



What Works *for*
**Children's
Social Care**

CARDIFF
UNIVERSITY
PRIFYSGOL
CAERDYDD

IMPACT OF SHARED DECISION-MAKING FAMILY MEETINGS ON CHILDREN'S OUT-OF-HOME CARE, FAMILY EMPOWERMENT AND SATISFACTION

A systematic review





What Works for Children's Social Care

Acknowledgements

We would like to thank Lisa Merkel-Holguin, Victoria Weisz, Hanneke Creemers, and Simone Onrust for their help in locating literature and sharing unpublished work. We would also like to thank Yvonne Taylor, Laura Cook, Shaheda Khatun, Daniel Latham and Wanda O'Connor for administrative support in conducting and finalising the review.

Conflict of interest

The authors do not have any conflicts of interest to declare.

Funding

Department for Education, England.

Authors

Nurmatov, UB., Division of Population Medicine, School of Medicine, Cardiff University; **Foster, C.**, CASCADE, School of Social Sciences, Cardiff University; **Bezeczyk, Z.**, CASCADE, School of Social Sciences, Cardiff University; **Owen, J.**, School of Geography and Planning, Cardiff University; **El-Banna, A.**, Warwick Clinical Trials Unit, University of Warwick; **Mann, M.**, Specialist Unit for Review Evidence, Cardiff University; **Petrou, S.**, Nuffield Department of Primary Care Health Sciences, Medical Sciences Division, University of Oxford; and, Warwick Clinical Trials Unit, University of Warwick; **Kemp, A.**, Division of Population Medicine, School of Medicine, Cardiff University; **Scourfield, J.**, CASCADE, School of Social Sciences, Cardiff University; **Forrester, D.**, CASCADE, School of Social Sciences, Cardiff University; **Turley, R.**, DECIPHer, School of Social Sciences, Cardiff University.

About What Works for Children's Social Care

The What Works Centre for Children's Social Care seeks better outcomes for children, young people and families by bringing the best available evidence to practitioners and other decision makers across the children's social

care sector. We generate, collate and make accessible the best evidence for practitioners, policy makers and practice leaders to improve children's social care and the outcomes it generates for children and families.

About CASCADE

The Children's Social Care Research and Development Centre (CASCADE) at Cardiff University is concerned with all aspects community responses to social need in children and families, including family support services, children in

need services, child protection, looked after children and adoption. It is the only centre of its kind in Wales and has strong links with policy and practice.

To find out more visit the Centre at: whatworks-csc.org.uk, or CASCADE at: sites.cardiff.ac.uk/cascade

If you'd like this publication in an alternative format such as Braille, large print or audio, please contact us at: wwccsc@nesta.org.uk



CONTENTS

ABBREVIATIONS AND DEFINITIONS	6	3. RESULTS	21
EXECUTIVE SUMMARY	7	3.1 Description of studies	21
1. INTRODUCTION	9	3.2 Risk of bias assessment	24
1.1 Description of the problem	9	3.2 Certainty of evidence assessment	24
1.2 Description of the service	10	3.4 Effectiveness of shared decision-making family meetings	24
1.3 How the service may work	11	3.5 Cost-effectiveness of shared decision-making family meetings	34
1.4 Why this review is needed	12	4. DISCUSSION	40
1.5 Objectives of this systematic review	13	4.1 Summary of main findings	40
2. METHODS	15	4.2 Discussion of findings	40
2.1 Protocol registration	15	4.3 Strengths and limitations	42
2.2 Population	15	4.4 Conclusions	44
2.3 Service of interest	15	5. AUTHOR CONTRIBUTION	45
2.4 Study eligibility criteria	15	6. REFERENCES	46
2.5 Outcome measures	16	7. APPENDICES	50
2.6 Search strategy	17	7.1 Appendix 1. List of databases and agencies searched	50
2.7 Study selection	17	7.2 Appendix 2. Scopus search strategy	51
2.8 Data extraction	17	7.3 Appendix 3. List of experts contacted	53
2.9 Study design categorisation	18	7.4 Appendix 4a. Description of excluded papers	54
2.10 Risk of bias assessment	18	7.5 Appendix 5. Included papers	59
2.11 Assessing the certainty of evidence using GRADE	19	7.6 Appendix 6. Study categorisations	61
2.12 Data analysis	19		



CONTENTS

7.7	Appendix 7. Descriptive characteristics of included studies	64
7.8	Appendix 8. Summary of shared decision-making family meetings	78
7.9	Appendix 9: Glossary table of economic terms	99
7.10	Appendix 10a. Critical appraisal of included RCTs assessed by the Cochrane Risk of Bias tool	100
7.10	Appendix 10b. Risk of bias graph: review authors' judgments about each risk of bias item presented a percentage across all included RCTs	101
7.10	Appendix 10c. Risk of bias summary: review authors judgments about each risk of bias item for each RCT	102
7.11	Appendix 11. Critical appraisal of included quasi-experimental studies assessed by the ROBINS-I tool	103
7.12	Appendix 12. GRADE summary tables	106
7.13	Appendix 13. Harvest plots for non-FGC meeting types	113
7.13	Appendix 13a. Care Entry	113
7.13	Appendix 13b. Reunification	115
7.13	Appendix 13c. Referrals for maltreatment	117



ABBREVIATIONS AND DEFINITIONS

CS1

Comparative study with efforts to make groups comparable

CS2

Comparative Study without efforts to make groups comparable

EMBASE

Biomedical and pharmacological information database

FGC

Family Group Conference

FGDM

Family Group Decision Making

FTC

Family Team Conference

FTM

Family Team Meeting

FUM

Family Unity Meeting

GRADE

Grading of Recommendations, Assessment, Development and Evaluations

PRISMA

Preferred Reporting Items for Systematic Reviews and Meta-Analyses

PROSPERO

International Prospective Register of Systematic Reviews

RCT

Randomised Controlled Trial

ROBINS-I

Risk of Bias in Non-Randomised Studies – of Services

TDM

Team Decision Making



EXECUTIVE SUMMARY

Background

Shared decision-making family meetings are a key method for facilitating the participation of family members in the safeguarding of children in need. Traditional child protection case conferencing is professionally driven, with social workers leading the assessment of families' problems and development of a service plan for families to comply with. Shared decision-making family meetings have arisen as a more participative alternative to this traditional model. A variety of different names are used internationally. The most familiar approach in the UK is family group conferences.

Objectives

This review comprehensively identified and assessed the evidence of the effect of shared decision-making family meetings in reducing the need for placing children in out-of-home care, and increasing family reunification, family empowerment and satisfaction, as well as reviewing the published literature on the cost-effectiveness of shared decision-making family meetings. Since an important rationale for shared decision-making meetings is upholding the participation rights of families, the review has considered the outcomes of family empowerment and satisfaction as proxies for this.

Methods

The systematic review protocol was prepared using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocol (PRISMA-P) guidelines and was registered on the International Prospective Register of Systematic Reviews (PROSPERO) (CRD CRD42019138011) prior to

the commencement of literature searches and analysis.

Empirical evaluation studies were included if they quantitatively compared a primary outcome of interest in families who took part in shared decision-making meetings with outcomes in a comparative group. Outcomes could be either for children and young people (0-18 years of age) at risk of entering or already in out-of-home care, or for their parents/carers/guardians, or for both. Data extraction, informed by the TiDIER template, was carried out to summarise the service used, study methods and results.

The primary outcomes considered were: number of children and young people entering and re-entering out-of-home care; number of children and young people reunified with their family following a period in care; quantitative measures of families' perception of empowerment in parenting situations; and quantitative measures of client satisfaction with the service. Secondary outcomes considered were adverse effects (substantiated child protection referrals or re-referrals) and cost-effectiveness.

The quality of included studies was assessed using the Cochrane eight domain-based evaluation for randomised controlled trials (RCTs) and quasi-RCTs. For other non-randomised comparative studies, the ROBINS-I tool was used. All publications that comprised a full economic evaluation underwent a further round of quality assessment against the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) guidance. The Grading of Recommendations, Assessment, Development and Evaluations (GRADE) was employed to judge the confidence in evidence and certainty of evidence in the primary outcomes. As meta-analysis could not be conducted, results for each outcome were summarised visually using Harvest Plots.



Results

This systematic review included 33 studies focusing on quantitative evaluation of shared decision-making meetings compared to control group services, which were usually termed "care as usual". For the seven outcome measures of interest, the evidence was inconclusive as to whether family group meetings are more effective than control services. The majority of studies (24 out of 33) were found to have a high risk of bias and the GRADE assessments found low or very low certainty of evidence for each primary outcome.

Seventeen of the 33 studies found a favourable result for the shared decision-making meetings compared to control group services for at least one outcome. However, all these studies had a high risk of bias except one whose risk of bias was moderate. When all the out-of-home care outcomes are pooled for all meeting types, the results could be interpreted as leaning in a positive direction for the effectiveness of shared decision-making meetings in preventing out-of-home care. However, none of the randomised-controlled trials identified a reduction of entry or re-entry to care, referrals or re-referrals for maltreatment, or increased satisfaction, parental empowerment or reunification with family when compared to control services. Satisfaction and parental empowerment were measured in only very few studies and the results were not conclusive. When different types of meetings were separated out – for example the results were considered just for family group conferences – this did not affect the overall view for all meeting types, namely that the evidence is inconclusive.

Seven studies included an economic analysis as part of their evaluation of shared decision-making meetings. Even though there was no strong evidence on the cost-effectiveness of shared decision-making meetings, there are encouraging indications that shared decision-making meetings could be cost saving. They are a relatively cheap service, and even a small impact on the use of public care would justify them as an investment on purely financial grounds.

Conclusion

The review does not provide conclusive evidence as to how shared decision-making meetings affect care entry, re-entry, family reunification, family empowerment or satisfaction, compared with usual services. The lack of strong evidence is at least partially attributed to a lack of high-quality quantitative and comparative evaluation studies and considerable variation in published results. Future research needs to measure the crucial outcomes of satisfaction and empowerment, with robust comparative research designs.

Family participation in decisions about children should be upheld as a fundamental principle within the child welfare system, but the results of this review suggest that more work is needed to improve the quality consistency of the services that are designed to achieve this. It may be that these meetings are not being run consistently well or they are not often enough part of a wider cultural change in children's services towards prioritising family participation.



1 INTRODUCTION

1.1 Description of the problem

Involving parents, children and other family members in ensuring child welfare is a key principle of the Children Act 1989 and is supported by Article 18 of the United Nations Convention on the Rights of the Child (UNCRC) (U.N. General Assembly, 1989). When a child is at risk of harm, it is argued that the immediate and wider family have a right to be involved in the key decisions about the child's upbringing, even when the current care of the child is regarded by many as seriously inadequate (Connolly and Masson 2014).

This systematic review assessed the research evidence around shared decision-making meetings aimed at increasing family participation in the child protection process. In the United Kingdom these are typically called Family Group Conferences (FGCs), though meetings with similar aims and structures are referred to by a variety of different names in the literature. In the US for example, FGCs are one model within the Family Group Decision Making (FGDM) approach, which also includes Family Unity Meetings and combinations of these two models (Center for the Study of Social Policy, 2002). This review considered all meetings aimed at improving shared decision-making with families. We outline more about the nature of such meetings below.

A main rationale for changing the way families are involved in meetings is the considerable evidence mainstream practice could do much more to involve families in decision-making. This is most obvious in relation to child protection case conferences. Case conferences are a key method of safeguarding children in need. These multi-agency meetings seek to establish a child's safety, promote their wellbeing, and identify whether they are at continuing risk of significant harm (Muench et al. 2017). Whilst parents attend

case conferences in the UK, traditional case conferencing is professionally driven, with social workers leading the assessment of families' problems and development of a service plan for families to comply with. Conferences may occur at different times within the life of a case, e.g. initial investigation when a child is at risk of entering care; or when a child is already in care and family reunification is being considered. Where conferences fail to find adequate solutions or families do not comply with resulting action plans, a child may enter or remain in care. There are, however, legitimate concerns about the ability of traditional case conferences to provide safe, respectful and proportionate services. Children and parents' negative experiences of traditional conferencing are well documented (Corby et al. 1996; Muench et al. 2017). In particular, the adversarial and even traumatising format is thought to hinder effective relationship building and engagement between families and their social workers (Darlington et al. 2012). A key driver for involving families more meaningfully in meetings is therefore the desire to address the fact that too often we are not doing well enough at present and a position that families' right to participation is not being upheld in mainstream practice.

A second driver is the belief that in doing so we are more likely to empower families to create their own solutions. A particular focus in this respect has been the potential that family involvement might have to reduce the need for children to be in out-of-home care. The rationale is that involving the wider family network will harness resources and ideas that might allow children to either remain with their parents or be cared for in the wider family. At a time when we are seeing large increases in the number of children in care (Thomas 2018) this is obviously an attractive possibility. It is by no means the only reason for sharing decision



making, as the involvement of families in decisions should be regarded as an important end in itself, however, reducing the removal of children from their families has consistently been an important aim when shared decision-making meetings have been introduced internationally.

This review focused on the evidence in relation to these two areas. First, do meetings such as Family Group Conferences and Family Group Decision-Making meetings involve families more effectively in decision-making in their lives? For instance, do parents report feeling more empowered or even just greater satisfaction with the service if there is such a meeting? Second, do the meetings help keep children out of state care and reduce the rate of referrals for maltreatment to children's services?

1.2 Description of the service

Alternative family-centred models have proliferated in response to the misgivings surrounding traditional conferencing. These include well-known models such as Family Group Conferencing (FGC), originating in 1980s New Zealand, and its US variant Family Group Decision Making (FGDM) which includes Family Unity Meetings. There are also a variety of similar approaches such as Team Decision Making, Family Involvement Meetings, Family Group Meetings, Family Team Meetings, Family Welfare Conferencing and Family Team Conferencing (Skaale Havenen and Christiansen 2014; Stabler et al. 2019). While there may be some differences in the design and implementation of the different models, as well as the aims, all meeting types included in this review adhered to the same principles of an organised planned meeting convening both social work practitioners and family members, with efforts to ensure the decision-making process was family led (see section 2.1 for further details). Therefore, we did not restrict the review to studies using the most well-known FGC model. We acknowledge the potential for heterogeneity between meeting types, particularly the degree to which meetings are family-led. However, as many decision-making meeting types are being used, it is important to examine outcomes in all meetings which seek to meaningfully involve the family in decision-making.

What Works for Children's Social Care have published two related evidence reviews to examine these services and will use the umbrella term of "shared decision-making meetings". This systematic review analysed the impact that these approaches have on whether children enter or remain in out-of-home care and whether families feel more satisfied or empowered, using a quantitative approach. The companion rapid realist review presents a detailed theory about how the services work and what has to happen for shared decision-making family meetings to improve the likelihood that shared decision-making meetings are effective (Stabler et al. 2019).

In shared decision-making meetings families are encouraged to actively participate in devising comprehensive plans for their child or children. Four stages are usually incorporated: referral (agreement that a conference is required and a coordinator, often impartial, is appointed); preparation (coordinator identifies, invites and discusses reasons for the meeting with the family network); the meeting (family group convenes with the coordinator and welfare practitioners, is provided with information to enable them to develop a plan for the child, and the plan is approved by the professionals providing it protects the child's needs); and follow-up (reviewing how the plan is working and monitoring adherence) (Barnsdale and Walker 2007). It should be mentioned that there is likely to be little planning where meetings are convened on an emergency basis, such as in Team Decision Making and Family Team Meetings and this may affect outcomes.

This process is generally underpinned by the following core values (Barnsdale and Walker 2007; Skaale Havenen and Christiansen 2014):

- Emphasis on the problem rather than the person – facilitated decision-making which focuses on healing and preventing harm rather than attributing blame.
- Extended family approach – those who care about the child involved and have useful information to share, beyond the domestic family unit, are invited to participate.
- Taking or sharing responsibility – family are given more freedom to interact and more



responsibility to help find good solutions to ensure child safety and wellbeing.

- Family empowerment and private time – family prepare an action plan, usually with the opportunity to do so in the absence of child protection workers.
- Cultural understanding – drawing on the cultural patterns and resources of the family and community involved.
- Community partnerships – aims to build new partnerships to secure community-based resources that can help the family in identifying home-grown solutions.

Differences in the make-up of models tend to concern whether or not the family can veto invitations to wider members, there is provision for private family time during the meeting, or age restrictions are applied for children's participation. Coordinator roles may vary in terms of their independence from child protection services, the case itself, and if responsibilities for meeting preparation and chairing are held by one individual or conducted separately by different coordinators (Skaale Havenen and Christiansen 2014). Meetings may differ according to whether referral to the family meeting is voluntary or not (Stabler et al. 2019). Some meeting types such as Team Decision Making happen when there is imminent risk of a child entering care, or when emergency placement occurs. Although families are invited and encouraged to attend these meetings, there is little planning time and the meeting will go ahead with or without the presence of key family members. This is in contrast to the FGC model which involves greater planning time and there is not always an immediate need for the child to be taken into care, or they may already be in care. As a result, we must be careful in interpreting outcomes such as care entry for these two meeting types when the risk of care entry is likely to be higher where emergency meetings have been convened.

1.3 How the service may work

The service is thought to work by encouraging partnership working and activating family resources to strengthen the safety net for children at risk of harm (Connolly 2006). Engaging the

family network in decision-making may make them more likely to engage with social workers in a meaningful way (Muench et al. 2017), and motivate parents to work alongside professionals to make agreed plans work (Faller 1981; Featherstone et al. 2018). Moreover, engaging with the wider family network may harness more resources for the family to ensure that the child can remain safely in the home (Appleton 2014). This could be due to making more people aware of the difficulties that the family is facing and therefore allowing them the opportunity to offer support (Morris 2007).

The programme theory developed from our companion realist review identified that across the three core stages of the participative meetings – pre-meeting preparation, meeting process and effective follow-up – there were three higher-level mechanisms that made shared decision-making meetings likely to be effective (Stabler et al. 2019):

- Enabling collaboration and engagement: Essentially, this mechanism is concerned with what social workers and other professionals do to enable true collaboration with families in a meeting, and how this creates family and wider network engagement in the meeting process.
- Building trust and reducing shame: Building trust between social workers and families can be an important mechanism for parents and the wider family to feel able to participate in a meeting in a way that is open, and solution focused. Related to this, where families may feel shame around their involvement with children's services, and issues that they have faced, proactively working with a social worker to manage this may help to build a more knowledgeable support network around the family and child.
- Enabling participation in decision making: One of the main outcomes from shared decision-making family meetings is to enable families to be involved in making important decisions about the care and safety of the child. This mechanism is enabled through the other two key mechanisms and is a pathway itself.

A conceptual model of the programme theory is presented in Figure 1.



1.4 Why this review is needed

To complement What Works for Children's Social Care realist review of service mechanisms and planned experimental work, there is a need for an up-to-date comprehensive systematic review of the effectiveness of shared decision-making meetings.

Existing literature reviews identify that a diverse evidence base is available (Barnsdale and Walker 2007; Skaale Havenen and Christiansen 2014; Dijkstra et al. 2016b). There is important evidence from qualitative and single group quantitative studies in the UK and elsewhere over the last few decades which discusses the complexities of this field of practice and reaches a positive conclusion about the potential of shared decision-making family meetings (Morris and Tunnard 1996; Crow and Marsh 1998; Lupton and Nixon 1999; Pennell and Burford 2000; Holland and O'Neill 2006).

The strengths of these studies are acknowledged. Qualitative research methods are better suited than quantitative for some purposes – for example, describing lived experience. Single-group quantitative studies can be important for identifying promising approaches. However, the focus in this systematic review is on comparative evaluation designs with quantification of outcomes. The reason for this is that in order to properly assess whether an innovative service or approach is more effective than conventional services (in this case, usually child protection case conferences) you need a comparison group and standardised measurement of one or more outcome. The decision to limit the scope of this review to such studies is not based on any hierarchy of evidence but on identifying the most appropriate type of evidence for a given research question – horses for courses, as Petticrew and Roberts (2003) put it.

The recent mapping exercise by the What Works Centre, for any activity to safely reduce out-of-home care (Brand et al. 2018), and the realist review (Stabler et al. 2019) identified that there are additional relatively newly published comparative studies in peer-reviewed journals available for examination (Hollinshead et al. 2017; Lambert et al. 2017; Dijkstra et al. 2018a) and several grey

literature reports which have not been included in previous reviews (Partnership for Strong Families 2012; YMCA Families United 2014; Beehler 2016), including two UK studies (Mason et al. 2017; Munro et al. 2017) which were the only UK studies meeting our inclusion criteria.

Extant literature reviews either partially examine our outcomes of interest, or do not use methods that combine study findings in a meaningful way. For example, one meta-analysis that employed a moderator analysis identified a small positive, but statistically insignificant, effect of family group conferencing on reducing out-of-home placement (Dijkstra et al. 2016b). However, as this study focused solely on out-of-home placement it is unclear whether the meetings improved other important outcomes such as family reunification, satisfaction or empowerment compared to usual care in the studies reviewed by Dijkstra et al. (2016b).

