given the evidence available. There were some external factors that impacted on the strength of the conclusions, including timing, data quality, sample sizes and sampling methods.

With respect to timing, it was not possible to assess the impact of ADR on the number of appeals or applications under s90 of the *Care Act*. This is due to the delay between the ADR process and any future appeals or applications. To assess this would require a longer-term evaluation of the impact of ADR (which the evaluators recommended). Such an evaluation would ideally also cover the longer-term impact of ADR on care matters, including the impact on costs to the NSW Children's Court, Legal Aid and Community Services.

The quality of quantitative data was impacted by the absence of effective systems for recording relevant information and the differences in data collected from one court, and from one pilot, to the next. For example, data on the proportion of matters that were referred to ADR during the evaluation had to be collected via a court file audit review. In addition, the mediators (Legal Aid Pilot) and Children's Registrars (DRCs) were using different post-conference reports until June 2011, leading to inconsistencies in the recording of the issues being discussed and the results of the conference. This impacted on the robustness of findings regarding the extent to which issues in dispute were narrowed or resolved through a matter being referred to ADR. As such, care needs to be taken when drawing direct comparisons between DRCs and the Legal Aid Pilot. A separate but related limitation arose from post-conference reports, which only record outcomes that have been achieved by the end of the conference and not those which were resolved in the period following the conference itself; it is possible, in other words, that the conferences were even more successful than the data show.

Several aspects of the evaluation were constrained by small sample sizes or potential for selective sampling, and several specific findings must therefore be considered provisional pending further confirmation. One example is the voluntary participant survey: the evaluators noted that parents may have been less likely to complete the survey when they were unhappy or upset by the outcome, and that the results may over-estimate the level of satisfaction among all participants. A second example is the professional surveys: some professionals were involved in a large number of conferences during the evaluation, and

may have completed multiple surveys. It is possible, therefore, that a few individual participant may have had a disproportionate impact on the overall survey response, particularly in the Legal Aid Pilot.

Finally, the highly localised nature of the Legal Aid Pilot means that its applicability to other locations or populations cannot be taken for granted. It is possible that similar mechanisms would operate differently when different officers of the court, child protection staff, or target populations were involved. This was not the case for the DRCs as the DRC pilot was run at Children's Courts across the state.

4.5.6 Conclusion

On the grounds that the DRCs and Legal Aid pilot had both delivered a range of positive outcomes to the parents and families involved in care proceedings in a relatively efficient manner, the evaluators supported the continued use of ADR in care and protection proceedings in the NSW Children's Court. There was a high degree of satisfaction among participants in the process, and these aspects of ADR were experienced particularly positively by Indigenous families. In addition, beneficial effects of conferences upon resolving or narrowing issues in dispute was indicated by the evaluation. The evaluation found a wide divergence of views concerning the aspects of cases that made resolution or narrowing of issues more or less likely. The Children's Court considers this evidence that all types of cases can potentially benefit from the ADR process, although further research on this might help refine referral processes to reduce the likelihood of referral for cases where improvement is unlikely to occur.

4.6 Family Group Conferencing

Family Group Conferencing (FGC) is the third Alternative Dispute Resolution strategy proposed by the Expert Working Party (see above). The Wood Report noted the lack of use of ADR, including Family Group Conferencing, despite an effective model and training program having been developed by UnitingCare Burnside, and speculated that this might be due to the increasingly "interventionist and forensic focus" in contemporary child protection practice.³⁷ In response, it recommended that sufficient funding be made available to implement ADR prior to care proceedings, with a specific focus on "placement plans, contact arrangements, treatment interventions".³⁸

FGC is designed to empower families to take greater ownership of solutions to the risk or harm to their children. Conferences are directed by an external (non-aligned) facilitator and attended by family and extended family members, the children and young people involved, as well as child protection caseworkers and their managers. The endorsed FGC model was based on the Institute of Family Practice model of FGS developed by UnitingCare Burnside (2007) and was originally based on the FGC model used in New Zealand since 1989.

³⁷Wood (2008). *Report* ["Wood Report"], Vol. 2 State of NSW through the Special Commission of Inquiry into Child Protection Services in NSW, Sydney, p. 483. Retrieved on 1 May 2012 from http://www.dpc.nsw.gov.au/_data/assets/pdf_file/0011/33797/Volume_2_-_Special_Commission_of_Inquiry_into_Child_Protection_Services_in_New_South_Wales.pdf.

³⁸Wood Report, p. 491 Note that some of these outcomes are outside the remit of FGC but of the broader ADR process.

4.6.1 *What was done?*

FGC was implemented as a trial through two FACS regions: Metro-Central (the Burwood, Central Sydney, Chatswood, Eastern Sydney, Epping, Lakemba, St George and Sutherland Community Service Centres) and Northern (the Ballina and the Clarence Valley CSCs). The Northern region trial also accepted referrals from the Tamworth CSC partway through the process. External facilitators were utilised as per existing FGC models, and neutral venues were used for meetings. The purpose of FGC was to provide families with an opportunity to arrive at solutions to the issues identified by Community Services. The intention of conferences was for families to develop their own Family Plan, which were formally recognised when endorsed by Community Services staff. Implementation was supported by a procedures manual.

In the Northern Region the UnitingCare Burnside Institute of Family Practice (IFP) was engaged by Community Services to provide an initial facilitator as well as develop a pool of suitable and trained individuals who could conduct conferences once the IFP contract had been completed. A similar model was used in the Metro-Central region, though the FGC project officer was primarily responsible for this process.

The FGC model is highly structured from the referral through to implementation. Referrals for FGC could be from child protection, out-of-home-care or restoration teams within Community Services, based upon eligibility criteria. Following referral, consent from the family was required, and a discussion with the family members identifying the issues that needed to be addressed was undertaken. When a matter was accepted for a conference, a conference facilitator was engaged who was then responsible for the majority of the conference preparation activities. This included reviewing the initial referral, contacting family members and all other people who proposed to attend the conference, managing the logistics of the meeting and advocating for the use of brokerage funds in order to support the attendance of family members, if required, and ensuring all participants understood what to expect at the conference and the expectations on them. Where children were not going to be present at a conference, it was the role of the facilitator and other support workers to understand the child's position regarding the issues and their wishes and to present these at the meeting.

The model for the actual conferences involved three specific components – Introduction and Information Sharing, Family Time, Ratification of the Family Plan. The facilitator was charged with managing these phases as well as supporting communication between all parties. The Family Plan was the immediate outcome of the conference. It was required to address the non-negotiable aspects of the Community Services concerns while also encouraging families to consider the broader issues that were impacting on their situation. Family Plans were primarily developed by the family but were required to be ratified by Community Services representatives present at the meeting. These plans were formally and informally monitored over time to ensure that the issues were being addressed as planned, or protective strategies were adhered to. The formal review process involved reconvening the FGC participants and reviewing the Family Plan actions, while the informal mechanism was the nomination of a Family Plan review or contact person (not the facilitator).