Two grey literature evidence reviews examined a wider range of outcomes, but neither are systematic syntheses or examined studies for risk of bias or quality of evidence (Barnsdale and Walker 2007; Skaale Havenen and Christiansen 2014). The first of these evidence reviews, completed for the Scottish Executive in 2007, surmised that family group decision-making may have a beneficial effect for children and families, but noted that findings were inconsistent between studies and there was a general lack of robust research designs (Barnsdale and Walker 2007). The second review in 2014, also noted the heterogeneity of study findings and research designs (Skaale Havenen and Christiansen 2014). These authors concluded that family group conferences may increase placement with relatives (as an alternative to public foster home) and the likelihood of family reunification but did not find sufficient research for whether family group conferences improve relationships between family and welfare services, prevent maltreatment or prevent entry into out-of-home care (although this may not always be the main aim of meetings, as discussed later). These authors also concluded that family group conferencing facilitates access to services beyond those offered by child welfare services in the short term, but not long term.



A systematic review not only enables more exhaustive identification of up-to-date relevant studies, but also examines study findings in light of their methodological conduct and risk of bias. This is important because the type of research design is not the only indicator of our confidence in a study's findings. For example, well conducted quasi-experimental studies may provide more reliable evidence than poorly conducted randomised controlled trials (RCTs). Methods such as Harvest Plots offer a sophisticated alternative for visually synthesising findings across diverse studies (Ogilvie et al. 2008), which together with GRADE (Grading of Recommendations, Assessment, Development and Evaluations), a transparent framework for developing and presenting summaries of evidence, help consider the overall strength and direction of the evidence base supporting each outcome. Finally, given the well-documented negative experiences of traditional case conferencing (Corby et al. 1996; Muench et al. 2017). It is important that an updated systematic review of shared decision-making meetings not only considers the impact on out-of-home care reduction, but also whether participative meetings lead to improved family satisfaction and empowerment, as proxies for promoting family members' rights to participation. It may be that even if outcomes for children do not improve as a result of a shared decision-making meeting, this approach could nonetheless be highly valued by family members because their opinions are taken seriously, which could still recommend their use.

This systematic review offers a robust, broader and up-to-date synthesis of the literature to maximise the value of the literature available through assessment of bias and certainty of evidence using standardised tools.

1.5 Objectives of this systematic review

1.5.1 Primary research questions

- Are shared decision-making meetings effective at reducing out-of-home placements and increasing reunification in families of children 0-18 years of age?

- Are shared decision-making meetings effective at improving family empowerment and satisfaction with child welfare services?

1.5.2 Secondary research questions

- Do shared decision-making meetings result in the adverse outcome of increased referrals for child maltreatment?
- Are shared decision-making meetings a cost-effective service?

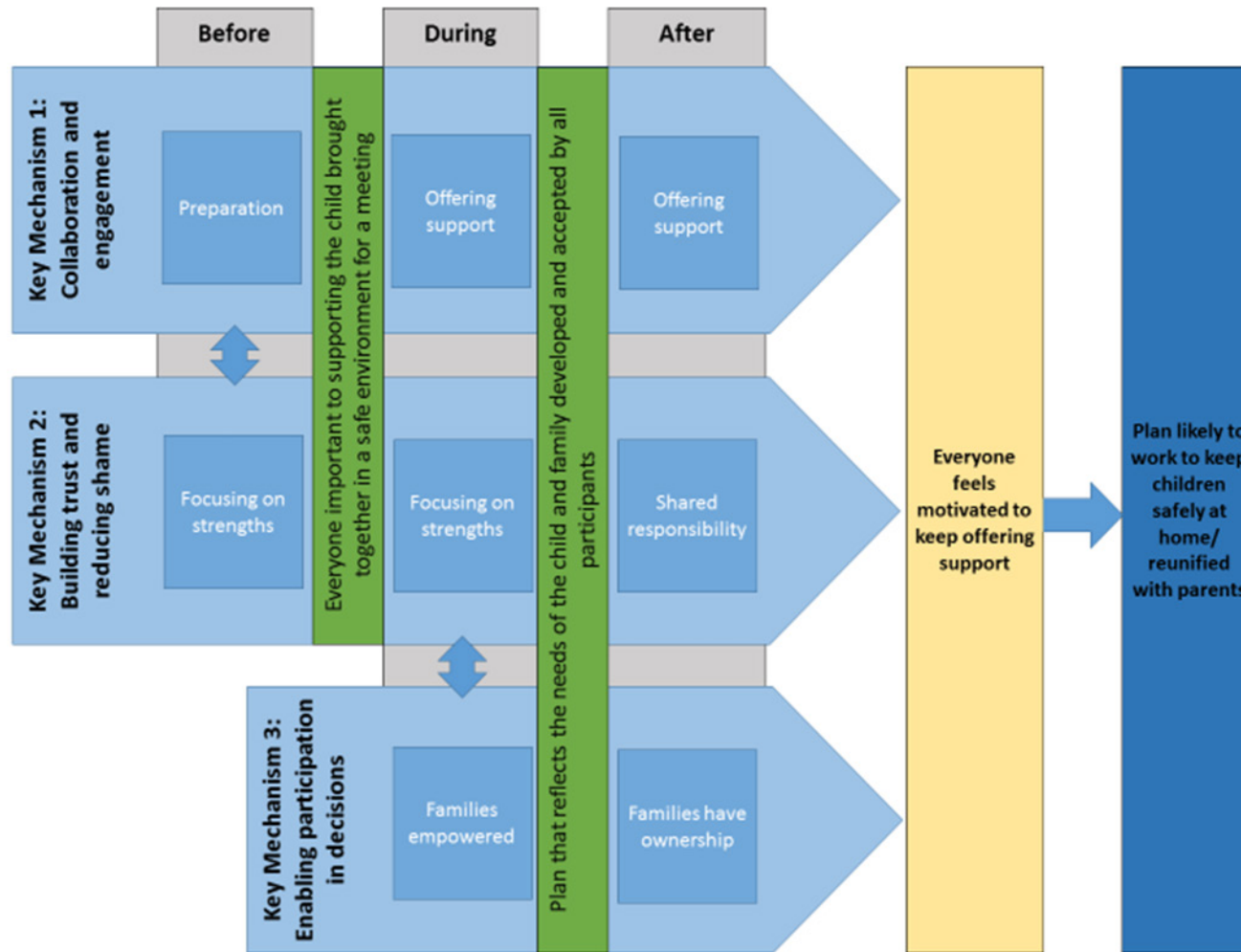


Figure 1. Overarching programme theory showing three key mechanisms of shared decision-making family meetings (Reproduced from Stabler et al. (2019)).



2 METHODS

2.1 Protocol registration

This systematic review protocol was prepared using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocol (PRISMA-P) guidelines and registered on International Prospective Register of Systematic Reviews (PROSPERO) (CRD CRD42019138011).

2.2 Population

Children and young people (0-18 years of age) who are either at risk of entering, or already in, out-of-home care and/or their parents/carers/guardians.

Children and young people may enter out-of-home care for a range of reasons including extreme risk of: abuse and neglect (e.g. sexual abuse, emotional abuse, physical abuse, supervisory neglect); where parents cannot provide good enough care for the children due to acute family problems (e.g. parental substance misuse); family in acute stress (e.g. financial crisis); child's disability; carer's illness or disability; socially unacceptable behaviour (pre-entry into juvenile court system).

Out-of-home care is defined as a child or young person being looked after by a local authority (or international equivalent), including those who are placed under a court order or a formal voluntary agreement with parents. Our definition of care does not extend to include care arrangements that are informal or those that do not specify continued statutory involvement (e.g. adoption).

2.3 Service of interest

Any service involving shared decision-making meetings between families and professionals. These services were defined as involving:

- an organised planned meeting convening both social work practitioners and family members

(and often extended family, friends and other professionals); that

- seeks to develop an action plan to maintain child safety, wellbeing and inform the next steps of social work involvement; whereby
- some efforts to ensure the decision-making process is collaborative and family-led – e.g. practitioners support families to develop family-led solutions to the situation of concern.

Both singular shared decision-making meetings and multi-component services that include a shared decision-making meeting element were eligible for inclusion. Studies were included if they compared meetings to either usual care or an alternative service.

2.4 Study eligibility criteria

Studies comparing a primary outcome of interest in the shared decision-making meeting service with outcomes in a comparative group were included. Both interventional and natural experiment studies were eligible for inclusion in the review. Interventional studies are those in which the circumstances of the service implementation are under the control of the researchers, e.g. RCTs. Natural experiments lack a consensus description (Craig et al. 2012; Leatherdale 2019), so for the purposes of this review we employed the broad definition applied by Medical Research Council guidance:

“By natural experiments, we mean events, interventions or policies which are not under the control of researchers, but which are amenable to research which uses the variation in exposure that they generate to analyse their impact. By natural experimental studies, we mean the methodological approaches to evaluating the impact on health or other outcomes of such events. The key features of these definitions are that (1) the intervention



is not undertaken for the purposes of research, and (2) the variation in exposure and outcomes is analysed using methods that attempt to make causal inferences" (Craig et al. 2012)

To meet our eligibility criteria, the studies had to include both a shared decision-making meeting service and control group. Within-group, cross-sectional designs were not included. This meant that RCTs and quasi-experimental designs, including natural experiments, were included so long as there were clearly defined shared decision-making meeting service groups and control groups. Quasi-experimental evaluation designs are also inconsistently defined in the literature (Craig et al. 2012). In this review, quasi-experimental describes evaluation designs in which participants receiving a service are compared to those who receive a different service or control but are not randomly assigned to groups. The quasi-experimental design may be used within interventional or natural experiment studies.

It was expected that a portion of the eligible studies would have conducted an economic evaluation alongside the main effectiveness study identified. In addition, searches were also carried out for studies that exclusively conducted economic evaluations of shared decision-making meetings using decision analytical modelling techniques. All types of partial and full economic evaluations of shared decision-making meetings were included in the review.

Studies with the following characteristics were not eligible for inclusion in the review and therefore were excluded: a) Populations focused on children and young people who are in need of out-of-home care, or have been in care, when ≥ 18 years old; b) Informal care arrangements that do not specify statutory involvement; e.g. homeless shelters that do not have statutory involvement or informal kinship care; c) uncontrolled pre-post evaluation; process evaluation or qualitative study (unless companion papers to an eligible study design); letters; commentary; expert opinion; case reports, literature reviews.

2.5 Outcome measures

Primary outcomes

- **Rates of out-of-home placement**
 - Number of children and young people entering out-of-home care.
 - Number of children and young people re-entering out-of-home care.
- **Reunification rates**
 - Number of children and young people reunified with their family following a period in care.
- **Family empowerment**
 - Quantitative measures of families' perception of empowerment in parenting situations, e.g. the Family Empowerment Scale (FES) (Dijkstra et al. 2018a).
- **Family satisfaction**
 - Quantitative measures of client satisfaction with the service, e.g. as collected by "Decision Process Ranking Scale" (Pennell and Burford 1995).

Secondary outcomes

- Adverse effects
- Substantiated referrals/re-referrals to a child protection authority.
- Economic data
 - Costs off-set due to shared decision-making meetings.
 - Cost difference between shared decision-making meetings and comparator.
 - Economic evaluations measuring benefit in monetary terms.
 - Economic evaluations incorporating incremental cost-effectiveness ratios (ICERs) that measure benefit in units specific to shared decision-making meetings e.g. number avoided care or use social care related quality of life as the outcome measure.



2.6 Search strategy

A comprehensive search for published and unpublished studies was conducted from June to July 2019. There were no language or geographical restrictions.

Ten databases were searched, covering a range of relevant disciplines: Child Development and Adolescent Studies, EconLit, EMBASE, MEDLINE, NHS Economic Evaluation Database (NHS EED), PsycINFO, Research papers in Economics (RePEc), Scopus, Social Policy and Practice and Sociological Abstracts. The search strategy was first developed in SCOPUS. The strategy was tested and refined using a subset of key papers already known to the team. The final SCOPUS strategy was then tailored to the remaining databases (see Appendix 2 for search history).

Supplementary forensic searches were also conducted to help identify further potential studies including grey literature and any ongoing studies. Approaches involved: forward and backward citation tracking of included studies; contacting a panel of international experts (see Appendix 3); and keyword searching websites of relevant agencies: Center for Family and Community Involvement; Child Welfare Information Gateway; Family Rights Group (UK); Open Grey; The Healthcare Management Information Consortium (HMIC), The Children's Social Care Innovation Programme; The American Humane Association (USA); The National Institute for Permanent Family Connectedness.

2.7 Study selection

Records resulting from the literature searches were exported into the reference manager software Endnote, and de-duplicated. Four authors (ZB, CF, MM and UN) screened titles and abstracts independently in duplicate for potentially eligible studies. Disagreement between researchers were resolved by consensus or arbitration involving a third author (AK or RT) where necessary. Full texts of potentially eligible studies were retrieved, and four authors (ZB, CF, MM and UN) evaluated whether these met inclusion or exclusion criteria. Again, disagreements were resolved by discussion among authors, with referral to a third author (AK

or RT) if necessary. A list of excluded papers along with reasons for their exclusion are provided in Appendix 4.

2.8 Data extraction

Each included study was data extracted independently in duplicate (ZB, CF, JO and AE), using an a priori designed data extraction form (CF and ZB). This contained two core components:

- i. Service description: this section extracted information about service activities and how they were implemented and was informed by the TiDIER checklist (Hoffman et al. 2014). Information was also gathered on whether the service included any of the mechanisms considered important for shared decision-making meetings by the WWC's companion realist review; referral, pre-planning of the meeting, the meeting itself, and details on implementation of the developed plan (Stabler et al., 2019). The data gathered in this component aided judgments about the comparability of services across the included studies.
- ii. Study data: Including the following: study aims, country of origin, study design, setting, sample size, service population characteristics, control characteristics, outcome measure used and analysis results.

If the study reported an economic evaluation, details were extracted using a third component, designed to extract methodological detail as well as the results of the economic evaluation. Where reported, costs and outcomes data were extracted in addition to the cost-effectiveness decision made by the study authors, including the thresholds used by decision makers to determine cost-effectiveness.

Any disagreements on eligibility for inclusion were settled by a third investigator. The key characteristics of the services and research designs, plus findings from each study are summarised and presented in descriptive summary tables (Appendices 7 and 8).



2.9 Study design categorisation

Limited reporting of research type and service implementation often precluded determining whether studies were interventional or natural experiments. Therefore experimental studies which were not RCTs were classified by two authors (ZB and CF) according to six categories

of their evaluation design, adapted from Leatherdale's (2019) schema summarised in the left panel of Table 1. Furthermore, it was recorded whether studies had used additional analytic techniques to reduce selection bias and improve the comparability between service and control groups (Craig et al. 2012; Craig et al. 2017), summarised in the right hand panel of Table 1.

Table 1: Classification of study evaluation designs and additional analytic techniques

Study evaluation designs	Additional analytic techniques for reducing selection bias
<p>RCT</p> <p>Quasi-experimental (QE)¹</p> <ul style="list-style-type: none"> ▪ Longitudinal pre-post ▪ Longitudinal Interrupted time series² ▪ Longitudinal time series post-test only³ ▪ Repeat cross-sectional pre-post ▪ Repeat cross sectional interrupted time series ▪ Cross-sectional post-test only <p><i>Exclude:</i> Uncontrolled interventional or natural experiment studies (e.g. pre-post study)</p>	<p>Selection on observables</p> <ul style="list-style-type: none"> ▪ Matching ▪ Propensity scores ▪ Regression adjustment <p>Selection on un-observables</p> <ul style="list-style-type: none"> ▪ Difference-in-differences ▪ Regression discontinuity ▪ Instrumental variables

1 Interrupted time series were defined as requiring at least two data points pre-intervention and post-intervention.

2 The design may or may not have an additional concurrent control group.

3 Requires at least two data points post intervention

2.10 Risk of bias assessment

The quality of included studies was assessed using the Cochrane eight domain-based evaluation for RCTs and quasi-RCTs (Cochrane Handbook, table 8.5.a (Higgins and Green 2011) which assess the study for risks of study bias due to; baseline confounding, participant selection, intervention group allocation, deviations from the study protocol, measurement of outcomes, missing data and selective reporting of results. Each domain was rated as low, unclear or high risk of bias. For

other non-randomised studies of shared decision-making meetings ROBINS-I tool was used (Sterne et al. 2016). Each parameter of trial quality was graded as low, moderate, serious or critical risk of bias. Studies are given an overall rating based on the highest score received on any domain. For example, if a study received a moderate risk of bias score for six domains, but high for two domains, the overall rating would be high. The breakdown of scores by domain is given in Appendix 10a-c to provide a more in-depth description of the risk of bias.



All publications that comprised a full economic evaluation underwent a further round of quality assessment against the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) guidance (Husereau et al. 2013).

One author (AE) assessed the quality of the economic evaluations. Four authors (UN, JO, ZB and CF) carried out all other critical appraisals of the included papers. Two authors completed each critical appraisal and any disagreement was resolved by consensus, or arbitration involving a third author (AK or RT).

2.11 Assessing the certainty of evidence using GRADE

The transparent international framework, the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) was employed to judge the confidence in evidence, the certainty of evidence in the importance of primary outcomes or values and preferences from included shared decision-making meeting studies (GRADE Working Group 2004). The five GRADE criteria: study limitations, consistency of effect, imprecision, indirectness, and publication bias) were used to judge the certainty of evidence. We justified our decisions to downgrade or upgrade the quality of studies based on the results and produced a summary of findings table based on GRADE assessment (see Appendix 12).

2.12 Data analysis

2.12.1 Harvest plots

Harvest plots were created in RStudio (<https://www.rstudio.com/>). The code used to create them is online, at: <https://github.com/CatherineFoster/Harvest-Plots>.

Studies were colour-coded according to ROB rating. Cochrane and ROBINS-I risk of bias judgments were grouped into three categories for ease of plotting; low, moderate and high risk of bias categories were used. Critical, serious (ROBINS-I) and high judgements (Cochrane) were all grouped into high, moderate (ROBINS-I) and unclear (Cochrane) judgements became

moderate, and the low risk of bias group applied to both risk of bias methods used.

Study types were differentiated into three categories according to bar height on the y-axis. Studies were categorised as either RCTs, experimental studies with efforts to improve group comparability (CS1) as described by (Craig et al. 2017), and experimental studies where no attempts to improve group comparability were reported in the paper (CS2) for example, no matching by group characteristics or regression adjustment. Group comparability was judged according to propensity score matching. Studies were then grouped according to the direction of effect (favouring the shared decision-making meeting, control or no difference) in relation to the research question. If outcome results were based on a formal statistical analysis, the direction of effect was automatically plotted based on this. If the statistical result was judged not to be meaningful, this is discussed in the results summary. Where the authors did not carry out statistical analysis, reviewers (CF, JO) agreed on the direction of effect to be included on the Harvest plot by first considering the author conclusion. Where the reviewers agreed that there was justification for the authors' conclusion of the trend direction this direction was included on the Harvest plot. In a small number of cases, the reviewers did not agree with the authors' conclusion regarding direction of effect and plotted the result as "No Difference" in all cases. The individual reasons for this, usually a small sample size making interpretation of effect unreliable, are stated in the results summary for each outcome. Please see section 3.4 for details of the numbered studies in the Harvest plots.

Harvest plots of data from all meeting types are first presented for each outcome, followed by data for FGCs only. Data for the other individual meeting types (FGDM, FUM, TDM, FTM, FTC) are shown in Appendix 13. The rationale for separating the studies was to avoid missing any potential differences in outcomes between meeting types by summarising the data as a whole. As FGCs are the original family-led model of shared decision-making from New Zealand, the approach most often now used in the UK and the most prevalent meeting type included in the review, it was decided



to present Harvest plots for FGCs only alongside the aggregated data on all meeting types to determine whether there were notable outcome differences for FGCs and other meeting types which may not be as family-led or participatory

2.12.2 Economic evaluation

The review was designed to capture and include all types of partial and full economic evaluations of shared decision-making meetings. The studies with economic data were grouped according to the approach taken. They could either be partial economic evaluations (cost analyses or cost-cost offset analyses) or they could be full economic evaluations that identify, measure and value costs and outcomes of the shared decision-making meetings with appropriate comparators. The different types of full economic evaluations include cost-effectiveness analyses, cost-utility analyses, cost-benefit analyses, cost-consequence analyses and cost-minimisation analyses; see Appendix 9 for a glossary of terms. The total number of partial and full economic evaluations was recorded, and the results summarised. The results were analysed to determine the number of studies that support the adoption of shared decision-making meetings on cost-effectiveness grounds, and where available an overall recommendation was made based on the results of both categories of economic analyses (partial or full economic evaluations).