4.6.2 *How was it evaluated?*

The Australian Institute of Criminology undertook evaluation of all aspects of the ADR trial, including Family Group Conferencing³⁹. As for the other two aspects of the trial (the Legal Aid Pilot and the Dispute Resolution Conferences) the design for the FGC evaluation was of processes and outcomes from the project. Specifically for FGC the evaluation methods included:

- the development of a program logic model and evaluation framework;
- a review of evaluation evidence regarding similar programs in Australia and overseas;
- conference observations;
- interviews with parents and family members who participated in a conference;
- interviews and focus groups with professionals involved in FGC;
- an online survey of professionals involved in the pilot;
- the preparation of case summaries based on participant records; and
- the analysis of administrative data, including data recorded by the Project Officer, data extracted from hardcopy records and data extracted from the Key information and Directory System (KiDS) for the intervention and comparison groups.

4.6.3 *What did the evaluation(s) find?*

The evaluation showed that referrals for FGC, and the numbers of actual conferences held, was low. During the evaluation period February 2011 to March 2012 there were 59 referrals for FGC of which 29 proceeded to a conference. Reasons for this were shared among systems issues such as narrow referral criteria and convoluted referral processes that were discouraging to caseworkers, organisational culture including lack of knowledge of the project, possible uncertainties or mistrust felt by families, and conference timeliness especially with regard to court processes that might be undertaken for the care and protection of the children.

Conferences were experienced positively by most parties, and led to the development of family plans in the majority of cases. Families experienced the conferences as less adversarial and respectful, while the professionals felt that the conferences assisted their working relationships with families and engaging a broad range of family members. For Indigenous participants, FGC processes were experienced as a more culturally appropriate means of decision making and a positive way to engage broad family networks. Facilitators were recognised as providing effective and culturally appropriate conference processes, and were generally seen as skilled and effective at engagement, even where there were previous negative experiences between the parties. The majority of family members, and almost all Community Services staff (between 95 and 100%), reported being suitably prepared for the conference and knew what to expect.

The evaluation found various indicators of success in the FGC trial at the development and implementation stages. Family Plans were satisfactorily developed and accepted in 26 of the

³⁹ The KTS Action Plan states in 11.1.(x) that the amendments to Contact Orders will not commence until parties have access to appropriate dispute resolution arrangements and access to administrative reviews. Thus implementation of recommendation 11.1 (x) is awaiting this evaluation

29 conferences. The plans were seen to be realistic for the family at the same time as being focussed on the safety and wellbeing of the children. When required, the FGC facilitator worked with the parties present to achieve a plan that suited the family and addressed Community Service staff concerns. High levels of conflict within the family was the principal barrier to development of a Family Plan in the three other matters. Nine cases were subsequently reviewed for degree to which actions and main goals had been started or achieved. It was found that there was at least some focused activity in eight of the nine cases. Where goals were not achieved, reported reasons included the inability to access support services that had been identified and, some actions appeared too ambitious or unrealistic.

In contrast to these positive findings, the evaluators also identified areas where the implementation deviated from the plan, and areas for improvement in the FGC processes. Three of these are discussed below. First, and most significantly, there was a lack of clarity regarding processes and responsibilities for ongoing monitoring and management of Family Plans. By the end of the evaluation period, review meetings had been conducted for only 16 of the 26 conferences at which a plan had been developed. [It is noted that the actual incidence of review meetings may be higher, because some cases with family plans may have been scheduled or undertaken outside of the evaluation period.] The evaluation found that the effectiveness of the review processes was variable due to lack of clarity regarding the process for undertaking the review and how to respond when some parties did not attend. This lack of an effective review mechanism impacted upon the potential for formal review of the implementation of Family Plans. On a related note, fifteen of the 26 Family Plans identified a review or contact person, but there was no formal review of the effectiveness of that role.

Timeliness was a second significant issue identified by the evaluation. In the vast majority of cases, conferences were not implemented within the recommended timeframe of six weeks from the date of referral – a timeframe adopted to reduce the potential for family circumstances to change. Conferences generally took between 10 and 14 weeks to complete. The primary reasons identified for this length of time were lack of availability of participants, and the volume of preparation work required. These delays may be responsible for the belief among community services staff identified by evaluators that conferences were not suitable for pressing child protection matters. On a separate note, conferences generally took less time to run than anticipated but a finding of the stakeholder survey was that the majority of participants (17 out of 19) felt the duration of the conferences was “just right”.

A third barrier to successful implementation identified through stakeholder feedback concerned interactions between FGC and the new ROSH threshold. In essence, the new threshold has the effect of restricting the matters allocated to a caseworker to the most serious and/or complex cases which are more likely to proceed to court rather than be dealt with informally. Some stakeholders indicated that this might make them unsuitable for FGC. It is not necessarily clear that this claim can be accepted at face value – the New Zealand model on which the FGC trial was based requires all matters to be referred to FGC prior to any court action, and responding to complex child protection cases through FGC is common practice.

4.6.4 *How sound are the findings?*

The evaluation provided some evidence suggesting FGC is viewed positively by participants. The evaluation model of initial and follow-up interviews, online survey and focus groups was an effective strategy for gauging experiences of parties to the FGC trial. That this model was consistent across the three ADR components provides an effective evidence base for this aspect of the Wood Report intentions and KTS Action.

Unfortunately, however, the validity and reliability of evidence concerning other aspects of the initiative were somewhat weaker due to the available assessment methods and problems with data. With respect to referrals and activities, sufficient data were gathered but their reliability was compromised by inconsistent use of the Family Plan review tool. Assessments of Family Plan goal achievement were hampered by lack of suitable data, as were assessments of the impact of FGC on child wellbeing and future child protection concerns. The low number of referrals, the brief timeframe under which families could be followed and potential for different readings of the indicators all fed into the lack of statistical power for these aspects of the evaluation.

Using a model that identified the likely ongoing costs for FGC (e.g. costs associated with facilitators, costs of having caseworkers and managers attend conferences) and dividing this by the number of matters that proceeded to conferences, the unit cost of each conference in this trial was found to be \$8,695. There are no data available to compare this to the savings or additional costs that this represents within child protection processes.

4.6.5 *Conclusions*

In general, the FGC trial was implemented well, and the evaluation identified that all participant groups had overwhelmingly positive experiences (“a high level of satisfaction”) of the process and of facilitators. In some cases this experience turned reluctant participants into advocates for the model. The effectiveness of FGC in providing a more appropriate model for dealing with issues in the child protection context is a standout feature of this trial and the evaluation report. FGC also appears to be particularly successful in ensuring culturally appropriate practices, according to reports from both family members and professional stakeholders. The Family Plans developed in these conferences were positively received, appeared to support cultural identity and, where appropriate, satisfied the Aboriginal Placement Principles of the *Children and Young Persons (Care and Protection) Act 1998*.

While the positive experiences of participants in the FGC trial is clearly established, and the better cultural fit for Indigenous participants in this model, there were clear implementation issues that reduced the numbers of conferences held and the monitoring and review of actions. Though the evaluation sought to identify outcomes from the model, a lack of data regarding the impacts of FGC on child wellbeing and future child protection concerns was not sufficient to draw any meaningful conclusions.

4.7 **Screening and Response Priority Tools**

A key finding of the Wood Report was that Department of Community Services was weighed down by responding to a range of reports that did not warrant a statutory response – “Too many reports are being made to DoCS which do not warrant the exercise of its considerable statutory powers”. This finding led to one of the Wood Report’s key recommendations –

raising the threshold for reporting such that only those children and/or young people suspected of being at risk of **significant** harm would be reported to that agency. The Wood Report also recognised that raising the reporting threshold introduced a further problem, that of consistently interpreting and applying the notion of ‘significant harm’ in the complex situations encountered by mandatory reporters day to day. Consequently, Justice Wood also recommended “clear guidelines for classifying risk of harm reports made and information given to the Helpline” and “sufficient quality training and guidelines” to equip workers with the resources to make such judgements.⁴⁰

In response to these recommendations, Community Services identified the Structured Decision Making (SDM) system tools as an appropriate mechanism for improving the consistency of reports and the handling of those reports at the initial point of contact. SDM Tools are designed to make best use of the information available at specific assessment points, by relying on actuarial science and comprehensive research into factors known to be most relevant to the safety, welfare and wellbeing of children. SDM tools are developed by the Children’s Research Centre (CRC) in the USA, and are used in various US jurisdictions, including in Queensland, the Northern Territory and South Australia.

4.7.1 What did the project set out to do?

Community Services engaged with the CRC to customise its Structured Decision Making Tools for use in NSW. These were the *Mandatory Reporter Guide* (an online decision making-tree for use by mandatory reporters in the community to decide whether a report should be made to Community Services), the *Structured Decision Making (SDM) Screening Tool*, and the *Response Priority Tool*. The SDM Screening Tool is implemented at the point at which the Community Services child protection Helpline is contacted and concerns are raised about the safety, welfare or wellbeing of one or more children or young people. The Screening Tool is implemented in order to assess whether the issues raised meet the threshold for Risk of Significant Harm. Where the threshold is not reached, the report is ‘screened out’ and does not proceed further in the Community Services system. When a report is ‘screened in’ for a child protection response the Response Priority Tool is applied in order to assess whether the situation requires a Community Services response within 24 hours, 72 hours, or 10 days. The CRC’s model includes localising the tool to specific conditions of the jurisdiction within which they are being implemented, training in their use, training for supervisors, and review of implementation quality.

4.7.2 What was done?

Adoption of the SDM Tools began with an initial field test in order to assess worker satisfaction and to assist with amending the tools to suit local conditions. Prior to rollout training in using the tools was conducted for all Helpline staff and their Team Leaders and Managers. Team Leaders also received training in supervision of staff utilising the tools. There were also numerous internal support and quality assurance strategies developed. These included the availability of dedicated, specialist SDM advice personnel, including Managers and Learning and Development staff, the development of “advanced learning

⁴⁰Department of Human Services - Community Services (2010). *Monitoring and Review of the Implementation of the Structured Decision Making (SDM®) Screening and Response Priority Tools at the Child Protection Helpline* ["SDM Review"]. Unpublished document held on file by Department of Human Services, Sydney.

scenarios” based on complex cases encountered at the Helpline, thematic review of types of reports staff found more complex, with strategies to respond to the complexity experienced, regular case readings by senior staff and managers, and the development of business rules and/or practice guidance regarding a number of specific case-related scenarios. External quality review strategies are an integral component of the CRC model. In the case of the NSW implementation the review strategies involved an inter-rater reliability study, where 12 case vignettes were screened and prioritised by both CRC and 47 Community Services Helpline staff to measure consistency and accuracy in the application of the tools. Formal case reading was another activity which occurred at 6 month intervals where, and case readings were undertaken by CRC and Community Services Helpline staff and the outcomes compared and measured in order to identify areas requiring further support, training or system development. The information developed through these processes was also fed into this evaluative aspect of the project.

4.7.3 How was it evaluated?

Where other evaluations of KTS projects have attempted to consider the effectiveness of the service model, due to the existing evidence supporting the validity of the SDM Tools, the implementers did not seek to test their general accuracy or appropriateness to child protection situations. While the project focussed on ensuring quality of implementation and providing support to staff using the tools, the evaluative aspects involved a review of Community Service data regarding reports to the Helpline, undertaking qualitative evaluation processes – specifically, semi-structured interviews – and an external review processes to assess the degree of uptake and assessment quality. Aspects of this latter component were repeated four times between April 2010 and February 2012.

The statistical component of the evaluation was undertaken for a roughly three month period in 2010 (from 24th January to 30th April), assessing the volume and pathways of contacts at the Helpline. The data for 2010 was compared against the same period in 2009 in order to evaluate the report volume change. This was undertaken internally by Community Services. The experiences and satisfaction of Helpline staff implementing the tools was assessed by the Community Services evaluation team through semi-structured interviews. These interviews were undertaken with caseworkers, team leaders and managers.

The external review components were undertaken in collaboration with the CRC. The first inter-rater reliability aspect involved 47 Helpline staff – caseworkers and team leaders – each independently reading and rating 12 case vignettes. This sought to assess the degree of worker consistency with regard to the assessment decision, the ROSH subcategories that were (or were not) used, the identification of parenting issues, and the response priority allocated. CRC staff assessed the assessments against these criteria, utilising a 75% similarity as the benchmark for satisfactory reliability.

Case reading was undertaken independently by the CRC. In the first round a total of 52 cases were assessed across three dates in March 2010. Purposive (as opposed to random) sampling was used to provide cases that had undergone numerous readings and assessments by Helpline staff. CRC review of cases focussed on the degree to which tools were used correctly and assessments reached the correct conclusions. This involved reviewing the degree to which caseworkers correctly identified the ROSH subcategories in the screening decision, correctly identified parental risk factors, and correctly completed the

Response Priority Tool. In addition, this CRC case reading process attended to the quality of case review provided by Team Leaders and Managers. This involved assessing the degree to which mistakes made by caseworkers were identified by Team Leaders and Managers.

Later rounds of the CRC case reading process involved slightly higher numbers of cases (specifically, the February 2011 review involved 60 cases and the February 2012 review involved 62) and used a random rather than purposive sampling of cases.

4.7.4 What did the evaluation(s) find?

Community Services system data provides detailed information on the significant reduction in numbers of reports to Community Services after the raising of the child protection reporting threshold, of which the SDM Tools are a vital component. Data comparing similar periods in 2009 and 2010 shows that for the three month period 24th January to 30th April 2010 phone calls to the child protection Helpline were reduced by 22% and total Helpline demand from all sources was reduced by 27%. Contacts that indicated concerns regarding a child or young person were reduced by 32%. Reports that met the threshold of ROSH were 53% lower than reports transferred to a CSC/JIRT for further assessment during the same period in 2009⁴¹. Within these changes the data shows that the spread of ROSH reports for different age groups and for Indigenous status were similar.