3 RESULTS

3.1 Description of studies

3.1.1 Results of search

Database searches returned 4,454 records, while forensic searches identified 320 additional records. After de-duplication 2,680 records were

screened at title and abstract level and 120 full-text papers were assessed based on the eligibility criteria. In total, 32 papers satisfied the inclusion criteria and were thus included in the systemic review. A full list of the excluded studies can be found at Appendix 4.

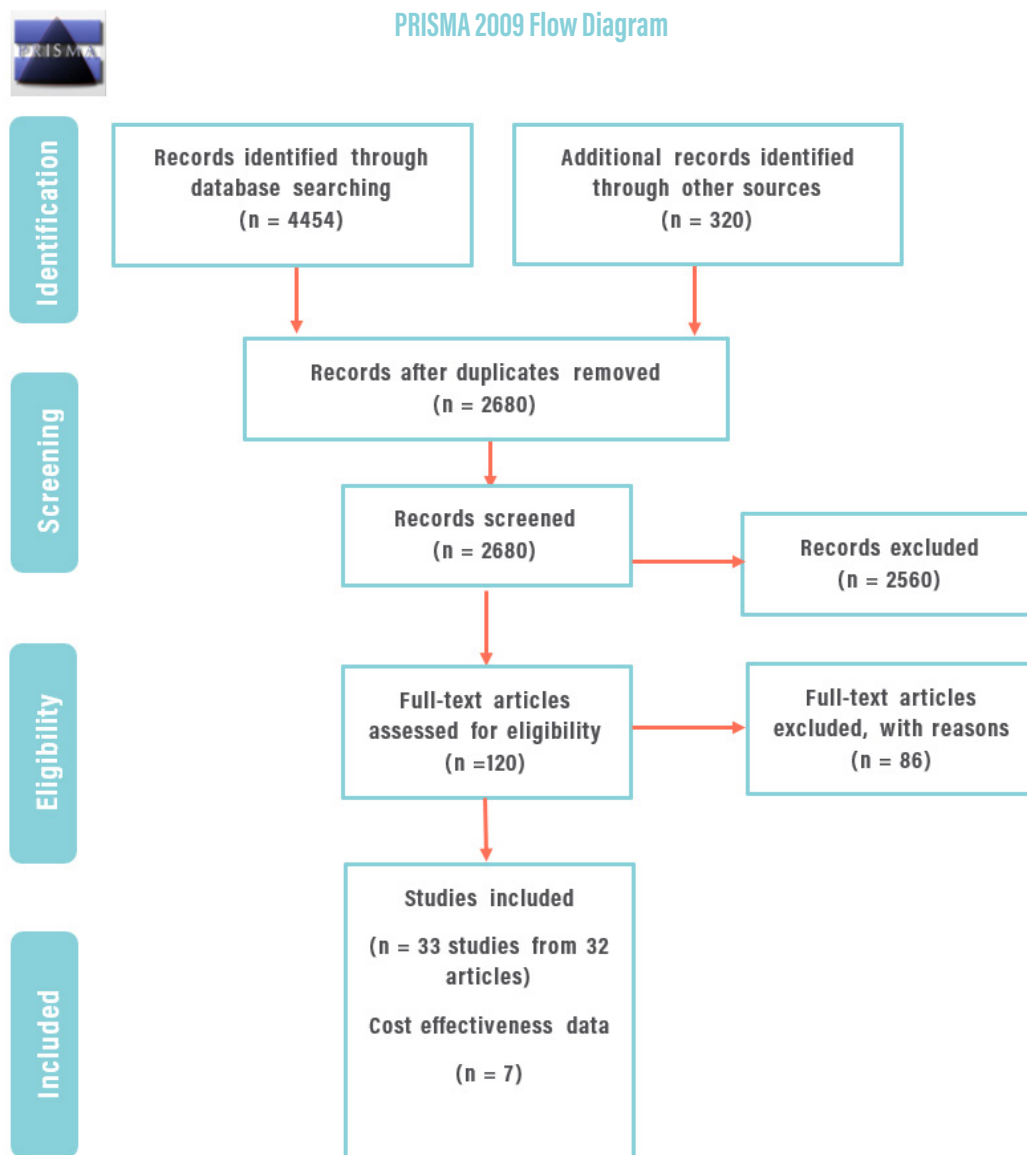


Figure 2. PRISMA flow diagram



The 32 articles included in this review reported on 22 published papers and 10 grey literature reports – related to 33 studies (see Appendix 5). Allan et al. (2015) contained two separate studies as did Dijkstra et al. (2016a), i.e. these two papers contained four separate studies. Allan et al. (2015) contained an RCT and quasi-experimental study and this differentiation is used when reporting the results. The studies contained in Dijkstra et al. (2016a) are differentiated as Study 4 and Study 5. In addition, two sources were used to extract data on one study (Partnership for Strong Families 2012; Perry et al. 2013), as outcome data on empowerment and satisfaction was more detailed in Partnership for Strong Families (2012). Of the 33 studies, seven reported on cost-effectiveness data and are captured in the economic analysis. One paper (Onrust et al. 2015) was included in the costs analysis only as it did not address any other outcomes this review investigated.

Eight of the included studies were RCTs and 25 were quasi-experimental studies. Nine of the quasi-experimental studies made efforts to improve the comparability between the shared decision-making meeting and comparison groups (CS1) e.g. through propensity score matching. Categorisation of studies can be found in full in Appendix 6.

Most studies were from the USA (n=23), with others from the Netherlands (n=6), the UK (n=2), Canada (n=1) and Sweden (n=1). Two papers were translated from Dutch into English for the purpose of this review (Wijnen-Lunenburg et al. 2008; Dijkstra et al. 2016a), please see Appendix 10.7 for study characteristics.

3.1.2 Sample population

The sample populations in the included studies varied in terms of the reason for referral to shared decision-making meetings and the level of risk of out-of-home care. The risk scenarios ranged from families with substance abuse problems (Huebner et al. 2012) to children who had been victims of physical or sexual abuse (Sundell and Vinnerljung 2004). Twenty-two included children still living at home but at risk of being removed, however, in a number of studies, children had already been removed and the outcomes focused

on re-entry to care rates (Chambers et al. 2016, Godinet et al. 2010) or reunification following the shared decision-making meetings (Pennell et al. 2010, Godinet et al. 2010). Several studies noted the possibility that more difficult cases were assigned to shared decision-making meetings rather than control services (Pennell and Burford 2000; Sundell and Vinnerljung 2004; Pennell et al. 2010; Crampton et al. 2011). The results may not be representative of the total population in some cases, or approximately matched in terms of case complexity as it is often more difficult to recruit these cases to research studies.

Age

The age of children included in the studies varied. Five studies specified the child's age in the inclusion criteria. The majority accepted children from birth to 17 or 18 years old, but one programme (the focus of two studies) was aimed specifically at children aged 2 to 12 years (Berzin 2006; Berzin et al. 2008)

The average age of children (as reported in 19 studies) ranged from 2.26 years (Hollinshead et al. 2017) to 10.46 years old (Dijkstra et al. 2016a - Part V).

Ethnicity

The ethnicity of children or parents was reported in 22 studies. In five studies over half the sample population were White, and in four over half the sample were African American.

The 'Ohana Conference program studied by Godinet et al. (2010) was targeted specifically towards native Hawai'ians and other Pacific Islander families only. The chief concern of the study by Crampton & Jackson (2007) was the disproportionate number of 'children of colour' in foster care, and FGDM as a potential alternative. No other studies focused on any particular ethnicities, even if there was a significant majority present in the sample population.

3.1.3 Shared decision-making meeting service

A number of different shared decision-making meeting models were included in this review. The majority were Family Group Conferences



(n=13; see Appendix 8 for a summary of the main components of the services). Two conferences, 'Eigen Kracht-conferentie' (Own Strength conference) and 'Ohana Conference' (Family or Kin Conference), were adaptations of the Family Group Conference model in the Netherlands and Hawaii respectively. Family Group Decision Making meetings were the next most frequent model described (n=9). In addition, studies which investigated Team Decision Making meetings (n=3), Family Team Meetings (n=3), Family Unity Meetings (n=1), and Family Team Conferencing (n=1) were also included.

In one study, Family Team Meetings were used as part of a larger service for families with substance misuse concerns (Huebner et al. 2012). Families taking part in the Sobriety Treatment and Recovery Teams (START) programme attended Family Team Meeting within 1-2 days after referral. They then received ongoing support and in-home services from the START team for an average of 14.2 months. The programme aimed to safely keep children at home where possible. As Family Team Meetings were one component of family support services in this instance, it is not possible to be sure any outcomes were solely due to the meetings. Inclusion of a study where shared decision-making meetings were part of another service is perhaps a limitation of the review, as discussed in section 4.3.

In the study by Chambers et al. (2016), Team Decision Making meetings were one of four strategies used as part of the Pomona Family First project. The other strategies were to find and maintain foster and kinship families who could support children and families locally; build community partnerships to better link families and services; and create self-evaluation tools using family outcome data to identify the area of progress and change. As with Huebner et al. (2012), effects cannot be isolated to the Team Decision Making meetings alone.

The study by Mason et al is of Leeds Family Valued, which also involved other aspects in

addition to FGCs. It is described as an approach to changing the whole system of children's services, which included training in restorative practice for a wide range of staff. It also included making FGCs mainstream in child welfare cases, on a scale said to be unprecedented in the UK, and providing new services to act on the recommendations of FGCs.

Eleven of the services were reported to include the four steps (referral, preparation, conference and implementation) typical of shared decision-making family meetings (see Appendix 8 for details of the components described for each study). However, full descriptions were often unavailable, particularly for 'implementation' which was not mentioned for 13 services and this was the case for both FGCs and other meeting types. The location of the service was rarely reported. Nine services specified where the service took place, and there was very little detail given, describing locations as 'neutral' or 'family friendly'. Twenty-three services were reported to include private family time.⁴

Families in the control/comparison groups received care as usual services. In general, the included studies provided little information about what 'care as usual' meant and what type of support families in control groups received. Seven studies provided descriptions of the usual care that was made available to families. These typically included service planning (Feldman 2017) and the production of care plans made by welfare workers in collaboration with the family (Sheets et al. 2009; Dijkstra et al. 2016a - Parts IV and V; Dijkstra et al. 2018a; Dijkstra et al. 2018b). Whilst similar to shared decision-making meetings in many ways, other types of meetings did not benefit from the presence of neutral facilitators and service providers, there was no family alone time and families were minimally prepared before the meetings (Partnership for Strong Families 2012; Perry et al. 2013), or families were greatly outnumbered in meetings by agency staff and other professionals (Sheets et al. 2009).

In one study, it was unclear if the comparison groups received 'care as usual' or attended a

4 Allan et al. (2015) – Larimer and Texas, Berzin (2006), Berzin et al. (2008) - Fresno and Riverside, Crampton & Jackson (2007), Dijkstra et al. (2016) - Parts IV and V, Dijkstra et al. (2018a), Dijkstra et al. (2018b), Godinet et al. (2010), Hollinshead et al. (2017), Munro et al. (2017), Partnership for Strong Families (2012) & Perry et al. (2013), Pennell & Burford (2000), Pennell et al. (2010), Sheets et al. (2009), Sundell & Vinnerljung (2004), Walker (2005), Weisz et al. (2006), Wheeler & Johnson (2003), Wijnen- Lunenburg et al. (2008), YMCA (2014)



similar type of meeting (a Family Team Meeting) as the service groups, and there were also concerns that a good number of service group families never received a FGC (Hollinshead et al. 2017). From one of the UK studies, which provided trend data, it was not clear what support the comparison groups from statistical neighbours (local authorities with similar characteristics) received, as it was noted that they were offered a FGC but it was not mandatory (Munro et al. 2017). Potential motivational differences can be seen for control group families who dropped out of the study (Dijkstra et al. 2016a - Part IV) or demanded an FGC, thereby over-riding the study protocol (Hollinshead et al. 2017).

3.2 Risk of bias assessment

The results of the risk of bias evaluations for all studies are reported in Appendix 10. Five of the eight RCTs were judged to have a high risk of bias. One RCT had a low risk of bias and two studies were unclear. The majority (n=19) of the quasi-experimental studies were considered to have a serious risk of bias. The remaining (n=6) had a moderate risk of bias. No quasi-experimental studies were rated as low or critical risk of bias.

3.2 Certainty of evidence assessment

According to the GRADE assessment, the body of evidence for the out-of-home care outcome, comparing shared decision-making meetings and care as usual (or equivalent for the control group) had low certainty. A judgement of low certainty indicates that the true effect may differ substantially from the estimate. Only 17 studies reported sample sizes, 14 studies were at serious risk of bias and there was a moderate to high degree of inconsistency in the results, such as effects in opposite directions (i.e. benefit and harm of shared decision-making meetings depending on what study is being considered).

The body of evidence concerning out-of-home care re-entry outcome was assessed as very low certainty. The evidence was downgraded due to the small number of studies, an inconclusive and high degree of inconsistency in the results, and methodological issues in the study designs such

as a lack of sample size calculations or power analysis or risk of bias.

The certainty in evidence for the outcome of reunification with family following a period in care was judged as low. The evidence was downgraded due to risk of bias of included studies, the heterogeneity in study methodology and sample sizes, inconsistency in the results, a weak direction of effect and the lack of sample size calculations or power analysis in their study designs.

The GRADE assessment of the certainty in evidence of effectiveness of shared decision-making family meetings on family empowerment was judged as low due to the same concerns. The evidence was downgraded due to the risk of bias, differences in the population, and methods for measuring empowerment, and a lack of sample size calculations or power analysis in study designs.

Finally, the certainty in the evidence on the effectiveness of the service on family satisfaction was judged very low. The evidence was downgraded due to the risk of bias assessment, the methodological limitations of the included studies – such as small number of studies and the lack of sample size calculations or power analysis, as well as the differences in how satisfaction was measured.

3.4 Effectiveness of shared decision-making family meetings

There was a high level of heterogeneity of included studies which precluded a meta-analysis. Therefore, systematic review findings summarised using Harvest plots (Ogilvie et al., 2008). For ease of plotting, studies were given number IDs which are displayed in the table below alongside the type of shared decision-making meeting evaluated in each study. The Harvest plots (Figure 3a-8b) display aggregated data for all meeting types, as well as FGC-only data for outcomes which >4 studies reported. Where the numbers were small, it was not meaningful to display separate Harvest plots for 1-2 FGCs.



Table 2: Number IDs for studies visualised in figures 3-8

Number ID	Author (Year)	Study Type
#1	Allan et al. (2015) Larimer	FUM & FGC
#2	Allan et al. (2015) Texas	FGC
#3	Beehlar (2016)	FTC
#4	Berzin (2006)	FGDM
#5	Berzin et al. (2008)	FGDM
#6	Chambers et al. (2016)	TDM
#7	Crampton et al. (2007)	FGDM
#8	Crampton et al. (2011)	TDM
#9	Dijkstra et al. (2016) Study 4	FGC (Eigen Kracht)
#10	Dijkstra et al. (2016) Study 5	FGC (Eigen Kracht)
#11	Dijkstra et al. (2018a)	FGC
#12	Dijkstra et al. (2018b)	FGC
#13	Feldman (2017)	FGDM
#14	Godinet et al. (2010)	FGC ('Ohana)
#15	Hollinshead et al. (2017)	FGC
#16	Huebner et al. (2012)	FTM
#17	Lambert et al. (2017)	FTM
#18	Mason et al. (2017)	FGC
#19	Munro et al. (2017)	FGC
#20	Onrust et al. (2015)	FGC
#21	Pennell & Burford (2000)	FGDM
#22	Pennell et al. (2010)	FTM
#23	Perry et al. (2013) & Partnership for Strong Families (2012)	FTC
#24	Sheets et al. (2009)	FGDM
#25	Sundell & Vinnerljung (2004)	FGC
#26	Teal (2013)	TDM
#27	Titcomb & LeCroy (2005)	FGDM
#28	Walker (2005)	FGC ('Ohana)
#29	Wang et al. (2012)	FGC
#30	Weisz et al. (2006)	FGC
#31	Wheeler & Johnson (2003)	FGC
#32	Wijnen-Lunenburg et al. (2008)	FGC
#33	YMCA (2014)	FGC



3.4.1 Primary outcome: Number of children entering out-of-home care

GRADE Assessment: Low Certainty of Evidence

The number of children entering out-of-home care was assessed in a total of twenty studies (five RCTs, five CS1 studies and ten CS2 studies). The studies shown in Figure 3a investigated out-of-home placement rates for approximately 620,711 participants (609,114 in shared decision-making meeting services and 11,597 in the control groups). This number includes eighteen out of twenty studies, as sample size was not explicitly stated in two studies (Mason et al. 2017; Munro et al. 2017). It was difficult to estimate exact numbers as some studies report the number of children who were the focus of services received, with more than one child being from the same family, and other studies report the number of families which may include one or more children. To complicate matters further, some studies reported the sample size for the service group only (Allan et al. 2015). Godinet et al. (2010) reported "cases" which was taken to mean families. The sample size reported above is therefore a mixture of numbers of families and children; these figures can be assumed to be a minimum as each family would have included at least one child. The main source of the difference in size between groups was due to a single study (Lambert et al. 2017) which included a sample of 604,498 in the control group and 8682 receiving shared decision-making meeting services.

The results from nine of the twenty studies were based on outcomes from statistical analysis examining between group differences. The remaining eleven did not formally test group differences or reported post-service within-group changes. For eight of these eleven studies, the reviewers agreed with the trend direction stated by the authors despite the low strength of evidence in the absence of a formal analysis.

In three of the studies (Pennell and Burford 2000; Dijkstra et al. 2016a; Mason et al. 2017) the authors stated that the results favoured the shared decision-making meeting service while the reviewers agreed that there was no basis to conclude any group difference. The reasons for these conclusions were that the numbers were

too small to evidence a group difference (Pennell and Burford 2000; Dijkstra et al. 2016a) or the methods used did not allow a group difference to be concluded (Mason et al. 2017).

The five RCTs concluded either no group differences or that fewer children entered care following care as usual services. Of the five CS1 studies, two CS1 reported fewer children entering care following shared decision-making meetings, two reported fewer entered care after care as usual services and one no difference. Of the ten CS2 studies, five reported no group differences and five reported fewer children entering care in the shared decision-making meeting groups. The evidence is therefore mixed, and the balance of evidence does not strongly favour shared decision-making meetings as an effective service to reduce the number of children entering out-of-home care. Overall, the studies have generated conflicting findings making it difficult to confirm or refute the role of shared decision-making meetings on children's out-of-home care.



Which Service (Shared Decision-Making Meetings or Control Services) Had Fewer Out of Home Placements?

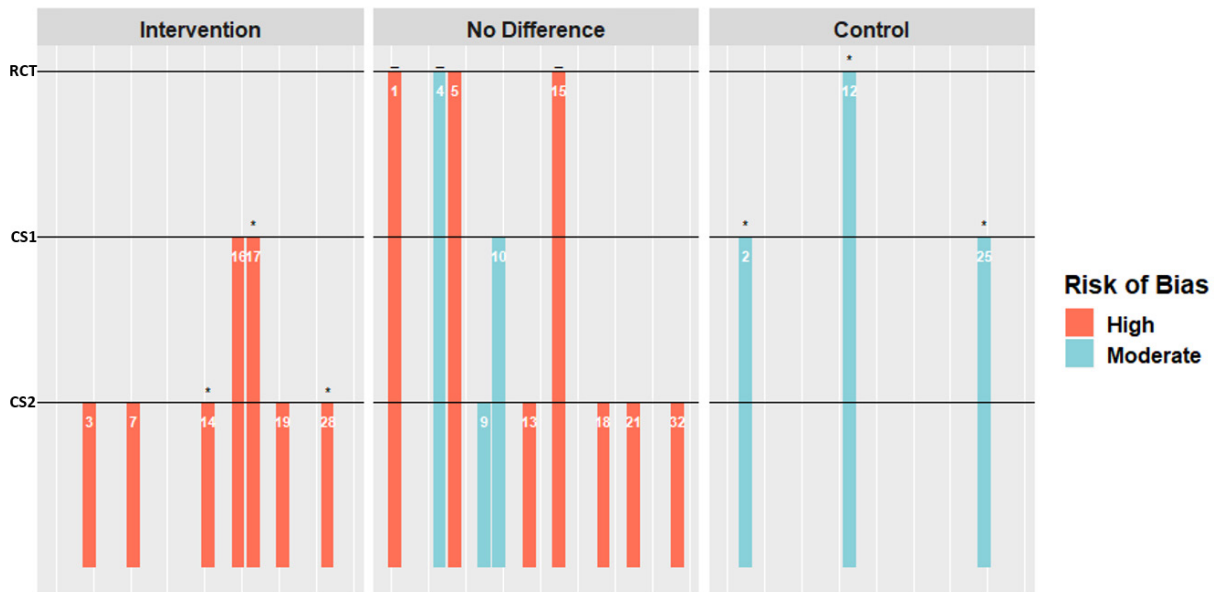


Figure 3a. Harvest plot showing the direction of effect for care entry. RCT = Randomised-controlled trial. CS1 = Comparative study with efforts to match groups or correct for differences. CS2 = Comparative Study without efforts to make groups comparable. Numbered studies are listed in Section 3.4. Studies are grouped according to whether the results suggest fewer children enter care in the service or control groups or show no difference between groups. * indicates a statistically significant difference between groups, _ indicates no statistical difference and studies without markers did not include statistical analysis.