Changes in child protection reporting as a result of the raising of the threshold is also reflected in the response rates allocated to reports that were 'screened in' at the Helpline. Before and after the KTS changes the potential response times were either <24 hours, <72 hours, and less than 10 days. Data from that period shows a significant numerical decrease as well as a shift in emphasis towards reports with greater degree of imminent risk to the subject child or young person. This indicates the possibility of the SCRPT tools being beneficial, though this positive relationship cannot be demonstrated here.

Aspects of the evaluation that addressed SCRPT directly were the satisfaction analysis conducted by Community Services through semi-structured interviews with Helpline staff, and the inter-rater review and case reading undertaken by CRC. The evaluation reported that the Helpline staff who were interviewed – caseworkers, team leaders and managers – had all reported positive experiences with the SDM Tools and changes associated with KTS. They reported that the tools supported their professional practice and allowed an opportunity to “exercise their clinical judgement”. It was identified, though, that the time taken to complete the SDM Tools was up to 45 minutes longer than the previous assessment process. This may be mitigated to some degree by a reduction in time spent waiting for a team leader to review the screening and response decisions. At the same time, they raised the issue that for some child protection situations the definitions provided by the tools were either unclear or in conflict with other aspects of the system. For example, the definition for sexual assault was not consistent with that used by JIRT.

The training provided to staff was also experienced positively. Specifically, managers identified that the timing of the training – occurring in November prior to the January implementation – allowed them to feel confident to answer caseworker questions at the point of using the tools.

⁴¹ As the category of ROSH did not exist in 2009 the closest equivalent measure is those reports that were assessed at the Helpline as needing further assessment through a local child protection office (CSC or JIRT).

The two CRC standard evaluative components of the implementation quality – the inter-rater reliability and the case reading – were implemented at the time and in the manner intended in the evaluation plan. These processes address consistency in decision-making among Helpline staff and quality of assessments.

The April 2010 inter-rater reliability process found positive results in the consistency of staff identification of issues of concern in the 12 case vignettes reviewed. When taken as a whole (responses averaged across the 12 case vignettes), the degree of Screening Tool agreement regarding ROSH subcategories was greater than 87%; for each of the parent risk factors the degree of agreement was greater than 91%; for the screening decision, the degree of agreement was 87%. When reviewed individually, three of the 12 cases had an agreement level of less than 75% (below the benchmark of 75%), while 8 of the 12 cases were screened with a consistency of greater than 89%.

In contrast, that same review found that the consistency level for the Response Priority Tool did not meet the threshold at any point. Eleven of the 12 cases had a consistency score of 60% while the final case had only 40% consistency. The report identified three possible explanations for the observed degree of inconsistency including the use of scenarios with limited information, the failure of some staff to use the definitions, and the lack of clarity in some of the definitions themselves. Immediate remediation action was recommended including a thorough review of the definitions that appeared problematic and refresher training on the use of the definitions. Both recommendations were actioned and the data from subsequent case reading shows a marked improvement in accuracy and consistency of both screening and response priority decisions.

The case reading strategy in February 2010 involved purposive sampling of 52 cases each of which had been completed by Helpline staff and reviewed by approximately four fifths of Helpline team leaders and managers. In this way the CRC reviewers had the greatest potential for comment upon the quality of review by team leaders and managers, as well as identifying the degree of accuracy of caseworker assessment.

The case reading process found generally positive results but some areas that could be improved. The CRC found that the Helpline caseworkers and team leaders were more likely to 'screen in' reports than the CRC case reviewers would recommend. Where cases were screened in, the correct ROSH sub-category was identified in the "vast majority" of cases. Parental risk factors were correctly identified in 79% of cases, though there was a trend of identifying parent risk factors without corresponding evidence within the case narrative. The degree of correct completion of the Response Priority Tool was 78%, which is much lower than the benchmark level of 90 to 95% sought by Community Services. Regarding case review undertaken by team leaders the CRC found that there was variability in the level of detail attended to by team leaders and managers.

Subsequent to this review a range of quality improvement measures were implemented. These included revising problematic definitions, providing refresher training, and ensuring Helpline staff undertook the screening and response allocation processes as two distinct steps rather than the former leading to a preconceived outcome. Extra support for Team Leaders and Managers was also provided, including access to in-house expert consultations from (casework specialists and the Principal Policy Officer) and regular feedback on the accuracy of screening and response priority decisions they approved for individual cases.

The fourth CRC review of SCRPT tools implementation was undertaken in February 2012. This review focused on correct decision making by Helpline staff and allocation of response time. The review was undertaken on 62 cases that had been reviewed by Helpline staff. In contrast to previous reviews this round did not deliberately select cases that had been reviewed by managers or team leaders but utilised a random selection process. As per previous reviews the CRC applied their own case reading tool to determine the quality of the decisions made by Helpline staff.

The review found an almost 90% consistency between Helpline and CRC decisions regarding cases screening in and screening out. The primary discrepancy here was cases being screened in which the CRC identified should have been screened out. This issue has persisted across time as reflected in the first case review process. There was a very high rate of correct decision making (as demonstrated in Table 3 – Percentage Agreement Between Helpline and CRC Staff by Case Reading Measure), as well as response time allocation. Response allocation time was correct in 87% of cases with disagreement in 6 individual cases. There was no specific pattern to these differences.

Table 3 – Percentage Agreement Between Helpline and CRC Staff by Case Reading Measure

<i>Case Reading Measure</i>	Percent Cases Handled Correctly (n = 62)
<i>Screening Decision Correct</i>	89%
<i>Correct ROSH Sub-Category Selected</i>	85%
<i>Parental Risk Factors Correctly Selected</i>	92%
<i>YPAR Used Correctly</i>	98%
<i>CS Additional Action Correctly Identified</i>	97%
<i>RP Decision Correct⁴²</i>	87%

Of particular note in this review, is the lower level of correct handling of cases in the ROSH Sub-Category Selected. This difference is rooted in the CRC assessment of incorrect screening decision. Essentially, those cases screened incorrectly will, by definition, have a ROSH Sub-Category selected (or not) that should not have been.

Longitudinal implementation of Helpline case reading by CRC can be used for monitoring improvement over time. Data from CRC (as presented in Table 4 – Comparative Case Reading Results Across Time) shows that practice by Community Services Helpline staff has steadily and significantly improved from March 2010 to February 2012.

Table 4 – Comparative Case Reading Results Across Time

<i>Measure</i>	<i>March 2010 (n=52)</i>	<i>July 2010 (n=50)</i>	<i>Feb 2011 (n=60)</i>	<i>Feb 2012 (n = 62)</i>
Screening Decision Correct	65%	84%	92%	89%
Correct ROSH Sub-Category Selected	63%	74%	87%	85%
Parental Risk Factors Correctly Selected	79%	78%	90%	92%
YPAR Used Correctly	90%	92%	97%	98%

⁴² Includes only those cases that were screened in by Helpline and thus had a Response Priority completed.

FACS Additional Action Correctly Identified	92%	98%	97%	97%
RP Decision Correct⁴³	78%	79%	83%	87%

This data shows that Community Services is approaching its target benchmark of 90 to 95% correct results in virtually all quality areas of SCRPT tool implementation and suggests that quality improvement processes have been effective in addressing the issues originally identified.

4.7.5 *How sound are the findings?*

The evaluation material reported here provided some clear information but by design left some questions unanswered which are interesting in the broader context of KTS. Firstly, the available data clearly showed the difference between pre- and post-KTS reporting patterns, but were not suitable for evaluating the contribution of the SDM Tools as distinct from other changes to the child protection system occurring at the same time. For example, the data evaluation noted the differences between agency-based mandatory reporters (e.g. Police, Health, Education etc. staff) and those who are in the non-Government sector or non-mandatory reporters. The report did not consider the potential impact of the Child Wellbeing Units to which the staff of Government agencies have access while the non-Government and non-mandatory reporters do not. On a similar note, the data provided does not allow the development of a picture of what is occurring to the concerns workers or individuals have about a child or young person but do not reach the threshold of ROSH.

The strategies designed to assess implementation quality have been utilised by the CRC in numerous situations and have the benefit of validation over a long period of time. They are easy to implement and robust for their purposes – assessing degree of consistency among local staff utilising the tools, and accuracy as assessed by the originators of the tools. Although this is not discussed in the evaluation documents, Community Services advises that the results of this process informed other quality improvement exercises. Advanced SDM training (called “One More Question”) was provided in Sept-Oct 2010, and was designed to help workers learn to use the definitions to elicit the relevant information from callers. Reviews of all definitions known to be problematic were conducted in 2010, 2011 and 2012. As a result of these reviews, definitions were updated to provide more clarity for caseworkers.

4.7.6 *Conclusions*

It is important to keep in mind the timing of the child protection data review and first SDM Tools reviews. Specifically, these were undertaken in the first three months post-proclamation of the changes to the child protection reporting threshold. It is reasonable to expect that there will be familiarisation and acculturation difficulties with consequent ongoing improvements in staff use of the tools so long as there are systems designed to support their implementation. The latter CRC review of the SCRPT tools implementation bear out that proposition.

The data showing changes to child protection reporting indicate that the Wood Report recommendation 6.1, intended to reduce the demand on Community Services, has been

⁴³ Includes only those cases that were screened in by Helpline and thus had a Response Priority completed.

achieved. The further important aspect of the recommendation, that there is consistency in assessment of reports, has also been achieved in the vast majority of cases.

4.8 Safe Families

According to the *Breaking the Silence Report*,⁴⁴ child sexual assault within Aboriginal communities is a “huge issue” that has “devastating, and lifelong, effects” on individuals and communities. It is “little understood”, “undetected” and “seldom reported”. Community members suggested that responses to perpetrators should be based on a clear understanding of the problem. The report found that socio-cultural factors, especially those associated with substance abuse, dispossession and fracturing of community, were implicated in its occurrence.⁴⁵ The *NSW Interagency Plan to Tackle Child Sexual Assault in Aboriginal Communities 2006-2011*⁴⁶ was developed in response to this report. It sought to strengthen the service system, undertake community development and improve understanding and reporting of child sexual assault. The Interagency Plan operated as part of the NSW Aboriginal Affairs Plan *Two Ways Together 2003-2012*.⁴⁷

Initiated in 2007/2008, Safe Families is a pilot early intervention program involving a coordinated approach between government agencies and communities to work together to tackle child sexual assault in five Aboriginal communities. Safe Families received enhancement funding under KTS, and forms part of the Action Plan’s commitment to work with Aboriginal children, their families and communities to reduce the number of children coming into contact with the child protection system, and improve support for those children already in the system.⁴⁸ It also responds to the Wood Commission’s extensive investigation of Aboriginal over-representation in child protection.⁴⁹

4.8.1 What did the project set out to do?

Safe Families was a four year trial conducted in five communities starting in April 2009: Bourke, Walgett, Brewarrina, Lightning Ridge and Wilcannia. The following discussion describes implementation in Wilcannia and Lightning Ridge due to their being the two communities on which the formative evaluation focussed.

Safe Families aimed to reduce the incidence of child sexual assault and reduce offending over the long term by building strong, healthy family and community relationships and empowering Aboriginal communities to keep children safe. A particular focus of Safe Families services is those families that do not meet the threshold for Community Services intervention. The partner agencies involved in the trial were New South Wales Health, FaCS-

⁴⁴Aboriginal Child Sexual Assault Taskforce (2006). *Breaking the Silence: Creating the Future* ["Breaking the Silence"] NSW Attorney General’s Department, Sydney. Retrieved on 1 December 2012 from [http://www.lawlink.nsw.gov.au/lawlink/acsat/acsat.nsf/vwFiles/80001%20CP%20Rep-all_sml.pdf/\\$file/80001%20CP%20Rep-all_sml.pdf](http://www.lawlink.nsw.gov.au/lawlink/acsat/acsat.nsf/vwFiles/80001%20CP%20Rep-all_sml.pdf/$file/80001%20CP%20Rep-all_sml.pdf)

⁴⁵*Breaking the Silence*

⁴⁶NSW Government (2007). *New South Wales Interagency Plan To Tackle Child Sexual Assault in Aboriginal Communities* ["Interagency Plan"] Department of Aboriginal Affairs, Sydney. Retrieved on 1 December 2012 from <http://www.daa.nsw.gov.au/data/files//NSW%20Govt%20Plan%20to%20Tackle%20Aboriginal%20Child%20Sexual%20Assault.pdf>

⁴⁷NSW Government (2003). *Two Ways Together. New South Wales Aboriginal Affairs Plan (2003-2012)* ["Two Ways Together"] Department of Aboriginal Affairs, Sydney. Retrieved on 1 December 2012 from http://www.daa.nsw.gov.au/publications/TWT%20CopBroch_LR_1.pdf

⁴⁸*Action Plan*, p. iii.

⁴⁹*Wood Report*, p. 735 and following.

CS, the NSW Police Joint Investigation and Response Team (JIRT) and the Office of Aboriginal Affairs.

Safe Families incorporated a mixture of government service provision and community development. The direct service delivery aspects of the project included the establishment of a co-located cross agency Safe Families team in each pilot location which undertook the delivery of the project as well as provision of case management to families. The community development aspects include relationship building amongst agencies to improve coordination of local service delivery, the formation of a Local Aboriginal Reference Group (LARG) to help identify community priorities to address child sexual assault and develop a local Safe Families Plan in each area, education and awareness-raising to remove barriers to reporting child sexual assault and reduce the incidence of child sexual assault, engagement to build relationships between communities and agencies (including development of a 'Community Story' to enable the Safe Families team to learn about the community), and informal practice with community members and groups to increase awareness of child sexual assault.

Safe Families, as with all aspects of the *Interagency Plan*, acknowledged the limited evidence base regarding addressing child sexual assault in Aboriginal communities and sought, through a commitment to evaluation, to build further evidence about what works. The evaluators explicitly drew attention to the fact the project was conducted on an "action learning model", using ongoing feedback to address some of these limitations within the evidence base.