Do Family Group Conferences Reduce Care Entry When Compared with Control Services?

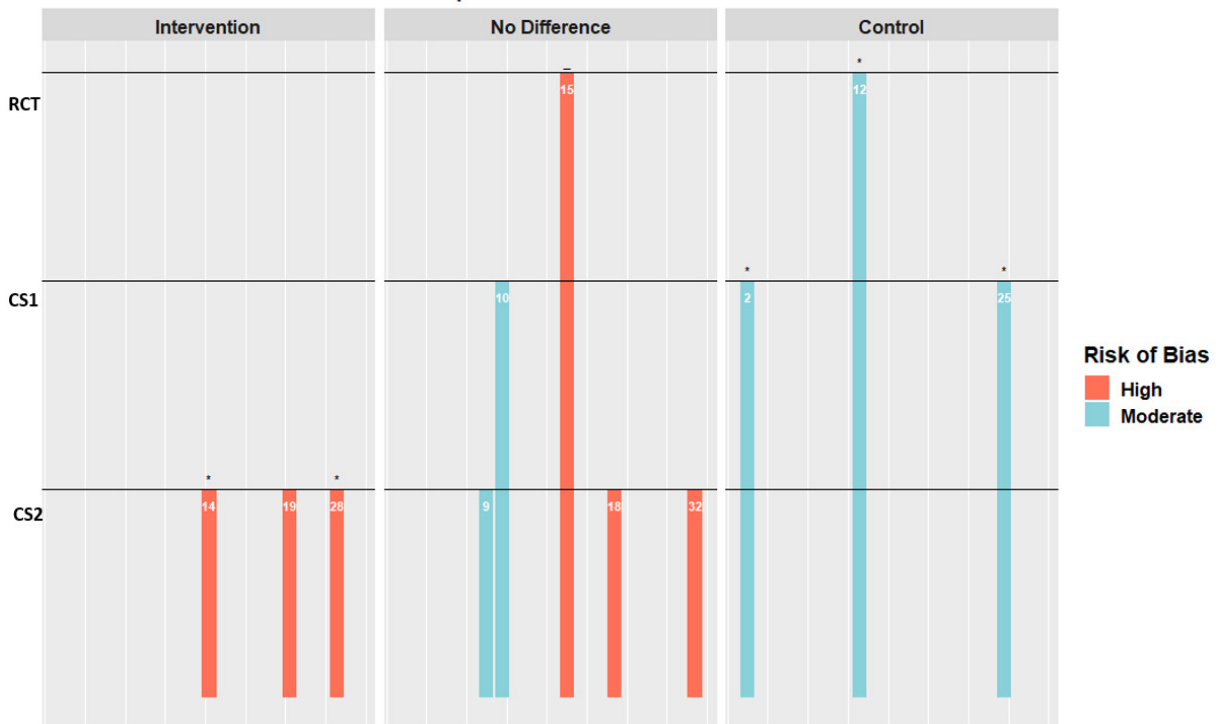


Figure 3b. Harvest Plot showing the same data as 3a for FGC services only (n=11 for care entry)



When we looked at FGCs only (Figure 3b), the results were similarly mixed, and it was not apparent that the inclusion of non-FGC meeting variations was undermining the apparent effectiveness of shared decision-making meetings (see Appendix 13 for the direction of effect for remaining study types).

3.4.2 Primary outcome: Number of children re-entering out-of-home care

GRADE Assessment: Very Low Certainty of Evidence

The number of children re-entering out-of-home care was assessed in three studies (one RCT, one CS1 and one CS2) all of which found a statistically significant difference between groups. The studies investigated out-of-home re-entry rates for 931 participants (513 in the shared decision-making meeting services and 418 in the control groups). While Chambers et al. (2016) and Godinet et al. (2010) reported family numbers, Partnership for Strong Families (2012) (same study as Perry et al. (2013)) reported individual children.

The RCT (Partnership for Strong Families 2012; Perry et al. 2013) of the FTC model found that fewer children were likely to re-enter care following care as usual services while the comparative studies both found that fewer children entered care following shared decision-making meetings (Figure 4).

However, it must be noted that in the studies by Chambers et al. (2016), evaluating a TDM model, and Godinet et al. (2010), evaluating 'Ohana conferencing, the numbers were too small to show strong evidence for shared decision-making meetings despite the statistically significant difference. In the study by Chambers et al. (2016), no children re-entered care in the shared decision-making meeting service and four re-entered in the control group, a sample too small to conduct a t-test by conventional standards. In the paper by Godinet et al. (2010) the rate of re-entry was 1.23 and 1.61 for the intervention and control groups respectively and the sample size for the control group was not specified. As reported by Partnership for Strong Families (2012)/Perry et al. (2013) foster care re-entry rates were

assessed between three family team pathways, two considered shared decision-making meeting service groups (Pathway 2 and Pathway 3) and a traditional pathway, considered a control group (Pathway 1). Re-entry was only higher than the control group for Pathway 3 which, unlike Pathway 2, included family alone time. The results are inconclusive with two small comparative studies supporting the benefit of shared decision-making meetings reducing the care re-entry rate, contradicted by a larger RCT that concluded that shared decision-making meetings with private family time (typical of most shared decision-making meeting models) were less effective than Pathways 1 and 2.

As there were only three studies reporting this outcome and one (study #14) using the FGC model which found lower re-entry rates for she shared decision-making meeting group (Godinet et al. 2010), it was not possible to separate by study type.

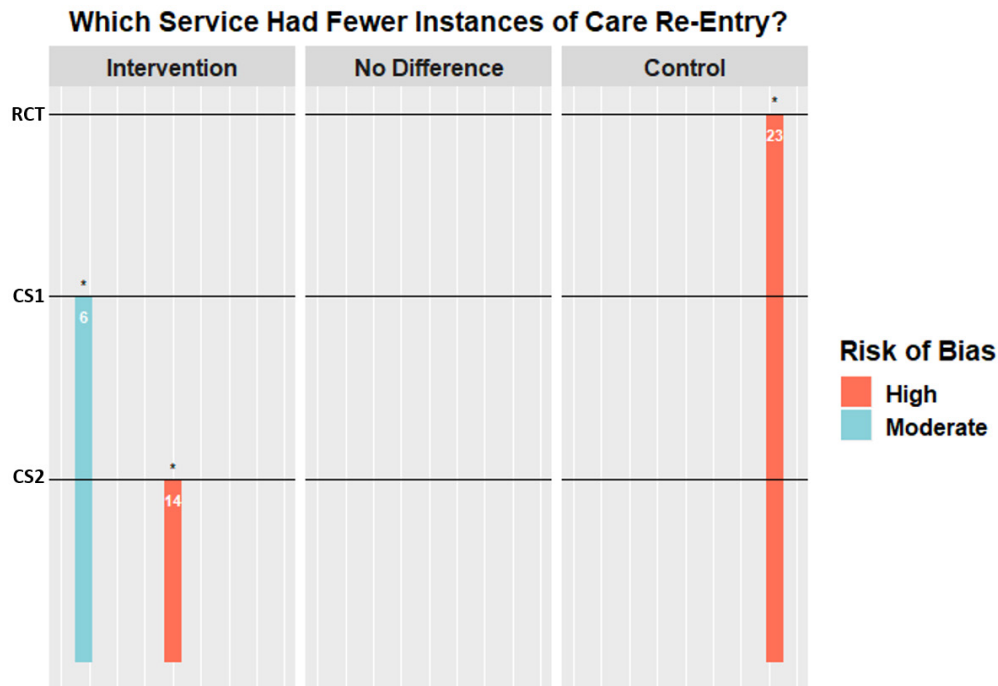


Figure 4. Harvest plot showing the direction of effect for care re-entry. RCT = Randomised-controlled trial. CS1 = Comparative study with efforts to match groups or correct for differences. CS2 = Comparative Study without efforts to make groups comparable. Numbered studies are listed in section 3.4. Studies are grouped according to whether the results suggest fewer children re-enter care in the service or control groups or show no difference between groups. * indicates a statistically significant difference between groups, _ indicates no statistical difference and studies without markers did not include statistical analysis.

3.4.3 Primary outcome: Reunification rate

GRADE Assessment: Low Certainty of Evidence

Reunification rates of children with parents or guardians were assessed in 13 studies (3 RCTs, 4 CS1 and 7 CS2 studies), including Wheeler & Johnson (2003) which only reported the shared decision-making meeting services sample size. The sample size was not reported by (Crampton et al. 2011). The studies which reported sample sizes together investigated reunification rates for 88,405 participants (10,475 in the shared decision-making meeting services and 77,930 in the control groups).

Results from ten of the fourteen studies were based on a statistical analysis giving a significant difference or lack thereof between groups, the remaining four did not test for statistical group differences. The reviewers did not agree with the conclusions in Weisz et al. (2006), where statistical analysis was not used, due to the small sample size: 7/33 (21%: 95% CI 10.7-37.8%) children

were reunified in the FGC group and 4/33 (12%: 95%CI 4.8-27.3%) in the control group which the reviewers agreed did not justify a conclusion in favour of the FGC service.

One RCT (Partnership for Strong Families 2012) found greater reunification following control services, two did not find a group difference. Of the CS1 studies, three found greater reunification for the service groups, one found no difference. For the CS2 studies three found greater reunification for the service groups, four found no difference (Figure 5). Despite the large overall sample sizes, as with the previous outcomes, heterogeneity in study methodology and sample sizes there is not a strong direction of effect. Five of the 13 studies reporting reunification rate used the FGC model, with two finding greater reunification in the FGC group and three finding no difference. Looking at FGCs only does therefore not change the mixed findings overall (Figure 5b). Results for each individual meeting type can be found in Appendix 13.

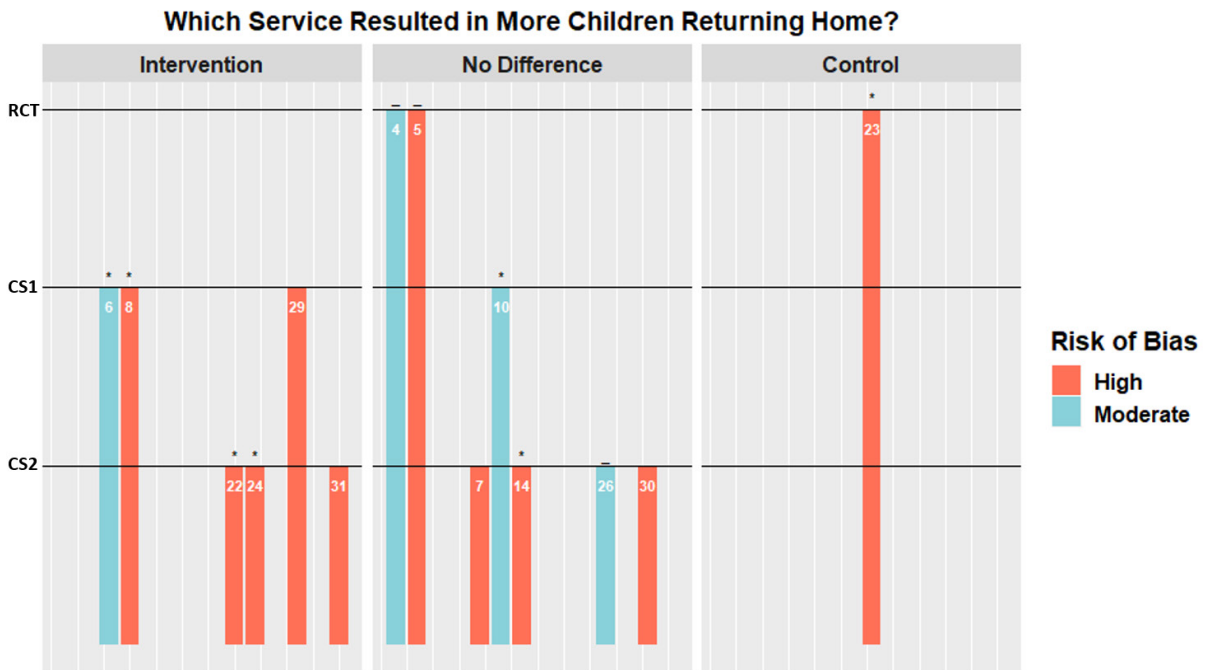


Figure 5a. Harvest plot showing the direction of effect for reunification following entry to care. RCT = Randomised-controlled trial. CS1 = C Comparative study with efforts to match groups or correct for differences. CS2 = Comparative Study without efforts to make groups comparable. Numbered studies are listed in Section 3.4. Studies are grouped according to whether the results suggest less children reunify with family in the service or control groups or show no difference between groups. * indicates a statistically significant difference between groups, _ indicates no statistical difference and studies without markers did not include statistical analysis.

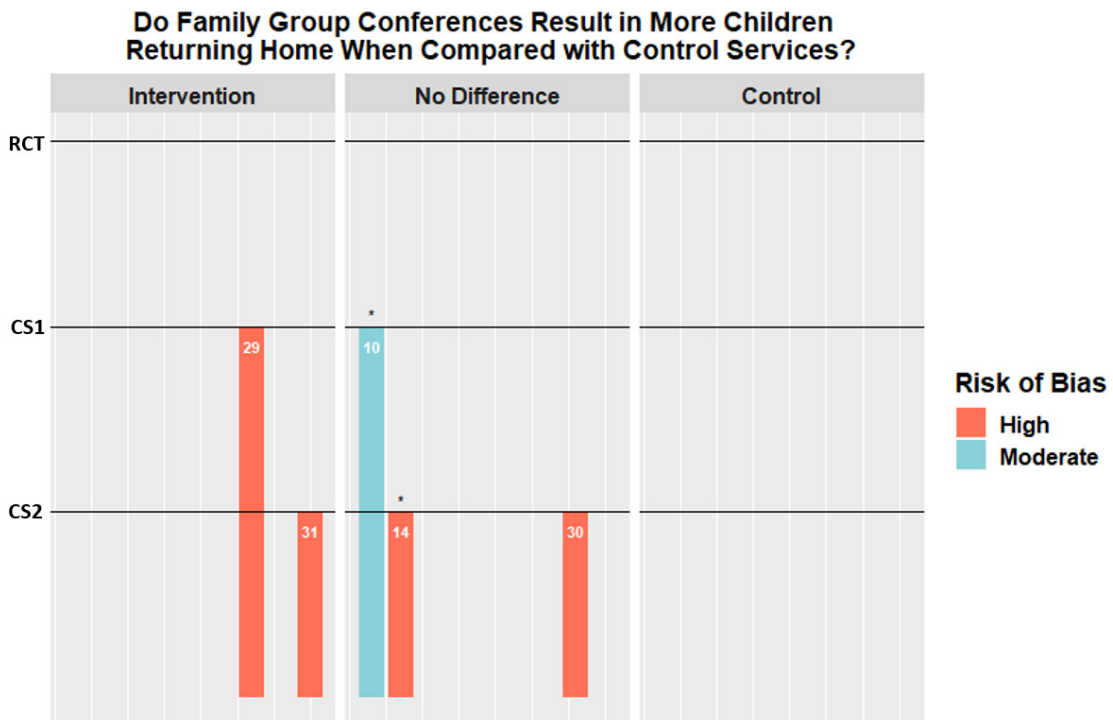


Figure 5b. Harvest plot showing the same information as 5a for FGC meetings only (n=5)



3.4.4 Primary outcome: Empowerment

GRADE Assessment: Low Certainty of Evidence

Only four studies investigated empowerment differences between families who participated in shared decision-making meetings and those receiving usual services. The studies investigated empowerment outcomes for an estimated 2,415 participants (1999 in the shared decision-making meeting services and 416 in the control groups). One study (Partnership for Strong Families 2012) reported separate scores for mothers and relatives and within that the sample size range of each of the three groups (Pathways 1, 2 and 3). For this study, the mean of the number of mothers and relatives who participated in the empowerment questionnaire was included in the calculation of estimated sample size.

Three of the studies were RCTs, two of which investigated FGCs and found no statistically significant differences, and one which investigated Family Team Conferencing did not conduct a statistical comparison but concluded that there

was no difference between groups. One Family Group Decision Making study (Sheets et al. 2009) found that parents and relatives in the intervention group rated their empowerment as higher than the control group who took part in permanency planning meetings. However, although this difference was statistically significant, in practice it was a 0.2-point and 0.3-point difference on a 5-point scale measuring empowerment for parents and relatives respectively. Whether this is a meaningful difference is open to interpretation, especially as there were substantially different numbers within the groups (303 parents, 636 relatives in the service, 121 parents and 50 relatives in the control group).

Overall the studies suggest no difference in empowerment between shared decision-making meetings and control services. Again, as there was a small number of studies for this outcome it was not necessary to separate by study type. Both FGC studies found similar rates of empowerment between groups.

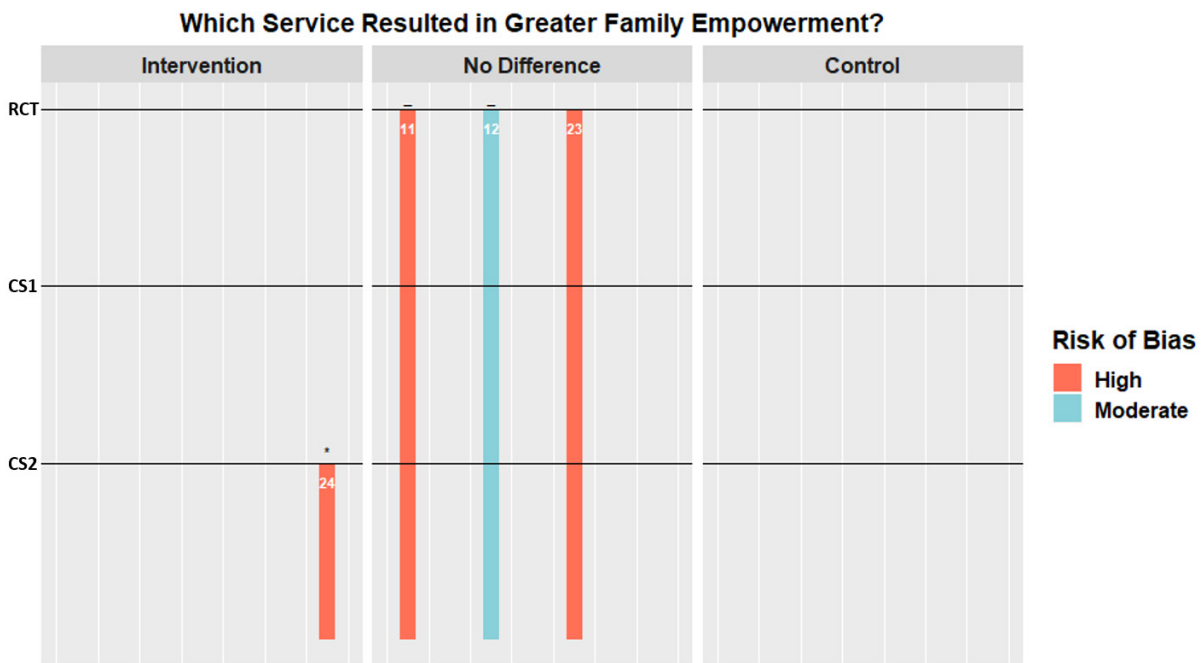


Figure 6. Harvest plot showing the direction of effect for empowerment. RCT = Randomised-controlled trial. CS1 = Comparative study with efforts to match groups or correct for differences. CS2 = Comparative Study without efforts to make groups comparable. Numbered studies are listed in Section 3.4. Studies are grouped according to whether the results suggest participants feel more empowered by services received in the service or control groups or show no difference between groups. * indicates a statistically significant difference between groups, _ indicates no statistical difference and studies without markers did not include statistical analysis.



3.4.5 Primary outcome: Satisfaction

GRADE Assessment: Very Low Certainty of Evidence

Levels of satisfaction were assessed in four studies (one RCT and three comparative studies) for 1,509 participants in total (1,199 in the shared decision-making meeting services and 310 in the control groups). Three studies reported results based on statistical analysis, two comparative studies of Family Group Decision Making reporting greater satisfaction in the intervention groups and the RCT of Family Team Conferencing reporting no group differences. The remaining FGC study (Walker 2005) which did not use a statistical comparison reported trend data based on a sample size of 17 in the shared decision-making meeting groups and 13 controls with 41% of the shared decision-making meeting group reporting a positive experience compared to 23% of the comparison group. It was decided that this study

would be included as a "No Difference" study on the Harvest plot given the small sample and lack of formal analysis.