4.8.2 What was done?

The Safe Families office opened in Wilcannia in June 2009, and by the time of the evaluation most local positions had been filled and the office was fully operational. The evaluators found the Wilcannia team were engaging in a range of activities designed to build relationships with local agencies. The team was also engaging with the community through activities such as community barbecues, providing information and education through school and community based activities and undertaking integrated case management with children and families. A number of specific activities were cited, including working on the establishment of the Local Aboriginal Reference Group (expressions of Interest had been called in December 2009 and probity checks were underway), and conducting 'informal practice' with individuals and families to increase awareness of CSA (such as supporting the Wilcannia Central School breakfast program one day per week, and encouraging drop-ins to the SF office). Formal case management was being undertaken with families who did not meet the threshold for Community Services intervention.

The Lightning Ridge Safe Families office opened in October 2009, and at the time of the evaluation all local positions had been filled. The evaluators found that the Lightning Ridge team was working on the establishment of the Local Aboriginal Reference Group (a community meeting to establish the group had been held and applications forms distributed), holding regular Elders morning teas at the Safe Families office and engaging with local agencies such as the Women's Legal Service, schools, youth services and the Safe House. No clients had been seen for formal services by this team, and there was little evidence of formal community information and education activities.

The Safe Families program was formally concluded on 30 June 2012, but as there was a significant underspend from previous years the program was extended for a further 12 months. Based, in part, on the evaluation findings discussed below, a new model was developed and rolled out in the Safe Families sites. It includes:

- Local Aboriginal Reference Groups (LARGs) helping to identify community priorities to address child sexual assault and developing local prevention plans;
- Case Coordination Groups to provide case management support to individuals and families where Aboriginal children and young people are at risk;
- Issues Panels, chaired by a local NGO, overseeing the child sexual assault prevention plan developed by the LARG; and
- The Regional Aboriginal Child Sexual Assault Group providing leadership and direction to the program, with high level matters escalated to the General Manager, Aboriginal Affairs, if required.

4.8.3 How was it evaluated?

Kristine Battye Consulting was engaged by Aboriginal Affairs to review the implementation and early operations of Safe Families in Wilcannia and Lightning Ridge, and Focus Communities in Nowra and Toomelah-Boggabilla. As the Focus Communities initiative did not receive KTS funding, information pertaining to Focus Communities has not been included in this synthesis. This synthesis reports on the Stage 1 evaluation as this is the component which has been completed to date.

In conducting their evaluation, the evaluators adopted an empowerment approach that attended to community relationships with evaluators and sought to “share power, control and decision making” with stakeholders at both the contracting and community levels.⁵⁰ Although the evaluation does not explicitly justify this approach with reference to standards such as the *Ethical principles and guidelines for indigenous research*,⁵¹ it is consistent with good practice in evaluation with Aboriginal people. Participatory evaluation practices are widely understood to strengthen interventions and support effective evaluation.⁵² This is in addition to the potential for mixed method evaluation designs to provide a broader evidence base for the evaluation of programs – especially those that are complex and related to “complex interventions”.

In concrete terms, the evaluators engaged in review and negotiation with the DAA as the contracting agency, and sought guidance from the local communities on conducting the evaluation and including the local community in its implementation. They also developed an evaluation method that incorporated both qualitative and quantitative research methods. The report of the Stage 1 review was based on a review of relevant literature, an analysis of the program logic for Safe Families, qualitative data collection via meet and greet site visits, semi-structured interviews with partner agency managers, small group meetings with Safe Families staff and Aboriginal governance groups, ‘yarning sessions’ for service users. It also

⁵⁰Kristine Battye Consulting (2011). *Evaluation of Program Initiatives Tackling Child Sexual Assault in Aboriginal Communities. Stage 1 Report* ["Safe Families Evaluation"]. Unpublished document held on file by Aboriginal Affairs NSW, Sydney, p. 64.

⁵¹AIATSIS (2011). *Guidelines for Ethical Research in Australian Indigenous Studies* ["Guidelines for Ethical Research in Australian Indigenous Studies"]. Australian Institute of Aboriginal and Torres Strait Islander Studies, Canberra.

⁵²Mikhailovich, Morrison and Arabena (2007). "Evaluating Australian Indigenous community health promotion initiatives: a selective review" ["Evaluating Health Promotion Initiatives"]. *Rural and Remote Health* 7:746.

incorporated analysis of quantitative data regarding relationships between partner agencies (via a partner agency survey), reporting of child sexual assault (provided by the NSW Ombudsman), community willingness to report child sexual assault (via yarning sessions), and output/operational data (provided by Aboriginal Affairs) were also included in the evaluation plan.

4.8.4 What did the evaluation(s) find?

The evaluation found the overall approach of the initiative was broadly in line with the available evidence regarding the contributors to successful interventions aimed at reducing and preventing child sexual assault in Aboriginal communities, and incorporated an Aboriginal worldview regarding the problem. It also identified key areas where the model deviated from the evidence. First, with respect to project management, the evaluators highlighted that the initiative had unrealistic timeframes for development and implementation, and that some of its aims had no corresponding strategies. Second, they identified two resource constraints: funding was allocated on a medium-term basis (in this case, for 2008/09-2011/12) rather than on a sustainable long term basis; and workforce development was limited (there was insufficient training, while the designated lead agency had limited experience in program development, implementation, and service delivery). Third, the evaluators suggested that attention was given to the “symptoms of the problem more than the cause” and that the mix of services offered may not have been appropriate: insufficient attention was paid to the perpetrators of child sexual assault, and rehabilitation programs were not available as part of the initiative; nor was there sufficient attention to providing culturally appropriate parenting education, or strategies to redefine cultural identity, such as through community healing activities.

Regarding implementation, the evaluation focussed its discussion on two areas – implementation and indicators of outcomes. The evaluators found that effective implementation had been impeded by a number of structural factors which led to issues such as lack of activity at the local level and lack of “ownership” of child sexual assault within these communities. The evaluators found that governance structures were complex and ineffective, and that workers lacked an understanding of the role of the lead agency. In addition, interagency staff in local offices operated as separate entities housed under one roof rather than building a new culture based on shared approach to the work. Regarding the model incorporating an Aboriginal worldview into the response to CSA, it was reported in the evaluation that this had not been reflected in the implementation.

There was poor coordination between agencies at the planning stage that lead to reduction in service options. For example, this impacted upon involvement of Safe House in both Wilcannia and Lightning Ridge. The evaluators also reported that the original plan anticipated NSW Health would provide counselling and possibly other clinical service delivery, but that this did not occur and the contribution of the local Aboriginal Health Worker was limited to case management. More generally, they found that capacity/accessibility issues relating to key services (such as drug and alcohol services and sexual assault counselling services) had not been addressed.

The evaluators identified that there was a lengthy delay between service commencement and the receipt of the first clients in both trial sites and suggested that this may have been due (at least in part) to high staff turnover, including senior management. This was a consistent problem in both locations, and impacted heavily on the relatively inexperienced

Community Services staff that had been appointed to the team. These problems were exacerbated by the fact that Safe Families staff, although located under the one roof, essentially still worked within the budgets, pay schedules, policy, procedures and reporting framework of their home organisations – and these varied considerably between the three organisations. In addition to staffing, the initiative experienced problems with facilities. The time taken to acquire and renovate office space delayed the start of the initiative in Lightning Ridge, and even once this was achieved there was limited accommodation for staff and no appropriate space for group work (e.g., parenting classes).

Conversely, the evaluators found that governance arrangements within each of the field offices worked reasonably well, somewhat more so in Wilcannia than in Lightning Ridge. They found that communication between the local Safe Families staff was generally effective. Staff appointed to Wilcannia were largely work-ready, passionate and eager to work together, and that the flat organisational structure adopted in that office helped create a unified team. They also found that the Safe Families team had developed productive relationships with youth workers and the Lightning Ridge Youth Centre. Nevertheless, they concluded that Safe Families had failed to build a new culture of collaboration amongst workers, and that the “flat” team structure meant that there was no single person with ultimate responsibility for the site if difficulties arose – although the situation improved with the appointment of the dedicated Departmental staff at the regional level. In general, service providers expressed confusion about the role of Safe Families within the broader service landscape – confusion which appears to have been exacerbated by the simultaneous implementation of another, unrelated initiative in Wilcannia.

The evaluation found that there were disparate activities and mixed degrees of effectiveness regarding stakeholder engagement. Initial components of the project required developing a Local Aboriginal Reference Group and a local Safe Families Plan at each site. In both locations the establishment of the Aboriginal Reference Groups had not been completed by the time of the evaluation, and the communities had not been significantly involved in service planning, though some community input was gathered in Wilcannia through regular Elders morning teas held at the Safe Families office. Perhaps reflecting this lack of stakeholder engagement, many community members reportedly viewed the Safe Families office as the local Community Services office, and this confusion appeared to pose a barrier to uptake of Safe Families services. The evaluators did note, however, that the Wilcannia team appeared to be achieving a more culturally appropriate way of working with their clients than Lightning Ridge.

Data regarding specific indicators was gathered through surveys, interviews, focus groups and yarning sessions. These data were to provide baseline measures against which change in the period between the Stage 1 and Stage 2 evaluation could be measured. This aspect of the evaluation focused on quality of relationships between Safe Families agencies, reporting of CSA and operational data relating to outcome indicators. The analysis of this data found that at the Wilcannia Safe Families office, staff shared a view that responding to CSA required a partnership approach and that the quality of agency relationships was consistently reported as positive. In Lightning Ridge the importance of partnerships was a shared ideology but the quality of relationships was reported as variable or negative.

The data also showed that there was little evidence of resources being provided to the community about CSA and that activities to achieve this were in their infancy. Consequently,

the evaluators found low levels of community knowledge regarding identification and responding to CSA, and where to go to for assistance. The Safe Families services were not nominated by community members as a place to go for help.

Finally, operational data from Safe Families teams such as numbers of referrals, participation of men in Safe Families initiatives, and overall involvement in community activities was not gathered by the Safe Families teams. In addition, data regarding reporting of CSA in Aboriginal communities was also not provided to the evaluators prior to the completion of the Stage 1 evaluation, though its provision was anticipated. As a result, these baseline measures were not developed. The evaluators urged that a culture of, and methods for, operational data collection be prioritised to support improved understanding of the program and its progress over time.

4.8.5 How sound are the findings?

There is little reason to doubt that the implementation of Safe Families has not been as successful as was originally hoped, but from a strictly methodological point of view it must be recognised that the Stage 1 evaluation has not provided irrefutable evidence. Nor was it able to achieve one of its main practical aims, namely to develop and record baseline measures to inform the planned outcomes evaluation.

The evaluators conducted a thorough evaluation, but were constrained by factors largely beyond their control. Based on advice from communities about what they considered would need to change in the short term for the initiative become effective, the evaluators planned to collect data on three indices: relationships between key agencies, reporting of child sexual assault, and measures relating to Safe Families operations/outputs. Contrary to their expectations, the evaluators were not provided with quantitative data relating to reported episodes of child sexual assault in the Safe Families sites, or with operational/output data such as referral rates, number of safety plans completed, attendance of men at educational events. This meant that the evaluators were unable to develop and record comprehensive baseline measures.

The evaluators were able to compensate for this to a degree. They gathered quantitative data on the quality of relationships between partner agencies, albeit from a relatively small sample group. Limited qualitative data was collected about community perceptions of child sexual assault. The data from the yarning sessions was of limited value because of the small sample size and the under-representation of men in the sessions. Some of the findings relating to the yarning sessions also need to be interpreted with caution because of potentially leading questions. Community members were provided with a scenario and asked to consider whether it depicted anything of concern, and participants were advised as part of the consent process and in the introduction to the yarning session that child sexual assault was the topic to be discussed.

Unfortunately, the evaluators did not seek to identify the success, or otherwise, of the “action learning model” which underpinned the approach to engaging with the limited evidence base upon which the program was founded. As a result the evaluation did not assess the ways in which the Safe Families and/or the Focus Communities initiatives were evolving as lessons were learned and areas for improvement were identified.

Finally, subsequent changes to the implementation model for Safe Families may limit the applicability of the baseline data and conclusions from this evaluation to the program in its

current form. This evaluation informed the development of a new model by the partner agencies;⁵³ another review will be undertaken in the 2012-13 financial year to gauge the overall effectiveness of the revised model, once it has been implemented in the communities for a period of 12 months.

4.8.6 Conclusions

Overall, the evaluation found that although there was variation of degree between locations, Safe Families had by the time of the evaluation achieved few of its intended goals of establishing strong community relationships and building community efficacy in relation to sexual assault of Aboriginal children. The program model and local structures and lack of organisational leadership were significant contributing factors to this situation. In addition, lack of foundational project building blocks – specifically the LARG and local Safe Families Plans – had meant that despite positive activity being done in the Wilcannia community it had not translated into achieving important project objectives.

The data to build baseline indicators regarding child sexual assault reporting was not available to the evaluators and therefore this aspect of the evaluation was not completed.

The misfit between goals and available project timeline, and the impending cessation of the project funding was seen as an issue of critical concern to the evaluators.

⁵³Aboriginal Affairs NSW, NSW Police Force, Department of Family and Community Services and NSW Ministry of Health (2011). *Safe Families. Response to the Evaluation of Safe Families by the Partner Agencies* ["Safe Families Response"]. Unpublished document held on file by Aboriginal Affairs NSW, Sydney.