It should be noted that in Feldman (2017), eight satisfaction questions were answered by participants and only one question ("How satisfied are you with the amount of help you received?") of the eight came out with the intervention group having a significantly higher score than the control group. In the Family Group Decision Making study by Sheets et al. (2009), all satisfaction ratings ranged from 4.1-4.5 for both groups on a 5 point scale, meaning that the actual differences are relatively small, despite a statistically significant difference ($p < 0.001$) between those participating in the service and control groups.

This outcome is based on only four studies, and the mixed evidence does not provide strong evidence that satisfaction is higher following any shared decision-making meetings than control services.



Figure 7. Harvest plot showing the direction of effect for satisfaction. RCT = Randomised-controlled trial. CS1 = Comparative study with efforts to match groups or correct for differences. CS2 = Comparative Study without efforts to make groups comparable. Numbered studies are listed in Section 3.4. Studies are grouped according to whether the results suggest participants feel more satisfied with services received in the service or control groups or show no difference between groups. * indicates a statistically significant difference between groups, _ indicates no statistical difference and studies without markers did not include statistical analysis.



3.4.6 Secondary outcome: Referrals and re-referrals for child maltreatment

GRADE not carried out for secondary outcomes

Referral and re-referral rates for maltreatment were assessed in 16 studies (6 RCTs, 5 CS1 and 5 CS2 studies). The studies investigated referrals and re-referral rates for 6,352 participants (2616 in the shared decision-making meeting services 3736 in the control groups), however this number does not include the control samples from Allan et al. (2015) and Wheeler & Johnson (2003) where the sample sizes for the comparison groups were not reported.

Four of the studies did not report results of any statistical analysis. One of these studies, Crampton and Jackson (2007) stated in the discussion that fewer service group participants received maltreatment referrals following case closure, but only stated that five referrals occurred in the comparison group. If there were any referrals for the service group, this is not reported. This study

is therefore not included in the Harvest plot as results could not be verified.

The six RCTs, two CS1 and two CS2 studies with mixed ROB all reported no difference in referral and re-referral rates between groups, one CS1 study found lower rates in the control group, one CS1 study found that referral rates were lower in the shared decision-making meeting services along with three CS2 studies. In this case the evidence would support the finding that shared decision-making meetings are no more effective in reducing referrals for child maltreatment when compared to control services. Put another way, we can conclude there is no evidence of increased risk of re-referral for child maltreatment when shared decision-making family meetings are used. Looking at the eight FGC studies alone (Figure 8b), the evidence is again mixed, with only one study finding lower referral or re-referral rates in the control group, and the remaining seven reporting no group differences.

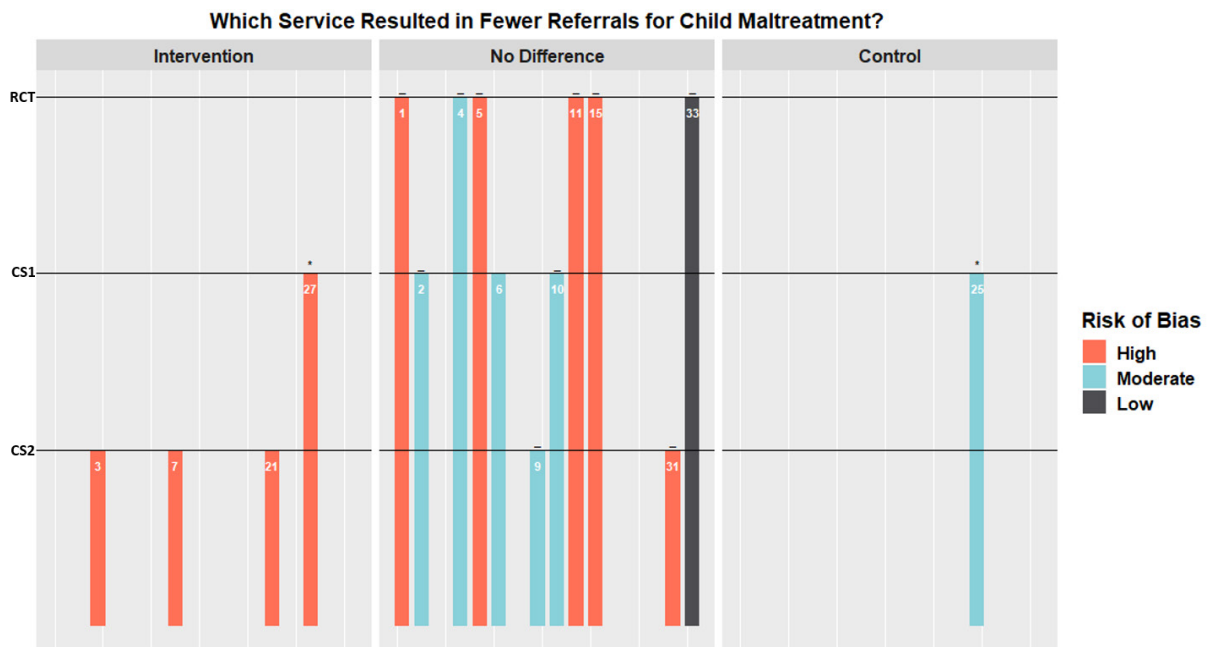


Figure 8a. Harvest plot showing the direction of effect for referrals or re-referrals for maltreatment. RCT = Randomised-controlled trial. CS1 = Comparative study with efforts to match groups or correct for differences. CS2 = Comparative Study without efforts to make groups comparable. Numbered studies are listed in Section 3.4. Studies are grouped according to whether more children are referred or re-referred to CPS for maltreatment in the service or control groups or show no difference between groups. * indicates a statistically significant difference between groups, _ indicates no statistical difference and studies without markers did not include statistical analysis.

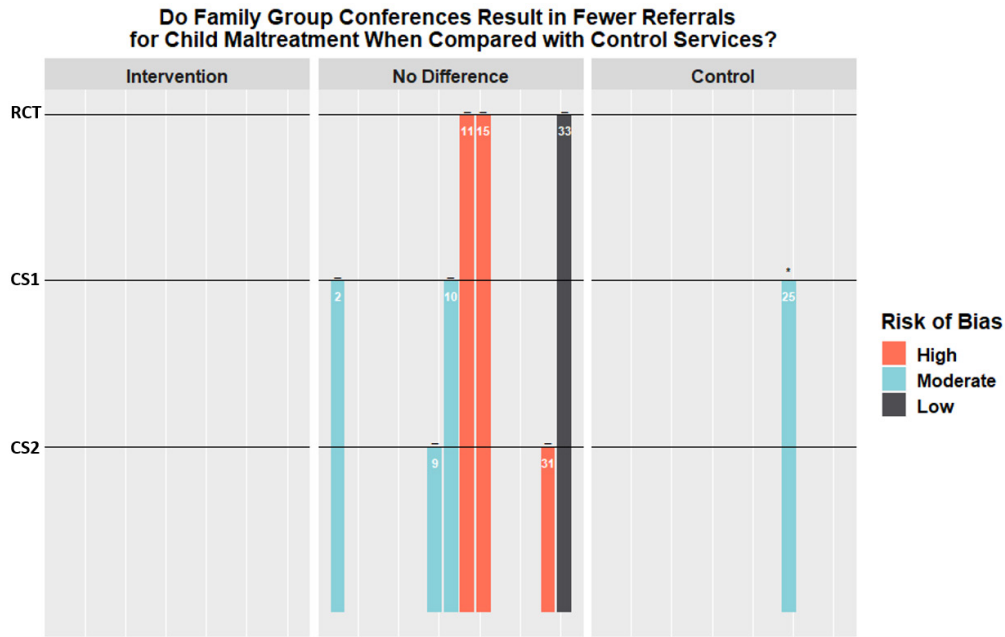


Figure 8b. Harvest plot showing the same data as Figure 8a for FGCs only

3.5 Cost-effectiveness of shared decision-making family meetings

Seven studies (Huebner et al. 2012; Partnership for Strong Families 2012; Onrust et al. 2015; Dijkstra et al. 2016a; Mason et al. 2017; Munro et al. 2017; Dijkstra et al. 2018b) included a form of economic analysis of shared decision-making meetings. Five studies were partial economic evaluations, three of these (Partnership for Strong Families 2012; Dijkstra et al. 2016a; Munro et al.

2017) were cost analyses, comparing costs in the shared decision-making meeting services to costs in the comparator group, and two (Huebner et al. 2012; Mason et al. 2017) were cost-cost offset analyses that offset the cost of the service against the costs saved as a result of the service. The final two studies (Onrust et al. 2015; Dijkstra et al. 2018b) were full economic evaluations in the form of cost-effectiveness analyses. A summary of the economic analysis approach and the authors' overall conclusion for each study is given in Table 3 below.

Table 3. Summary of the economic data

Study	Approach	Author's Conclusion
Partial economic evaluations		
Dijkstra et al. (2016a)	Cost analysis	Legal costs do not differ Healthcare costs greater in the shared decision-making meeting group
Huebner et al. (2012)	Cost-cost offset analysis	Cost-saving
Mason et al. (2017)	Cost-cost offset analysis	Cost-saving
Munro et al. (2017)	Cost analysis	Lower weekly cost of care per child in the shared decision-making meeting group
Partnership for Strong Families (2012)	Cost analysis	No difference in service costs
Full economic evaluations		
Dijkstra et al. (2018b)	Cost-effectiveness analysis	Not cost-effective
Onrust et al. (2015)	Cost-effectiveness analysis	None made



3.5.1 Cost analyses

All studies in Table 3 compared the costs of shared decision-making meetings to the costs of usual care except for the Partnership for Strong Families (2012) study. This study compared the costs of three versions of the shared decision-making meetings referred to as family team conferences (FTC):

1. FTC as usual: Family care counsellor (FCC) meets with the family to conduct the FTC. FTC facilitators and service providers are not invited to the FTC. There was no family alone time and families were not prepared for the meeting
2. FTC new: The FCC and FTC facilitator meet with the family. Service providers are also invited to the meeting and the families are adequately prepared for the meeting.
3. FTC new and family time: All the characteristics of 'FTC new' plus alone time for the family

The Partnership for Strong Families (2012) study did not focus on the costs of the service itself but on the services delivered to families in each of the three arms, these included the costs for domestic violence, mental health, parenting classes and supports, substance abuse and other services. The largest costs per family were for mental health services however, there was no statistical difference between these costs and the cost of other services between the three shared decision-making groups. All costs were reported in US \$, however the price date is unknown.

Dijkstra et al. (2016a) compared the costs for families where FGCs were offered and implemented to the costs for families that received usual care. They considered three categories of costs across both groups of families; the cost to deliver shared decision-making meetings, in addition to the healthcare costs and legal care costs incurred by the families. Since this programme evaluated the service across seven youth care institutions, the cost of FGCs differed depending on the methods adopted for implementation. The additional cost of FGCs compared to usual care ranged from €0 to €4000. Healthcare costs were significantly higher in the shared decision-making meeting

group, however there was no significant difference in legal costs between the two groups. All costs were reported in Euros, the price date was not specified.

The final cost analysis by Munro et al. (2017) also compared the costs for families that received FGCs to those that did not. In particular, this study estimated the average weekly cost of care per child across the two arms. Munro et al. (2017) estimated a lower weekly average cost of care for families that are involved in FGCs, £381 compared to £447. All costs were reported in UK £ and valued at 2014 prices.

3.5.2 Cost-cost offset analyses

The study by Huebner et al. (2012) focused on families where substance abuse is present; it estimated the cost of the programme (of which shared decision-making meetings were only one aspect, alongside family mentors and substance misuse treatment) and the costs offset as a result of state out-of-home placements avoided. The cost of the programme was given as \$2,676,188, even though Huebner et al. (2012) indicate that this value does not include the first year initiation costs, the time period covered is unclear and the individual costs that make up the total cost value are not described. The cost of each state care placement was valued at \$30,000; this assumes an average 15.8 months state care duration and an average daily rate of foster care at \$31.28. All costs were reported in US \$ and valued at 2010 prices.

There were three groups of children considered by Huebner et al. (2012), children that are accepted in the START programme and have a group family meeting as part of their management, children that are referred to the programme but are not accepted due to capacity issues, and children that are not considered for the programme. The proportion of children that entered out-of-home care were 21%, 23% and 42%, respectively. This was an unanticipated outcome since it implies that those referred and not accepted have similar rates of care placement to children that are accepted. The authors attempt to justify this by suggesting that social workers who were not involved in the programme learned about the service through social workers that were involved and adopted



some of the methods used by the FGC workers in their cases. Appropriate methods of blinding in future studies are needed to ensure that results are reflective of the service given. Nonetheless, Huebner et al. (2012) conclude that the START programme is cost-saving. If only children accepted are included in the analysis, then for every \$1 spent on the START programme \$1.07 was saved in state care placement costs. When the children referred and not accepted were also considered then this value increased to \$2.22 for each \$1 invested in the programme.

The second cost-cost offset analysis carried out by Mason et al. (2017) compared the core FGC component of the Leeds Family Valued programme to usual care i.e. social work involvement without an FGC. The other strands of the Leeds Family Valued programme were not considered in the analysis. The cost estimated for each arm included the cost of identifying families that need support, assessing their needs, the service itself followed by the stage of reviewing achievements. All cost data were provided by Leeds County Council and the sources of unit costs are given. Costs are reported in UK £ but the price date is unclear. The majority of costs were referenced at 2008 prices except for the cost of the FGC service itself, this was referenced at 2011 prices. All costs were combined without adjustments suggesting that the final cost may not accurately reflect the true cost of FGCs. The cost of providing an FGC service was estimated as £2,418 per family and for usual care services the cost was £1,943 per family. Even though the cost of FGCs are greater, this is offset by the time that each family spends in the social care system so that FGCs overall are cost saving. Mason et al. (2017) found that FGC families spent on average 10 weeks less in the social care system than families that received usual social work services. The cost per month of keeping a family in the social care system was £302, which is equivalent to a £755 saving per family as a result of time saved.

3.5.3 Cost-effectiveness analyses

Two studies (Onrust et al. 2015; Dijkstra et al. 2018b) carried out a cost-effectiveness analysis to determine the cost-effectiveness of FGCs. Both

studies were based in the Netherlands. Onrust et al. (2015) evaluated FGCs as a service for families where either a child or a parent has an intellectual disability and Dijkstra et al. (2018b) evaluated the use of FGCs for families with multi-complex problems that were referred to the child and youth protection services.

As both studies were full economic evaluations, the quality of their reporting was assessed against the CHEERS checklist. Four out of the 24 items in the checklist were not relevant: the application of discount rates (Attema et al. 2018) was not necessary as both studies adopted a short time horizon; the measurement and valuation of preference-based outcomes (York-Health-Economics-Consortium 2016) as these do not apply in this context; and the choice of model and model assumptions, since neither study adopted decision modelling-based approaches. The reporting quality was assessed against the remaining 20 CHEERS criteria. Dijkstra et al. (2018b) scored positively on all 20 items, indicating that it was a good quality economic evaluation. Onrust et al. (2015) can also be considered a good quality economic evaluation. It scored positively on 18 out of the 20 checklist items but did not meet the criteria for two items - the perspective of the study was not explicitly defined, however the study can be assumed to adopt a health and social care perspective based on the types of costs listed and included in the evaluation. In addition, the price date of all costs was not given. These should be clearly stated, and any adjustments made clearly reported. See Appendix 9 for a glossary of economic terms.

Dijkstra et al. (2018b) measured costs and outcomes at 6 and 12 months, with three outcomes considered in their cost-effectiveness analysis, namely child maltreatment risk, empowerment and social support. FGCs were as effective as usual care in reducing the risk of child maltreatment, empowering parents and the level of social support given at 12 months. In addition to the costs of meetings, three groups of costs were measured, healthcare, non-healthcare and child welfare costs. All costs were reported in Euros at 2015 prices.



Dijkstra et al. (2018b) estimated incremental cost-effectiveness ratios (ICERs) for each outcome, this is the value of additional investment required to achieve a one unit gain in outcome (McCabe et al. 2008). At 12 months, the ICER for social support is €13,335, for empowerment it is €28,337 and for child maltreatment, shared decision-making meetings are dominated in health economic terms, i.e. they result in higher costs and lower effects, on average, so no ICER is presented. Cost-effectiveness is determined by comparing the estimated ICER to a pre-defined cost-effectiveness threshold, the maximum value a decision maker attaches to an outcome. A visual representation of this comparison can be presented on a cost-effectiveness plane, as illustrated in Figure 9,

the difference in effectiveness between the intervention and its comparator is plotted on the x-axis and the difference in costs on the y-axis. Interventions with an ICER that fall in the top left quadrant of the plane are rejected on cost-effectiveness grounds, they are associated with an increased cost and reduced effectiveness, on average. Interventions that have an ICER that falls in the bottom right quadrant are accepted, they have greater benefits and result in cost savings, on average. For the remaining two quadrants, decisions are made based on the threshold defined by decision makers, recognising that investment is needed to achieve greater effects. ICERs that fall below the threshold line are cost-effective and vice-versa.

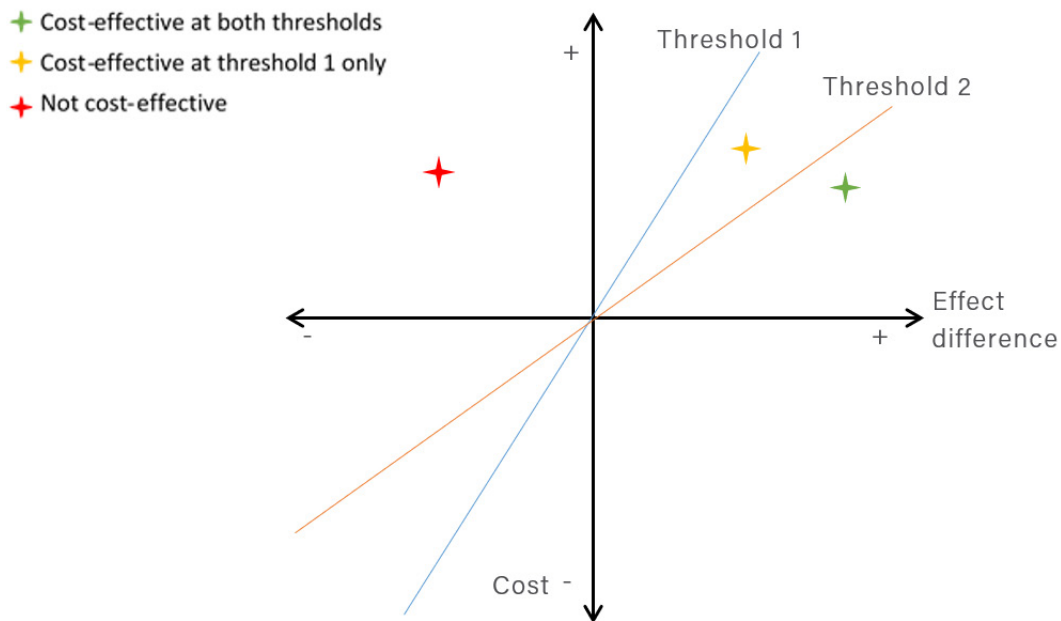


Figure 9. Cost-effectiveness plane



With further mathematical techniques it is possible to generate a large number of ICER estimates for each outcome to take account of variation around the underlying costs and outcomes. These ICERs can be summarised to generate a probability of cost-effectiveness at each threshold i.e. the proportion of ICERs that fall below the threshold.

Dijkstra et al. (2018b) estimate the probability of cost-effectiveness by generating 1000 ICER

estimates for each outcome. An example of the cost-effectiveness plane for the social support outcome is given in Figure 10, each blue point represents an ICER estimated for this outcome. The results in Figure 10 were converted to a cost-effectiveness probability at two thresholds; €0 and €10,000. Table 4 provides a summary of these probabilities for social support in addition to the other two outcomes measured.