5. Appendix 2 – Status of KTS Evaluations

Table 5 – Status of KTS Evaluations

<i>Expected</i>	<i>Commenced/Ongoing</i>	<i>Completed prior to KTS</i>	<i>Included in Synthesis</i>
Aboriginal Foster Carers Aboriginal Student Liaison Officers Accreditation for Voluntary OOHC carers Growing Partnerships (AbSec and ACWA) Home School Liaison Officers Interagency Conference JIRT referral unit Joint Domestic Violence Training Kaleidoscope Manager casework qualifications (research and training access) NGO Capacity Building OOHC Education coordinators Regional Governance Reparative parenting Services for Children of Parents with Mental Illness Toomelah/Boggabilla Project Working with men	Sustaining NSW Families Children under 10s Sexualised Behaviours Drug & Alcohol Unit Restoration and Short Term court orders Whole Family Teams OOHC Health Assessments Family Preservation Services Getting On Track In Time Aboriginal IFBS SAY Night Patrols Protecting Aboriginal Children Together	Brighter Futures Hey Dad! Indigenous Dads, Uncles & Pops Bail Assistance Line Mothering at a Distance New Street ⁵⁴	Family Case Management Family Referral Services Child Wellbeing Units Alternative Dispute Resolution Family Group Conferencing Screening and Response Priority Tools Safe Families and Focus Communities

⁵⁴Process reviews were completed for these programs prior to KTS, and outcomes evaluations are ongoing. They have been placed in this table according whether these outcomes evaluations are likely to be concluded before the conclusion of the KTS Strategic Impact and Outcomes evaluation in June 2014.

6. Appendix 3 - References

- Aboriginal Affairs NSW, NSW Police Force, Department of Family and Community Services and NSW Ministry of Health (2011). *Safe Families. Response to the Evaluation of Safe Families by the Partner Agencies*. Unpublished document held on file by Aboriginal Affairs NSW, Sydney.
- Aboriginal Child Sexual Assault Taskforce (2006). *Breaking the Silence: Creating the Future*. NSW Attorney General's Department, Sydney. Retrieved on 1 December 2012 from [http://www.lawlink.nsw.gov.au/lawlink/acsat/acsat.nsf/vwFiles/80001%20CP%20Rep-all_sml.pdf/\\$file/80001%20CP%20Rep-all_sml.pdf](http://www.lawlink.nsw.gov.au/lawlink/acsat/acsat.nsf/vwFiles/80001%20CP%20Rep-all_sml.pdf/$file/80001%20CP%20Rep-all_sml.pdf).
- AIATSIS (2011). *Guidelines for Ethical Research in Australian Indigenous Studies*. Australian Institute of Aboriginal and Torres Strait Islander Studies, Canberra.
- ARTD Consultants (2012). *Family Case Management. Final Report*. Unpublished document held on file by Department of Premier and Cabinet, Sydney.
- ARTD Consultants (2012). *Stage 1 Family Case Management. Interim Evaluation Report*. Unpublished document held on file by Department of Premier and Cabinet, Sydney.
- Department of Human Services - Community Services (2010). *Monitoring and Review of the Implementation of the Structured Decision Making (SDM®) Screening and Response Priority Tools at the Child Protection Helpline*. Unpublished document held on file by Department of Human Services, Sydney.
- Department of Premier and Cabinet (2010). *Family Case Management Evaluation Design and Execution. Request for Proposal*. Unpublished document held on file by Department of Premier and Cabinet, Sydney.
- Department of Premier and Cabinet (2010). *Keep Them Safe: A shared approach to child wellbeing*. Department of Premier and Cabinet, Sydney. Retrieved on 1 May 2012 from http://www.dpc.nsw.gov.au/data/assets/pdf_file/0004/57145/Keep_Them_Safe.pdf.
- Ernst & Young (2011). *FRS Evaluation. Final Report*. Unpublished document held on file by NSW Health, Sydney.
- KPMG (2012). *Keep Them Safe Workforce Survey. Final Report*. Department of Premier and Cabinet, Sydney. Retrieved on 4 December 2012 from http://www.dpc.nsw.gov.au/data/assets/pdf_file/0006/146670/KTS_WS_-_Final_Report_no_appendices.pdf.
- KPMG (2013). *Keep Them Safe Location Based Evaluation. Report*. Department of Premier and Cabinet, Sydney. Retrieved on 15 May 2013 from http://www.dpc.nsw.gov.au/prem/documents/keep_them_safe/publications_non-automated/Location_Based_Evaluation.pdf.
- Kristine Battye Consulting (2011). *Evaluation of Program Initiatives Tackling Child Sexual Assault in Aboriginal Communities. Stage 1 Report*. Unpublished document held on file by Aboriginal Affairs NSW, Sydney.
- KTS Evaluation Steering Committee (2011). *Keep Them Safe: A shared approach to child wellbeing. Interim Review Plan*. Department of Premier and Cabinet, Sydney. Retrieved on 1 May 2012 from http://www.dpc.nsw.gov.au/data/assets/pdf_file/0003/137406/12-04-12_-_Interim_Review_Plan.pdf.

- KTS Evaluation Steering Committee (2013). *Interim Review. Report*. Department of Premier and Cabinet, Sydney. Retrieved on 15 May 2013 from http://www.dpc.nsw.gov.au/prem/documents/keep_them_safe/publications_non-automated/Report_of_the_Interim_Review.pdf.
- Mikhailovich, Katja, Paul Morrison and Kerry Arabena (2007). "Evaluating Australian Indigenous community health promotion initiatives: a selective review". *Rural and Remote Health* 7:746.
- Nous Group (2011). *Keep Them Safe: Review of NSW Child Wellbeing Units. Final Report*. Unpublished document held on file by Department of Premier and Cabinet, Sydney.
- NSW Government (2003). *Two Ways Together*. New South Wales Aboriginal Affairs Plan (2003-2012). Department of Aboriginal Affairs, Sydney. Retrieved on 1 December 2012 from http://www.daa.nsw.gov.au/publications/TWT%20CopBroch_LR_1.pdf.
- NSW Government (2007). *New South Wales Interagency Plan To Tackle Child Sexual Assault in Aboriginal Communities*. Department of Aboriginal Affairs, Sydney. Retrieved on 1 December 2012 from <http://www.daa.nsw.gov.au/data/files//NSW%20Govt%20Plan%20to%20Tackle%20Aboriginal%20Child%20Sexual%20Assault.pdf>.
- Urbis (2011). *Implementation Plan for Evaluation of Keep Them Safe*. Department of Premier and Cabinet, Sydney. Retrieved on 1 May 2012 from http://www.dpc.nsw.gov.au/data/assets/pdf_file/0019/125146/Urbis_Final_KTS_Implementation_Plan_-_publication_version.pdf.
- Urbis (2012). *Keep Them Safe – Whole Family Teams: Year 2 Evaluation Final Report*. Unpublished document held on file by NSW Health, Sydney.
- Wood, James (2008). *Report*, Vol. 1. State of NSW through the Special Commission of Inquiry into Child Protection Services in NSW, Sydney. Retrieved on 1 May 2012 from http://www.dpc.nsw.gov.au/data/assets/pdf_file/0010/33796/Volume_1_-_Special_Commission_of_Inquiry_into_Child_Protection_Services_in_New_South_Wales.pdf.
- Wood, James (2008). *Report*, Vol. 2. State of NSW through the Special Commission of Inquiry into Child Protection Services in NSW, Sydney. Retrieved on 1 May 2012 from http://www.dpc.nsw.gov.au/data/assets/pdf_file/0011/33797/Volume_2_-_Special_Commission_of_Inquiry_into_Child_Protection_Services_in_New_South_Wales.pdf.