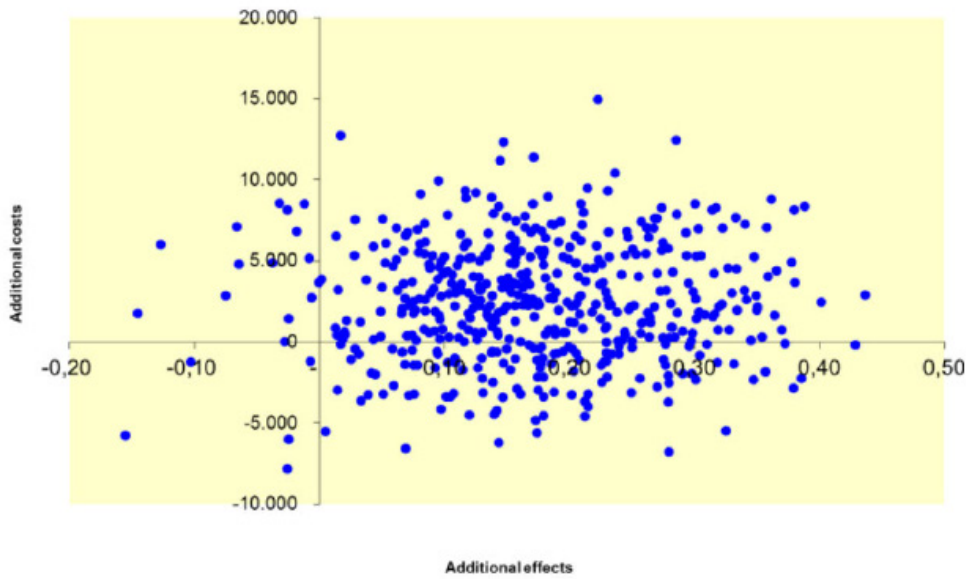


Figure 10. Cost-effectiveness plane: Social support

Table 4. Cost-effectiveness probability at 12 months

Outcome	Threshold	
	€0	€10,000
Child maltreatment	30%	33%
Empowerment	26%	36%
Social support	26%	46%



The results indicate that shared decision-making meetings may not be a cost-effective alternative to usual care due to the low probabilities reported in Table 4. However, the probability of cost-effectiveness increases with an increasing cost-effectiveness threshold and justification for setting the maximum threshold at €10,000 was not given. As the authors have not clearly stated a pre-defined cost-effectiveness threshold since this is most likely unknown, they should aim to report cost-effectiveness probabilities over a wider range of thresholds. This will determine the threshold point at which shared decision-making meetings becomes cost-effective and it will be up to policy makers to decide if they are willing to invest to achieve the measured benefits of shared decision-making meetings.

Onrust et al. (2015) also carried out a cost-effectiveness analysis of FGCs based on outcomes framed around child-functioning, child-rearing and the wider environment. Rather than estimate ICERs for each outcome they present an ICER for the total areas of concern. FGCs showed a decrease in the number of areas of concern compared to the usual care group. Total costs for each arm were based on formal care received by the young person and the cost of shared decision-making meetings; costs were greater in the shared decision-making meeting group. All costs were reported in Euros but the price data are not stated. When the incremental costs were applied to the incremental benefits to generate an ICER, for the loss of each area of concern an additional investment of €2,180 is required. Onrust et al. (2015) also estimate the probability of cost-effectiveness at several thresholds; if this is set at €0, the cost-effectiveness probability is 3%, increasing to 60% if the threshold is €2,500 and to 99% if the threshold is €5,000.

The cost-effectiveness analyses suggest that at a zero cost-effectiveness threshold, FGCs have a very low probability of cost-effectiveness. A zero threshold assumes that decision makers within local authorities will not be willing to invest to achieve the desired outcomes for children in their care. However, this would not be the case and some level of investment is of course expected. Within the UK healthcare context a threshold of

£20,000 to £30,000 for each quality-adjusted life year gained is the general agreed upon threshold (McCabe et al. 2008). A similar threshold has not been defined within children's social care or for any of the outcomes described in the two cost-effectiveness analyses identified (Onrust et al. 2015; Dijkstra et al. 2018b). Decision makers are therefore required to make judgements of cost-effectiveness based on the thresholds they perceive as appropriate for each outcome.

If decision makers judge the value of a one unit reduction in an area of concern as €5000, FGCs will be considered cost-effective as estimated by Onrust et al. (2015). For Dijkstra et al. (2018b) the decision is less clear and further analysis is needed to present the value at which the service becomes cost-effective. This study shows that FGCs will require an investment greater than €10,000 and decision makers will need to make a judgement on whether this would be an appropriate threshold value.



4 DISCUSSION

4.1 Summary of main findings

This systematic review included 33 studies focusing on evaluation of shared decision-making meetings. For the seven outcome measures of interest, the evidence about whether shared decision-making meetings are more effective than control services was inconclusive. Most studies (24 out of 33) were found to have a high risk of bias and the GRADE assessments found low or very low certainty of evidence for each primary outcome.

Of the 33 studies, 17 found favourable results for the shared decision-making meetings compared to control group services for at least one outcome. However, all but one of these studies had a high risk of bias. The exception - Chambers et al. (2016) which evaluated the TDM model - had a moderate risk of bias. There was no RCT on shared decision-making meetings that identified a reduction of entry or re-entry to care, referrals or re-referrals for maltreatment, or increased satisfaction, empowerment or reunification with family when compared to control services.

Four studies found that comparison services were more effective for three outcomes; fewer children entering care (Sundell and Vinnerljung 2004; Allan et al. 2015 - Larimer and Texas; Dijkstra et al. 2016a - Part V), greater levels of reunification with family and fewer cases of re-entry to care (Perry et al. 2013) and lower referrals or re-referrals for maltreatment (Sundell and Vinnerljung 2004). All had a moderate risk of bias, except Perry et al. (2013) which had a high risk. Twenty-one individual studies did not evidence group differences for at least one outcome.

The overall evidence for the effectiveness of shared decision-making meetings is therefore weak at present, yet there is also no evidence to say such meetings increase numbers in care or compromise

child safety. There is no evidence that any one type of meeting, e.g. FGC or Team Decision Making is more effective than others, however this is difficult to make a definitive judgment on as there are only two to three examples of meeting types other than FGCs or Family Group Decision Making.

The three outcomes of care entry, care re-entry and reunification are defined in the included studies as separate outcomes, but in fact all three relate to the number of children in out-of-home care. Pooling these three outcomes relating to numbers of children in care shows that 15 studies suggest favourable outcomes for the shared decision-making meetings, 17 show no difference and five suggest more favourable outcomes for the control services. Given the limitations and quality of the studies, as well as their mixed results, it is not possible to make any strong conclusions. However, pooling the out-of-home care outcomes could be interpreted as leaning in a positive direction for the effectiveness of shared decision-making meetings in preventing out-of-home care.

The full economic evaluations suggest that shared decisions-making meetings could be cost-effective, however this depends on the type of outcome measured and the threshold used to decide on cost-effectiveness. Four out of the five partial economic evaluations show there to either be no difference or a reduction in costs with shared decision-making meetings. There was not strong evidence on the cost-effectiveness of shared decision-making meetings, however, there are encouraging indications that shared decision-making meetings could be cost-saving and potentially cost-effective.

4.2 Discussion of findings

Given the importance of family members' rights to participation in child protection decision-making,



it is disappointing to find so little strong evidence about the outcomes of shared decision-making meetings, which are probably the most widespread attempt to increase participation. Although studies of shared decision-making meetings have been carried out, few have used strong comparison groups and there is considerable variation in the outcomes from those studies that do exist. What are the implications of this lack of strong evidence?

There is a need for robust studies on the impact of shared decision-making meetings. In various forms such meetings are used extensively in children's services, not just in England but in many countries. Not only policy makers and practitioners, but also parents and children, deserve stronger evidence about the difference that these meetings can make. Informed consent to taking part in such meetings would be more meaningful if we were able to provide more evidence about the impact that they have.

It is particularly important to have studies with a good comparison group. This is important for several reasons. It is noteworthy that many of the outcomes including evaluation of empowerment and service satisfaction were relatively positive in the comparison groups of the studies we reviewed as well as the shared decision-making meeting condition. This suggests that simply investigating the experiences of the people involved in shared decision-making meetings does not provide strong evidence of the impact they have; it may be that people are often appreciative of an attempt to help them. In addition, a valid comparison group provides the strongest test of the difference that something makes. If we believe that shared decision-making meetings can lead parents, children or others to feel more involved in child protection processes, or if it is felt that they can reduce the need for children to enter public care, the strongest test of these propositions is the impact the meetings have compared to children and families who receive services as usual.

It is also worth considering why studies produced such varied findings. The findings are different from those in our Signs of Safety review (Sheehan et al. 2018) – which found very little evidence. For this literature, there were a fair number of studies, but the quantitative evaluation evidence is of

poor quality and the findings vary considerably between studies. Obviously, a recurring issue is that we do not do enough robust evaluations in children's services. However, there are other important explanations to consider. One is that it is likely that there is considerable variation in how well shared decision-making meetings, and indeed treatment as usual services, are actually delivered. Unfortunately, very few studies in this review reported on fidelity of the shared decision-making meetings to the model in question, and none reported on treatment as usual service fidelity, so it is not possible to make any empirical conclusions. Clear description of core meeting components and attendees of meetings is vitally important to include in published studies of shared decision-making meetings. Without this it is not possible to determine what sets these meetings apart from traditional case conferencing. However, it is very likely that the passion and skill with which services are provided has a substantial impact on the difference they make. In future studies, evaluation of the actual service delivery would provide valuable insight alongside the measurement of outcomes.

A related issue is the extent to which such meetings can be delivered as a standalone service. Some have argued that they are best seen as a manifestation of a restorative approach to service delivery (Sen and Webb 2019). Given that consideration of wider system changes often did not feature in the research studies, it is possible that many of the studies focused on family meetings as a discrete service without considering the wider system changes needed to make them effective. On the other hand, one of the reasons why so many studies found no impact may be that often "service as usual" incorporates key elements of collaborative family work. The social workers and other professionals in comparison groups were presumably often, or at least sometimes, struggling to involve family members as much as they can. This is an important point to highlight for future research investigations. As previously noted, details of treatment as usual services were usually not given, and a potential lack of differences between services could have led to the high rate of "no difference" results for many of the outcomes reviewed in this report. Researchers



must take care to define and report the services received by each group so that outcomes following shared decision-making meetings are not unfairly reported.

These are far from simply research issues: there are policy and practice implications from these considerations. It seems unlikely from the literature that family group involvement in decision-making is a “magic wand” that will improve either service experience or outcomes for children. Indeed, this is probably the last thing that proponents of such meetings would argue. Instead, it seems likely that inclusion of the extended family and community members involved in the child’s life in a meaningful way in decisions about their lives is not just about providing a specific type of meeting, it is fundamentally about the quality of that experience. It is therefore crucial that those delivering services focus on the quality of practice in and around such meetings. The accompanying What Works for Children’s Social Care report that examined the literature on this topic (Stabler et al., 2019) addresses some of these issues. Ensuring a consistently high quality of meetings must therefore be the primary focus for anybody involved in delivering shared decision-making meetings.

The lack of strong evidence that such meetings have consistently achieved their aims also suggests that those delivering services should carefully evaluate both the quality and the impact that family involvement in decision-making through meetings such as these is having. It seems particularly important to be clear what the intended aim of a service delivering shared decision-making meetings is. If it is to empower parents and other family members and uphold their rights, then monitoring whether these outcomes are being achieved seems essential. If the aim is to reduce the number of children in out-of-home care, then this should not be taken for granted but should be evaluated within the service. There may be disparities between the intended aims of shared decision-making meetings from the perspective of children, adult family members, professionals and researchers evaluating the service effectiveness (Holland and O’Neill 2006; Mitchell 2019) Yet, some of the papers included in this review have

been analyses of large administrative datasets (Wang et al. 2012; Lambert et al. 2017) which have investigated numbers of children entering care and family reunification following Family Group Decision Making meetings and FGCs. It is unlikely that the researchers had any communication with those who designed and facilitated the meetings, therefore, it is possible that in some cases, studies may have been evaluated against aims they were not designed to meet. In future studies, it would be useful for authors to describe their level of involvement with the service they investigated or evaluated.

In summary, while we acknowledge the difficulty of conducting comparative research on this topic, the existing quantitative evaluations are not of sufficient quality or detail to make any firm conclusions. The variation in results between studies, with almost equal numbers finding favourable outcomes for shared decision-making meetings and no differences between services, may be explained by differences in study design, meeting implementation and data analysis. In future, a collaborative multi-site RCT, involving a range of experts from social work practitioners to statisticians and experienced trial managers would produce greater certainty and confidence in the reported findings. If greater consensus could be reached about how studies are designed, and the tools used to measure outcomes (e.g. standardised empowerment and satisfaction questionnaires) it would be more feasible to compare results between studies and reach a consensus on shared decision-making meeting effectiveness.

4.3 Strengths and limitations

This systematic review has a number of strengths. It represents the most comprehensive review of evidence to date on the role of shared decision-making meetings on children’s out-of-home care, family empowerment and satisfaction. Strengths include comprehensive searches of 12 international electronic databases without any language or geographical restrictions, all available sources of grey literature, including relevant agencies in the field, reference lists of included studies, and citation tracking. In addition, we were able to contact an international panel of experts



to identify unpublished or on-going studies. These contacts resulted in three relevant unpublished reports or papers (Weisz et al. 2006; Wijnen-Lunenburg et al. 2008; Dijkstra et al. 2016a). The Cochrane and ROBINS-I risk of bias judgments provided a formal assessment of the quality of the research, and the GRADE assessments provided an objective view on the strength of the evidence, something which has not been addressed in previous literature reviews (Barnsdale and Walker 2007; Skaale Havenen and Christiansen 2014; Dijkstra et al. 2016b)

In the literature search, a considerable volume of published studies on shared decision-making meetings was noted, however the majority were not comparative (see Appendix 4). As this review aimed to evaluate the differences in outcomes between shared decision-making meetings and care as usual services, studies without a comparison group receiving different services were not eligible for inclusion.

Unfortunately, meta-analysis was not appropriate given the substantial heterogeneity of included studies in terms of the populations, services, outcomes and study designs. The classic meta-analysis aggregates effect sizes from a sample of studies to a summary effect size. All studies in a meta-analysis must use essentially the same index for all domains (service, study designs, population characteristics, well-defined outcome measures and statistical analysis), this was not the case with the 33 studies which met our eligibility criteria in terms of outcomes and comparative study design. However, Harvest plots were included to visually illustrate and summarise the included studies according to study type, direction of effect and risk of bias.

One of the key limitations of the review, stemmed from the limited and sometimes unclear reporting within included studies. It was difficult to evaluate many of the studies due to the limited information given in the publications, especially regarding the characteristics of the shared decision-making meetings and control group treatments or services. This meant that we could not compare differences between shared decision-making meeting and control services and it was often necessary to assign an "unclear" risk of bias judgment to

elements of studies as meeting characteristics and study designs were not explicitly stated (see Appendix 10a-c). One of the key elements of shared decision-making meetings is that the meetings are family-led, however the general format of the meetings was often not reported making it difficult to determine who attended and how much input professionals had. Often studies laid out the core values of these meetings but did not allude to how closely they were adhered to in the study in question. There were exceptions to this such as the YMCA Families United (2014) report and Pennell et al. (2010). As a result, it was also not possible to compare different ways in which family meetings were conducted.

Few studies reported empowerment and satisfaction data for both service and control groups. Empowerment and satisfaction were reported in six and 14 studies respectively, but only four studies reported data for both groups. From the small number of studies with comparative data it is not possible to determine whether they outperform control services. This is an important limitation of the international evidence base, given the argument that the primary rationale for shared decision-making family meetings should be family participation as a right, rather than any expected placement outcome (Morris and Connolly 2012; Connolly and Masson 2014). Empowerment and satisfaction scales are only proxies for families perceiving their rights to be upheld and other measures relating to rights need to be considered in future research.

Several studies did not employ consistent reporting, for example reporting exact numbers for one outcome but only narrative reporting and figures from which it is impossible to extract exact data (Walker 2005). Studies also differed in whether they reported numbers of families (where more than one child may be subject to CPS investigation), individual children or "cases" which complicated determining the exact sample size for each outcome, and indeed the total sample size of the review.

One aspect of this review that could be considered either a strength or a limitation is the inclusion of multiple meeting types, and studies where shared decision-making meetings were part of



a wider service or intervention. While all shared decision-making meetings attempt to engage with family members in decisions being made about the child(ren), the different circumstances under which meetings are held, and particular models used (e.g. FGC or Team Decision Making) mean that in reality the services reported in this review are not directly comparable. However, we decided to include a range of shared decision-making meeting models given that the same outcomes are assessed for various meeting types and we had the opportunity to identify whether any one meeting type appeared to be more effective than others, from which we may have been able to identify underlying reasons. However, the results were mixed and therefore inconclusive, regardless of the meeting type, which, in a way, is more informative than reaching this conclusion for a single shared decision-making model. Finally, we have attempted to acknowledge the differences between shared decision-making meetings by showing the results for each outcome disaggregated by study type where there are enough studies to do so. In doing this, there is the potential for the results to be misunderstood with such small sample sizes for some study types, especially where there may be only one study using a certain model, for example, for the outcome of reunification, only one study each investigated Team Decision Making and Family Team Conferences. The Team Decision Making study found greater reunification in the service group and the Family Team Conference found greater reunification in the control group. However, it cannot be concluded that either model is inferior, and we encourage readers to focus on the aggregated results. We are keen to stress that the results are mixed overall, with no strong evidence for or against any one model of shared decision-making meetings.

4.4 Conclusions

Involving parents, children and other family members in key decisions about their lives is inherently a good thing. The lack of strong evidence that meetings designed to do this succeed in this aim, and that they often do not reduce the need for children to be in care, should not lead us to reject such meetings. Indeed, the opposite is true: the

lack of evidence should spur us to consider how we might more effectively and consistently involve children and their family members in decisions about their lives. In doing so we need to consider not just how to provide such meetings well, but also the myriad other ways in which we could and should involve people better. Ultimately, therefore, the most important message from this review is that we have much to learn about involving families in services that statutory agencies deliver. We hope that the findings prompt innovation and adaptation, evaluation and reflection, as we seek to find the best way of involving children and their families in these important decisions about their lives.



5 AUTHOR CONTRIBUTION

UN contributed to the protocol, registered PROSPERO, conducted electronic searches, screened EndNote Libraries, critical appraisal of included papers, GRADE assessment and contributed to the manuscript writing. **CF** screened EndNote Libraries, data extraction, the ROB assessment and study categorisations, designed and created the Harvest plots and contributed to the first draft and editing of the manuscript. **ZB** screened EndNote Libraries, data extraction, the ROB and GRADE assessments, and contributed to the first draft of the manuscript. **JO** screened EndNote Libraries, data extraction, the ROB assessment, and contributed to the first draft of the manuscript. **AE** carried out the cost-analysis and contributed to the first draft of the

manuscript. **MM** led the database and literature searches, screened EndNote Libraries and contributed to the first draft of the manuscript. **AK** provided a supervisory role, undertook third reviews and contributed to manuscript writing and editing. **RT** was responsible for conception of the review, design of review methods, design of literature search strategy, and protocol writing. **RT** also provided a methodological supervisory role during the conduct of the review, acted as a third reviewer for screening studies, and contributed to manuscript writing and editing. **JS** was responsible for conception of the review, editing of the manuscript and he also provided a supervisory role. **DF** contributed to the first and final draft of the manuscript and provided a supervisory role.



6 REFERENCES

- Allan, H. et al. 2015. *No Place Like Home. Final Progress Report*. Available at: <https://library.childwelfare.gov/cwig/ws/library/docs/gateway/Blob/103415.pdf?w=+NATIVE%28%27recno%3D103415%27%29&up-p=0&rpp=10&r=1&m=1>
- Appleton, J. V. 2014. Working alongside one another... *Child Abuse Review* 24, pp. 313-316.
- Attema, A. E. et al. 2018. Discounting in economic evaluations. *Pharmacoeconomics* 36(7), pp. 745-758.
- Barnsdale, L. and Walker, M. 2007. *Examining the Use and Impact of Family Group Conferencing*. Edinburgh: Scottish Executive.
- Beehler, P. 2016. *Site Visit Report: Family Group Decision-Making: Engaging, Encouraging and Empowering Families to Succeed (FGDM-EEE)*. Available at: <https://www.childwelfare.gov/topics/management/funding/funding-sources/federal-funding/cb-funding/cbreports/fgdm/>
- Berzin, S. C. 2006. Using sibling data to understand the impact of family group decision-making on child welfare outcomes. Fresno study. *Children and Youth Services Review* 28(12), pp. 1449-1458.
- Berzin, S. C. et al. 2008. Does family group decision making affect child welfare outcomes? Findings from a randomized control study. *Child welfare* 87(4),
- Brand, S. L. et al. 2018. Mapping the evidence about what works to safely reduce the number of children and young people in statutory care: a systematic scoping review. London:
- Chambers, R. M. et al. 2016. An innovative child welfare pilot initiative: Results and outcomes. *Children and Youth Services Review* 70, pp. 143-151.
- Connolly, M. 2006. Up front and personal: Confronting dynamics in the family group conference. *Family process* 45, pp. 345-357.
- Connolly, M. and Masson, J. 2014. Private and public voices: Does family group conferencing privilege the voice of children and families in child welfare? *Journal of Social Welfare and Family Law* 36(4), pp. 403-414.
- Corby, B. et al. 1996. Parental participation in child protection work: Rethinking the rhetoric. *The British Journal of Social Work* 26, pp. 475-492.
- Craig, P. et al. 2012. *Using natural experiments to evaluate population health interventions: guidance for producers and users of evidence*. Available at: <https://mrc.ukri.org/documents/pdf/natural-experiments-guidance/> [Accessed: 06/09/2019].
- Craig, P. et al. 2017. Natural experiments: an overview of methods, approaches, and contributions to public health intervention research. *Annual review of public health* 38, pp. 39-56.
- Crampton, D. S. and Jackson, W. L. 2007. Family group decision making and disproportionality in foster care: a case study. *Child welfare* 86(3),
- Crampton, D. S. et al. 2011. Does community and family engagement enhance permanency for children in foster care? Findings from an evaluation of the family-to-family initiative. *Child welfare* 90(4),
- Crow, G. and Marsh, P. 1998. *Family group conferences in child welfare*. Blackwell Science.



Darlington, Y. et al. 2012. Parents' perceptions of their participation in mandated family group meetings. *Children and Youth Services Review* 34(2), pp. 331-337.

Dijkstra, S. et al. 2018a. A Randomized Controlled Trial on the Effectiveness of Family Group Conferencing in Child Welfare: Effectiveness, Moderators, and Level of FGC Completion. *Child maltreatment*,

Dijkstra, S. et al. 2016a. De inzet van familienetwerkberaden in de jeugdbescherming.,

Dijkstra, S. et al. 2016b. The effectiveness of family group conferencing in youth care: A meta-analysis. *Child abuse & neglect* 62, pp. 100-110.

Dijkstra, S. et al. 2018b. Cost-effectiveness of Family Group Conferencing in child welfare: a controlled study. *BMC public health* 18(1), p. 848.

Faller, K. C. 1981. Social work with abused and neglected children: A manual of interdisciplinary practice. . Simon and Schuster.

Featherstone, B. et al. 2018. Let's stop feeding the risk monster: towards a social model of 'child protection'. *Families, Relationships and Societies* 7(1), pp. 7-22.

Feldman, L. H. 2017. Using Family Group Decision Making to Assist Informal Kinship Families. *Child welfare* 95(4), pp. 41-67.

Godinet, M. T. et al. 2010. Disproportionality, Ohana conferencing, and the Hawai'i child welfare system. *Journal of Public Child Welfare* 4(4), pp. 387-405.

GRADE Working Group. 2004. *Grading quality of evidence and strength of recommendations*. . Available at: [Accessed: 20/05/2019].

Higgins, J. and Green, S. 2011. *Cochrane Handbook for SRs of interventions 5.1.0 (updated March 2011)*. . Available at: [Accessed: 20/05/2019].

Holland, S. and O'Neill, S. 2006. 'We Had

to be There to Make Sure it was What We Wanted' Enabling children's participation in family decision-making through the family group conference. *Childhood* 13(1), pp. 91-111.

Hollinshead, D. M. et al. 2017. Effectiveness of family group conferencing in preventing repeat referrals to child protective services and out-of-home placements. *Child abuse & neglect* 69, pp. 285-294.

Huebner, R. A. et al. 2012. The impact of Sobriety Treatment and Recovery Teams (START) on family outcomes. *Families in Society* 93(3), pp. 196-203.

Husereau, D. et al. 2013. Consolidated health economic evaluation reporting standards (CHEERS) statement. *Cost Effectiveness and Resource Allocation* 11(1),

Lambert, M. C. et al. 2017. The impact of family group decision-making on preventing removals. *Children and Youth Services Review* 78, pp. 89-92.

Leatherdale, S. T. 2019. Natural experiment methodology for research: a review of how different methods can support real-world research. *International Journal of Social Research Methodology* 22(1), pp. 19-35.

Lupton, C. and Nixon, P. 1999. Empowering practice?: a critical appraisal of the family group conference approach. Policy Press.

Mason, P. et al. 2017. Leeds Family Valued. Evaluation report, July 2017.

McCabe, C. et al. 2008. The NICE cost-effectiveness threshold. *Pharmacoeconomics* 26(9), pp. 733-744.

Mitchell, M. 2019. Reimagining child welfare outcomes: Learning from Family group conferencing. *Child & Family Social Work*,

Morris, K. 2007. An evaluation of outcomes: Camden family group conference service. Camden, UK:

Morris, K. and Connolly, M. 2012. Family decision making in child welfare: Challenges in developing a knowledge base for practice. *Child Abuse Review* 21(1), pp. 41-52.



Morris, K. and Tunnard, J. 1996. Family Group Conferences: Messages from UK practice and research. Family Rights Group London.

Muench, K. et al. 2017. Children and parent participation in Child Protection Conferences: A study in one English Local Authority. *Child Care in Practice* 23(1), pp. 49-63.

Munro, E. R. et al. 2017. Daybreak Family Group Conferencing: children on the edge of care: July 2017.

Ogilvie, D. et al. 2008. The harvest plot: A method for synthesising evidence about the differential effects of interventions. *BMC Medical Research Methodology* 8(1), p. 8.

Onrust, S. A. et al. 2015. Family group conferences within the integrated care system for young people with ID: a controlled study of effects and costs. *BMC health services research* 15(1), p. 392.

Partnership for Strong Families. 2012. Family Connection Grant: Family Empowerment through Family Team Conferencing. Final Report. . Gainesville, FL:

Pennell, J. and Burford, G. 1995. Family group decision making: New roles for "old" partners in resolving family violence: Implementation Report. . Memorial University of Newfoundland, School of Social Work:

Pennell, J. and Burford, G. 2000. Family group decision making: protecting children and women. *Child welfare* 79(2),

Pennell, J. et al. 2010. Expedited family group engagement and child permanency. *Children and Youth Services Review* 32(7), pp. 1012-1019.

Perry, R. et al. 2013. Family team conferencing: Results and implications from an experimental study in Florida. *Child welfare* 92(6), p. 63.

Petticrew, M. and Roberts, H. 2003. Evidence, hierarchies, and typologies: horses for courses. *Journal of Epidemiology & Community Health* 57(7), pp. 527-529.

Sen, R. and Webb, C. 2019. Exploring the declining rates of state social work intervention in an English local authority using Family Group Conferences. *Children and Youth Services Review* 106, p. 104458.

Sheehan, L. et al. 2018. Signs of Safety: Findings from a Mixed-methods Systematic Review Focussed on Reducing the Need for Children to Be in Care. What Works Centre for Children's Social Care Cardiff.

Sheets, J. et al. 2009. Evidence-based practice in family group decision-making for Anglo, African American and Hispanic families. *Children and Youth Services Review* 31(11), pp. 1187-1191.

Skaale Havenen, K. J. and Christiansen, O. 2014. Knowledge Review on Family Group Conferencing. Experiences and Outcomes.

Stabler, L. et al. 2019. How can shared decision-making meetings between professionals and family members safely reduce the number of children in care? A rapid realist review. London:

Sterne, J. A. et al. 2016. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 355,

Sundell, K. and Vinnerljung, B. 2004. Outcomes of family group conferencing in Sweden: A 3-year follow-up. *Child abuse & neglect* 28(3), pp. 267-287.

Teal, B. L. 2013. The effectiveness of team decision making in Children's Division on child welfare outcomes. Capella University.

Thomas, C. 2018. The Care Crisis Review: Factors contributing to national increases in numbers of looked after children and applications for care orders. London: Family Rights Group,

Titcomb, A. and LeCroy, C. 2003. Evaluation of Arizona's family group decision making program. *Protecting Children* 18, pp. 58-64.

Walker, L. 2005. A Cohort Study of 'Ohana Conferencing in Child Abuse and Neglect Cases. *Protecting Children* 19(4), pp. 36-46.

Wang, E. W. et al. 2012. Expediting permanent placement from foster care



systems: The role of family group decision-making. *Children and Youth Services Review* 34(4), pp. 845-850.

Weisz, V. et al. 2006. Nebraska family group conferencing: Evaluation report. Lincoln, NE: University of Nebraska, UN-L Center on Children, Families, and the Law, Nebraska Court Improvement Project,

Wheeler, C. and Johnson, S. 2003. Evaluating family group decision making: The Santa Clara example. *Protecting Children* 18(1-2), pp. 65-69.

Wijnen-Lunenburg, P. et al. 2008. De familie aan zet. De uitkomsten van Eigen Krachtconferenties in de jeugdbescherming met betrekking tot veiligheid, sociale cohesie en regie. . Voorhout, the Netherlands:

YMCA Families United. 2014. YMCA of San Diego County Families United Family Group Conferencing: Project Summary and Findings. US:

York-Health-Economics-Consortium. 2016. *Preference-Based Measures*. yhec.co.uk: Available at: <https://www.yhec.co.uk/glossary/preference-based-measures/> [Accessed: 29/10/2019].



7 APPENDICES

7.1 Appendix 1. List of databases and agencies searched

Published works were searched for in the following databases:

- Child Development and Adolescent Studies
- EconLit
- EMBASE
- Health Management Information Consortium (HMIC)
- IDEAS/RePEc (Research papers in Economics) search
- MEDLINE (National Library of Medicine's bibliographic database)
- NHS Economic Evaluation Database (NHS EED)
- PsycINFO
- Scopus Elsevier
- Sociological abstracts PROQUEST
- Social Policy and Practice

Grey literature was searched for via the following relevant agencies:

- Center for Family and Community Involvement
- Child Welfare Information Gateway
- Department for Education Innovations Programme
- Family Rights Group (UK)
- National Institute for Permanent Family Connectedness
- Open Grey
- Research in Practice
- The American Humane Organisation (USA)



7.2 Appendix 2. Scopus search strategy

The Scopus search strategy was adopted for all other search databases.

((TITLE-ABS ("California Title IV-E Waiver Demonstration project" OR "Sobriety Treatment and Recovery Teams" OR "Transitioning youth to families intervention" OR "case conferencing" OR "family case conferencing" OR "participative case conferencing")) OR (TITLE-ABS (("Family first") W/4 (program* OR service* OR model* OR initiative* OR project OR projects))) OR (TITLE-ABS ((daybreak) W/4 (program* OR service* OR model* OR initiative* OR project OR projects))) OR (TITLE-ABS (("family centred") W/4 (program* OR service* OR model* OR initiative* OR project OR projects))) OR (TITLE-ABS ((client OR family OR families OR consumer*) W/2 (engag* OR participat* OR involv* OR collabor* OR partnership*) AND (decision* OR planning OR meeting* OR plans))) OR (TITLE-ABS ((client OR family OR families OR consumer*) W/3 (decision* OR planning OR plans))) OR (TITLE-ABS ("Family group" W/5 (decision OR engag* OR meeting* OR involve* OR conferenc* OR service*))) OR (TITLE-ABS ("Family team" W/5 (decision OR engag* OR meeting* OR involve* OR conferenc* OR service*))) OR (TITLE-ABS ((family) W/2 (meeting* OR model* OR program* OR conferenc*))) OR (TITLE-ABS (("Family valued") W/4 (meeting* OR model* OR program* OR conferenc* OR service* OR initiative*))) OR (TITLE-ABS (("Family unity" OR "family centred") W/4 (meeting* OR model* OR program* OR conferenc* OR service* OR initiative*))) OR (TITLE-ABS ((fgdm OR fgc) W/3 (meeting* OR model* OR program* OR conferenc* OR service* OR initiative* OR plans))) OR (TITLE-ABS (("private family") W/2 (meeting* OR conferenc* OR decision* OR time)))) AND ((TITLE-ABS (("Client outcomes" OR placement OR placements OR "state custody" OR "more empower*" OR "out of home care"))) OR (TITLE-ABS ((removal* OR removing OR remove*) W/8 (home OR family OR families))) OR (TITLE-ABS ((enter* OR entry) W/3 (care OR "child protection" OR "social services")))) OR (TITLE-ABS ((permanence OR permanency OR

"substantiated problems" OR "re-referral*" OR "re-entry" OR "supervision order" OR reunified OR reunification OR "family stabili*"))) OR (TITLE-ABS ((transition*) W/5 (care OR "social service*" OR "welfare service*")))) OR (TITLE-ABS ((number) W/3 (professional OR service* OR agenc*)))) OR (TITLE-ABS ((case*) W/2 (clos*)))) OR (TITLE-ABS (exit* W/2 (care OR expedit*)))) OR (TITLE-ABS (child* W/3 (maltreatment OR neglect OR abuse OR safety OR welfare OR wellbeing)))) OR (TITLE-ABS ((satisf*) W/3 (famil* OR parent* OR mother* OR father* OR consumer* OR client*)))) OR (TITLE-ABS ((empower*) W/5 (famil* OR parent* OR mother* OR father* OR feeling* OR perspect* OR perceiv*)))) OR (TITLE-ABS ((respect OR dignity OR dignified) W/3 (famil* OR parent* OR mother* OR father* OR outcome OR experienc* OR satisf* OR perspect* OR perceiv*)))) OR (TITLE-ABS ((famil* OR parent* OR relatives) W/2 (functioning OR perspectives OR experienc* OR nurturing OR attachment OR knowledge)))) OR (TITLE-ABS ((respectful OR digni*) W/3 (program* OR service* OR model* OR initiative* OR project OR projects)))) OR (TITLE-ABS ((respect OR respected OR digni*) W/3 (outcome OR outcomes OR increased OR improved OR finding OR findings OR results)))) OR (TITLE-ABS ("economic evaluation" OR "cost effectiveness" OR "cost-utility" OR "cost benefit" OR "cost analysis" OR "cost measure*"))) OR (TITLE-ABS ((cost OR costs OR costing OR economic) W/2 (apprais* OR assess* OR analysis* OR analyses* OR study OR evaluat* OR estimat* OR decision OR burden OR expenditure)))) OR (TITLE-ABS ((cost OR costs OR costing OR economic OR value) W/2 (decision* OR threshold)))) OR (TITLE-ABS (value W/2 money))) OR (TITLE-ABS (model* W/2 (economic OR decision OR decisionmaking)))) OR (TITLE-ABS (costbenefit* OR costeffect* OR "return on investment"))) OR (TITLE-ABS ((costs OR cost) W/2 (effect* OR utility OR benefit))))) AND ((TITLE-ABS-KEY (("social care" OR "social work" OR "child protection" OR "welfare service*" OR "social service*" OR "social worker*" OR "welfare system" OR "child welfare" OR "care system"



OR "foster care" OR "child protective service*" OR "youth service*")) OR (SRCTITLE ("social care" OR "social work" OR "child protection" OR "welfare service*" OR "social service*" OR "social worker*" OR "welfare system" OR "child welfare" OR "care system" OR "foster care" OR "child protective service*" OR "youth service*")) AND (EXCLUDE (PUBYEAR, 1988) OR EXCLUDE (PUBYEAR, 1987) OR EXCLUDE (PUBYEAR, 1986) OR EXCLUDE (PUBYEAR, 1985) OR EXCLUDE (PUBYEAR, 1984) OR EXCLUDE (PUBYEAR, 1983) OR EXCLUDE (PUBYEAR, 1982) OR EXCLUDE (PUBYEAR, 1981) OR EXCLUDE (PUBYEAR, 1980) OR EXCLUDE (PUBYEAR, 1979) OR EXCLUDE (PUBYEAR, 1978) OR EXCLUDE (PUBYEAR, 1977) OR EXCLUDE (PUBYEAR, 1976) OR EXCLUDE (PUBYEAR, 1975) OR EXCLUDE (PUBYEAR, 1974) OR EXCLUDE (PUBYEAR, 1973) OR EXCLUDE (PUBYEAR, 1972) OR EXCLUDE (PUBYEAR, 1971) OR EXCLUDE (PUBYEAR, 1969) OR EXCLUDE (PUBYEAR, 1968) OR EXCLUDE (PUBYEAR, 1967) OR EXCLUDE (PUBYEAR, 1964) OR EXCLUDE (PUBYEAR, 1952))



7.3 Appendix 3. List of experts contacted

Name	Country
Mr Abyd Quinn-Aziz	UK
Dr Stephanie Berzin	USA
Professor Gale Burford	USA
Professor Edward Cohen	USA
Dr Hanneke Creemers	Netherlands
Dr Sharon Dijkstra	Netherlands
Dr John Fluke	USA
Professor Anna Gupta	UK
Dr Dana Hollinshead	USA
Dr Ruth Huebner	USA
Dr Matthew Lambert	USA
Dr Lisa Merkel-Holguin	USA
Dr Mary Mitchell	UK
Professor Kate Morris	UK
Dr Simone Onrust	Netherlands
Professor Joan Pennell	USA
Dr Eugene Wang	USA
Professor Victoria Weisz	US



7.4 Appendix 4a. Description of excluded papers

References	Reason for exclusion
Ahn, H., Hartzel, S., & Shaw, T. (2018). Participants' satisfaction with family involvement meetings: Implications for child welfare practice. <i>Research on Social Work Practice, 28</i> (8), 952-963.	Uncontrolled study
Akakpo, T. F. (2008). Staff attitudes and beliefs about family involvement of delinquent children in residential programs. Michigan State University. School of Social Work.	Uncontrolled study
Akin, B. A., Brook, J., Lloyd, M. H., & McDonald, T. P. (2017). Effect of a parenting intervention on foster care reentry after reunification among substance-affected families: A quasi-experimental study. <i>Child maltreatment, 22</i> (3), 194-204.	Not a family group meeting service
Allan, H., Harlaar, N., Hollinshead, D., Drury, I., & Merkel-Holguin, L. (2017). The impact of worker and agency characteristics on FGC referrals in child welfare. <i>Children and Youth Services Review, 81</i> , 229-237.	Uncontrolled study
Asscher, J. J., Dijkstra, S., Stams, G. J. J., Deković, M., & Creemers, H. E. (2014). Family group conferencing in youth care: characteristics of the decision making model, implementation and effectiveness of the Family Group (FG) plans. <i>BMC public health, 14</i> (1), 154.	Trial protocol
Ban, P. (1994). Preliminary findings on family decision making project in the Victorian child protection system. <i>Australian Social Work, 47</i> (1), 34-36.	Descriptive
Bell, M. (1996). An account of the experiences of 51 families involved in an initial child protection conference. <i>Child & Family Social Work, 1</i> (1), 43-55.	Not a family group meeting service
Bell, M., & Wilson, K. (2006). Children's views of family group conferences. <i>British Journal of Social Work, 36</i> (4), 671-681.	Uncontrolled study
Berzin, S. C., Thomas, K. L., & Cohen, E. (2007). Assessing model fidelity in two family group decision-making programs: Is this child welfare intervention being implemented as intended?. <i>Journal of Social Service Research, 34</i> (2), 55-71.	Uncontrolled study
Burford, G., Pennell, J., Macleod, S., Campbell, S., & Lyall, G. (1996). Reunification as an extended family matter. <i>Community Alternatives, 8</i> , 33-55.	Uncontrolled study.
Burke, T. K., Allen-Eckard, K., Kemp, S. P., Ware, J., Ackroyd, A., & Munoz, S. (2003). Community family support meetings: Adding community resources to family decision making. <i>Protecting Children, 18</i> , 104.	Uncontrolled study
Carson, G. (2010). It's a family affair. <i>Community Care, 22</i> .	Descriptive
Chandler, S. M. (2013). The application of collaboration models to family group conferencing. <i>Journal of Policy Practice, 12</i> (1), 3-22.	Descriptive
Cleek, E. N., Wofsy, M., Boyd-Franklin, N., Mundy, B., & Howell, T. J. (2012). The family empowerment program: An interdisciplinary approach to working with multi-stressed urban families. <i>Family process, 51</i> (2), 207-217.	Descriptive
Connolly, M. (2006). Up front and personal: Confronting dynamics in the family group conference. <i>Family process, 45</i> (3), 345-357.	Qualitative study
Cooper, S. (2007). Ways to keep children from going into care. <i>Children Now</i> .	Descriptive



References	Reason for exclusion
Crampton D. (2003) Family Group Decision Making in Kent County, Michigan: The Family and Community Compact. <i>Protecting Children</i> , 18, 81-83.	Results not reported
Crampton, D. (2007). Research Review: Family group decision-making: A promising practice in need of more programme theory and research. <i>Child & Family Social Work</i> , 12(20), 202-209.	Literature review
Crampton, D. S. (2004). Family Involvement Interventions in Child Protections: Learning from Contextual Integrated Strategies. <i>J. Soc. & Soc. Welfare</i> , 31, 175.	Descriptive
Crea, T. M., Crampton, D. S., Abramson-Madden, A., & Usher, C. L. (2008). Variability in the implementation of Team Decisionmaking (TDM): Scope and compliance with the Family to Family practice model. <i>Children and Youth Services Review</i> , 30(11), 1221-1232.	Uncontrolled study
Crea, T. M., Usher, C. L., & Wildfire, J. B. (2009). Implementation fidelity of team decisionmaking. <i>Children and Youth Services Review</i> , 31(1), 119-124.	Uncontrolled study
Devaney, C., & Byrne, P. (2015). The value of family welfare conferencing within the child protection and welfare system. <i>Child care in practice</i> , 21(4), 340-356.	Qualitative study
Drywater-Whitekiller, V. (2014). Family group conferencing: An indigenous practice approach to compliance with the Indian Child Welfare Act. <i>Journal of Public Child Welfare</i> , 8(3), 260-278.	Descriptive study
Edwards, D., Parkinson, K., Fisher, T., & Owen, J. (2019). Looked after children as decision makers: family group conferences in practice. <i>Child Care in Practice</i> , 1-13.	Case studies
English, D. J., Brummel, S., & Martens, P. (2009). Fatherhood in the child welfare system: Evaluation of a pilot project to improve father involvement. <i>Journal of Public Child Welfare</i> , 3 (3), 213-234.	Uncontrolled study
Gallagher, M., Smith, M., Wilkinson, H., Cree, V., Wosu, H., Stewart, J., & Hunter, S. (2011). Engaging with families in child protection.	Qualitative study
Garcia, J. A. (2003). Transforming Relationships in Practice and Research: What is the Stanislaus Model? <i>Protecting Children</i> , 18, 22.	Discussion of the evaluation plan
Greeno, E. J., Murray, K., & Rushovich, B. (2013). Using multi-informed fidelity data to determine the impact of a neutral child welfare facilitator for permanency decision teams. <i>Child welfare</i> , 92(6).	Uncontrolled study
Gunderson, K, Cahn, K, & Wirth, J. (2003). The Washington State long-term outcome study. <i>Protecting Children</i> (18), 42-47.	Uncontrolled study
Gunderson, K., Cahn, K. & Wirth, J. (2003). The Washington State Long-Term Outcome Study. <i>Protecting Children</i> , 18, 42-7.	Uncontrolled study
Hall, J., Pennell, J. & Rikard, R. V. (2015). Child and Family Team Meetings. In Gal, T. & Duramy, B. (Eds). <i>International Perspectives and Empirical Findings on Child Participation: From Social Exclusion to Child-Inclusive Policies</i> (p. 207-226). Oxford University Press. Available at: https://www.researchgate.net/profile/RV_Rikard/publication/280567671_Child_and_Family_Team_Meetings_The_Need_for_Youth_Participation_in_Educational_Success/links/55ccc2ea08aef0f107835e0e/Child-and-Family-Team-Meetings-The-Need-for-Youth-Participation-in-Educational-Success.pdf	Qualitative study



References	Reason for exclusion
Hess, P. M., McGowan, B. G., & Botsko, M. (2000). A preventive services program model for preserving and supporting families over time. <i>Child Welfare, 79</i> (3).	Not a family group meeting service
Horwitz, M. (2003). Family Group Decision Making and Permanency Planning. <i>Protecting Children, 18</i> , 93.	Descriptive
Huebner, R. A., Posze, L., Willauer, T. M., & Hall, M. T. (2015). Sobriety treatment and recovery teams: implementation fidelity and related outcomes. <i>Substance use & misuse, 50</i> (10), 1341-1350.	Uncontrolled study
Huebner, R. A., Robertson, L., Roberts, C., Brock, A., & Geremia, V. (2012). Family preservation: Cost avoidance and child and family service review outcomes. <i>Journal of Public Child Welfare, 6</i> (2), 206-224.	Not a family group meeting service
Huebner, R. A., Willauer, T., Posze, L., Hall, M. T., & Oliver, J. (2015). Application of the evaluation framework for program improvement of START. <i>Journal of Public Child Welfare, 9</i> (1), 42-64.	Uncontrolled study
Jones, L. P., & Finnegan, D. (2004). Family unity meetings: Decision making and placement outcomes. <i>Journal of Family Social Work, 7</i> (4), 23-43.	Uncontrolled study
Kim, J., Imburgia, T. M., Armstrong-Richardson, E., Jagers, J. W., & Hall, J. A. (2018). Effects of case characteristics on teamwork in family meetings. <i>Child & Family Social Work.</i>	Uncontrolled study
Kim, J., Trahan, M., Bellamy, J., & Hall, J. A. (2019). Advancing the innovation of family meeting models: The role of teamwork and parent engagement in improving permanency. <i>Children and Youth Services Review, 100</i> , 147-155.	Uncontrolled study
Kinchin, I., Doran, C. M., McCalman, J., Jacups, S., Tsey, K., Lines, K., ... & Searles, A. (2017). Delivering an empowerment intervention to a remote Indigenous child safety workforce: Its economic cost from an agency perspective. <i>Evaluation and program planning, 64</i> , 85-89.	Not a family group meeting service
Kumpfer, K. L., & Magalhães, C. (2018). Strengthening Families Program: An evidence-based family intervention for parents of high-risk children and adolescents. <i>Journal of Child & Adolescent Substance Abuse, 27</i> (3), 174-179.	Review
Lewandowski, C. A., & Pierce, L. (2002). Assessing the effect of family-centered out-of-home care on reunification outcomes. <i>Research on Social Work Practice, 12</i> (2), 205-221.	Not a family group meeting service
Lewandowski, C. A., & Pierce, L. (2004). Does family-centered out-of-home care work? Comparison of a family-centered approach and traditional care. <i>Social Work Research, 28</i> (3), 143-151.	Not a family group meeting service
Litchfield, M., Gatowski, S., & Dobbin, S. (2003). Improving outcomes for families: Results from an evaluation of Miami's Family Decision Making Program. <i>Protecting Children (18)</i> , 48-51.	Uncontrolled study
Lorentzen, B. L. (2008). Effects of family group decision making in a voluntary family maintenance program. (Doctoral dissertation 3331701, University of California.	Does not include outcomes of interest
Lupton, C., & Stevens, M. (1997). Family outcomes: Following through on family group conferences. Social Services Research and Information Unit.	Uncontrolled study



References	Reason for exclusion
Marcynyszyn, L. A., Bear, P. S., Geary, E., Conti, R., Pecora, P. J., Day, P. A., & Wilson, S. T. (2012). Family Group Decision Making (FGDM) with Lakota families in two tribal communities: tools to facilitate FGDM implementation and evaluation. <i>Child welfare, 91</i> (3).	Qualitative study
Marsh, P., & Crow, G. (2003). Family group conferences and child protection in a multicultural community-1998. <i>Protecting Children, 18</i> (1-2), 131-33.	Uncontrolled study and qualitative
McCrae, J. S., & Fusco, R. A. (2010). A racial comparison of Family Group Decision Making in the USA 1. <i>Child & Family Social Work, 15</i> (1), 41-55.	Uncontrolled study
Moffett, J. (2015). Site Visit Report: Homes for Black Children: Nurturing the Resiliency in Wayne County Families: Rethinking the Family Decision-Making Model as Community-Centered Child and Family Work.	No control data for two outcomes of interest and unable to determine which condition participants belonged to
Nyberg, Eva (2003). Family Group Conferencing in Sweden. <i>Protecting Children</i> .	Descriptive
Papworth, A. (2010). Power to the family. <i>Professional Social Work, 20</i> -21. Available at: https://www.basw.co.uk/system/files/resources/basw_114838-10_0.pdf (Accessed 11/06/2019)	Descriptive
Patterson-Trimble-Johnson, K. (2011). A study of a family preservation/ family support program in Region 12 of Coastal Southeast Georgia (Doctoral dissertation, Fielding Graduate University).	Not a family group meeting service
Pennell, J. & Anderson, G. (2007). Widening the Circle: The Practice and Evaluation of Family Group Conferencing with Children, Youths, and Their Families. Roslyn Heights: Libra Publishers Incorporated.	Descriptive
Pennell, J. (2003). North Carolina family group conferencing project: research summary. <i>Protecting Children, 18</i> (1), 70-3.	Not empirical research
Pennell, J. (2006). Restorative practices and child welfare: Toward an inclusive civil society. <i>Journal of social issues, 62</i> (2), 259-279.	Uncontrolled study
Pennell, J. (2006). Restorative practices and child welfare: Toward an inclusive civil society. <i>Journal of social issues, 62</i> (2), 259-279.	Uncontrolled study
Pennell, J., & Burford, G. (2000). Family group decision making: protecting children and women. <i>Child welfare, 79</i> (2).	Not empirical research
Quinnett, E., Harrison, R. S., & Jones, L. (2003). Empirical Research on the San Diego Model of Family Unity Meetings. <i>Protecting Children: Promising Results, Potential New Directions: International FGDM Research and Evaluation in Child Welfare, 18</i> , 98-103.	Uncontrolled study
Rauktis, M. E., Huefner, J., & Cahalane, H. (2011). Perceptions of fidelity to family group decision-making principles: examining the impact of race, gender, and relationship. <i>Child Welfare, 90</i> (4).	Uncontrolled study
Robinson, J. S. D., Litchfield, M., Gatowski, S., & Dobbin, S. (2002). Family conferencing: A success for our children. <i>Juvenile and family court journal, 53</i> (4), 43-48.	Uncontrolled study 10.4 Appendix 4a. Description of excluded papers
Sandau-Beckler, P., Reza, S., & Terrazas, A. (2005). Familias Primero: Family Group Decision Making in El Paso County, Texas. <i>Protecting Children, 19</i> (4), 54-62.	Uncontrolled study



References	Reason for exclusion
Swain, D. (1995). Family group conferences in child care and protection and in youth justice in Aotearoa/New Zealand. <i>International Journal of Law, Policy and the Family</i> , 9(2), 155-207.	Not empirical research
Thoennes, N. (2003). Family group decision making in Colorado. <i>Protecting Children</i> , 18(1-2), 74-80.	Not a controlled study
Thomas, K. L., Cogen, E. & Duerr Berrick, J. (2003). California's Waiver Evaluation of FGDM: A unique opportunity. <i>Protecting Children</i> , 18, 52- 56.	Results not reported
Thomas, K. L., Berzin, S. C., & Cohen, E. (2005). Fidelity of family group decision making: A content analysis of family conference and case plans in a randomized treatment study. <i>Protecting Children</i> , 19(4), 4-15.	Does not include outcomes of interest
Treichel, C. J., & Bemis, A. (2003). Minnesota s Evaluation Study of Family Group Decision Making. <i>HHS</i> .	Uncontrolled study
Tyuse, S. W., Hong, P. P., & Stretch, J. J. (2010). Evaluation of an intensive in-home family treatment program to prevent out-of-home placement. <i>Journal of evidence-based social work</i> , 7(3), 200-218.	Not a family group meeting service
Vuorenmaa, M., Halme, N., Perälä, M. L., Kaunonen, M., & Åstedt-Kurki, P. (2016). Perceived influence, decision-making and access to information in family services as factors of parental empowerment: a cross-sectional study of parents with young children. <i>Scandinavian journal of caring sciences</i> , 30(2), 290-302.	Uncontrolled study
Walton, E., Roby, J., Frandsen, A., & Davidson, R. (2003). Strengthening at-risk families by involving the extended family. <i>Journal of Family Social Work</i> , 7(4), 1-21.	Uncontrolled study
Weigensberg, E. C., Barth, R. P., & Guo, S. (2009). Family group decision making: A propensity score analysis to evaluate child and family services at baseline and after 36-months. <i>Children and Youth Services Review</i> , 31(3), 383-390.	Does not include outcomes of interest
Williams, Angela (2003). Preliminary Results from the Nebraska FGC Evaluation. <i>Protecting Children</i> .	Descriptive study
Xu, Y., Ahn, H., & Bright, C. L. (2017). Family involvement meetings: Engagement, facilitation, and child and family goals. <i>Children and Youth Services Review</i> , 79, 37-43.	Uncontrolled study
Huebner, R. A., Willauer, T., & Posze, L. (2012). The impact of Sobriety Treatment and Recovery Teams (START) on family outcomes. <i>Families in Society</i> , 93(3), 196-203.	Part of another intervention and therefore unable to isolate effects due to meeting alone



7.5 Appendix 5. Included papers

Study ID	Peer-Reviewed Publications
#4	Berzin, S. C. (2006). Using sibling data to understand the impact of family group decision-making on child welfare outcomes. <i>Children and Youth Services Review</i> , 28(12), 1449-1458.
#5	Berzin, S. C., Cohen, E., Thomas, K., & Dawson, W. C. (2008). Does family group decision making affect child welfare outcomes? Findings from a randomized control study. <i>Child welfare</i> , 87(4).
#6	Chambers, R. M., Brocato, J., Fatemi, M., & Rodriguez, A. Y. (2016). An innovative child welfare pilot initiative: Results and outcomes. <i>Children and Youth Services Review</i> , 70, 143-151.
#7	Crampton, D., & Jackson, W. L. (2007). Family group decision making and disproportionality in foster care: a case study. <i>Child Welfare</i> , 86(3).
#8	Crampton, D. S., Usher, C. L., Wildire, J. B., Webster, D., & Cuccaro-Alamin, S. (2011). Does community and family engagement enhance permanency for children in foster care? Findings from an evaluation of the family-to-family initiative. <i>Child welfare</i> , 90(4).
#9	Dijkstra, S., Asscher, J. J., Deković, M., Stams, G. J. J., & Creemers, H. E. (2018). A Randomized Controlled Trial on the Effectiveness of Family Group Conferencing in Child Welfare: Effectiveness, Moderators, and Level of FGC Completion. <i>Child maltreatment</i> , 1077559518808221.
#10	Dijkstra, S., Creemers, H. E., Van Steensel, F. J., Deković, M., Stams, G. J. J., & Asscher, J. J. (2018). Cost-effectiveness of Family Group Conferencing in child welfare: a controlled study. <i>BMC public health</i> , 18(1), 848.
#13	Feldman, L. H. (2017). Using Family Group Decision Making to Assist Informal Kinship Families. <i>Child Welfare</i> , 95(4), 41-67.
#14	Godinet, M. T., Arnsberger, P., Li, F., & Kreif, T. (2010). Disproportionality, Ohana conferencing, and the Hawai'i child welfare system. <i>Journal of Public Child Welfare</i> , 4(4), 387-405.
#15	Hollinshead, D. M., Corwin, T. W., Maher, E. J., Merkel-Holguin, L., Allan, H., & Fluke, J. D. (2017). Effectiveness of family group conferencing in preventing repeat referrals to child protective services and out-of-home placements. <i>Child abuse & neglect</i> , 69, 285-294.
#17	Lambert, M. C., Johnson, L. E., & Wang, E. W. (2017). The impact of family group decision-making on preventing removals. <i>Children and Youth Services Review</i> , 78(C), 89-92
#20	Onrust, S. A., Romijn, G., & de Beer, Y. (2015). Family group conferences within the integrated care system for young people with ID: a controlled study of effects and costs. <i>BMC health services research</i> , 15(1), 392.
#21	Pennell, J., & Burford, G. (2000). Family group decision making: protecting children and women. <i>Child welfare</i> , 79(2).
#22	Pennell, J., Edwards, M., & Burford, G. (2010). Expedited family group engagement and child permanency. <i>Children and Youth Services Review</i> , 32(7), 1012-1019.
#23	Perry, R., Yoo, J., Spoliansky, T., & Edelman, P. (2013). Family team conferencing: Results and implications from an experimental study in Florida. <i>Child welfare</i> , 92(6), 63.**



Study ID	Peer-Reviewed Publications
#24	Sheets, J., Wittenstrom, K., Fong, R., James, J., Tecci, M., Baumann, D. J., & Rodriguez, C. (2009). Evidence-based practice in family group decision-making for Anglo, African American and Hispanic families. <i>Children and Youth Services Review</i> , 31(11), 1187-1191
#25	Sundell, K., & Vinnerljung, B. (2004). Outcomes of family group conferencing in Sweden: A 3-year follow-up. <i>Child abuse & neglect</i> , 28(3), 267-287.
#27	Titcomb, A., & LeCroy, C. (2003). Evaluation of Arizona's family group decision making program. <i>Protecting Children</i> , 18, 58-64.
#28	Walker, L. (2005). A Cohort Study of 'Ohana Conferencing in Child Abuse and Neglect Cases. <i>Protecting Children</i> , 19(4), 36-46.
#29	Wang, E. W., Lambert, M. C., Johnson, L. E., Boudreau, B., Breidenbach, R., & Baumann, D. (2012). Expediting permanent placement from foster care systems: The role of family group decision-making. <i>Children and Youth Services Review</i> , 34(4), 845-850.
#31	Wheeler, C. E., & Johnson, S. (2003). Evaluating family group decision making: The Santa Clara example. <i>Protecting Children</i> , 18(1-2), 65-69.

* Two studies were included from the reports by Allan et al. (2015) and Dijkstra et al. (2016a).

** Two sources were used to extract data on one study (Partnership for Strong Families 2012; Perry et al. 2013).



7.6 Appendix 6. Study categorisations

Author (year)	Study type	Study design ¹	Efforts to make groups comparable ²
1. Allan et al. (2015) Larimer	Quasi-experimental	Cross-sectional post-test only quasi-experimental	Yes- matching
2. Allan et al. (2015) Texas	RCT	-	-
3. Beehlar (2016)	Quasi-experimental	Longitudinal time series post-test only quasi-experimental	
4. Berzin (2006)	RCT	-	-
5. Berzin et al. (2008)	RCT	-	-
6. Chambers et al. (2016)	Quasi-experimental	Outcome 2 (re-entry): Cross-sectional post-test only quasi-experimental Outcome 3 (reunification): Longitudinal time series post-test only quasi-experimental	Yes- matching
7. Crampton et al. (2007)	Quasi-experimental	Cross-sectional post-test only quasi-experimental	No
8. Crampton et al. (2011)	Quasi-experimental	Cross-sectional post-test only quasi-experimental	Yes- regression adjustment
9. Dijkstra et al. (2016) Part IV	Quasi-experimental	Cross-sectional post-test only quasi-experimental	No
10. Dijkstra et al. (2016) Part V	Quasi-experimental	Cross-sectional post-test only quasi-experimental	Yes- regression adjustment
11. Dijkstra et al. (2018a)	RCT	-	-
12. Dijkstra et al. (2018b)	RCT	-	-
13. Feldman (2017)	Quasi-experimental	Cross-sectional post-test only quasi-experimental	No
14. Godinet et al. (2010)	Quasi-experimental	Cross-sectional post-test only quasi-experimental	No



15. Hollinshead et al. (2017)	RCT	-	-
16. Huebner et al. (2012)	Quasi-experimental	Cross-sectional post-test only quasi-experimental	Yes- matching
17. Lambert et al. (2017)	Quasi-experimental	Cross-sectional post-test only quasi-experimental	Yes- regression adjustment
18. Mason et al. (2017)	Quasi-experimental	Longitudinal time series post-test only quasi-experimental	No
19. Munro et al. (2017)	Quasi-experimental	Outcome 1 (out-of-home care): Trend data: Longitudinal pre-post quasi-experimental Outcome 1 (out-of-home care): Wiltshire comparison group: Cross-sectional post-test only quasi-experimental	No
20. Onrust et al. (2015)	Quasi-experimental	Cross-sectional post-test only quasi-experimental	No
21. Pennell & Burford (2000)	Quasi-experimental	Longitudinal pre-post quasi experimental	No
22. Pennell et al. (2010)	Quasi-experimental	Cross-sectional post-test only quasi-experimental	No
23. Perry et al. (2013) & Partnership for Strong Families (2012)	RCT	-	-
24. Sheets et al. (2009)	Quasi-experimental	Cross-sectional post-test only quasi-experimental	No
25. Sundell & Vinnerljung (2004)	Quasi-experimental	Cross-sectional post-test only quasi-experimental	Yes- regression adjustment
26. Teal (2013)	Quasi-experimental	Cross-sectional post-test only quasi-experimental	No
27. Titcomb & LeCroy (2005)	Quasi-experimental	Cross-sectional post-test only quasi-experimental	Yes- matching
28. Walker (2005)	Quasi-experimental	Cross-sectional post-test only quasi-experimental	No



29. Wang et al. (2012)	Quasi-experimental	Cross-sectional post-test only quasi-experimental	Yes- regression adjustment
30. Weisz et al. (2006)	Quasi-experimental	Cross-sectional post-test only quasi-experimental	No
31. Wheeler & Johnson (2003)	Quasi-experimental	Cross-sectional post-test only quasi-experimental	No
32. Wijnen-Lunenburg et al. (2008)	Quasi-experimental	Cross-sectional post-test only quasi-experimental	No
33. YMCA (2014)	RCT	-	-

Note: ¹Categorised using the designs described by Leatherdale (2019), ²Informed by Craig et al. (2017).



7.7 Appendix 7. Descriptive characteristics of included studies

Study authors and ID number	Country	Design	Target population	Sample size (int)	Sample size (con)	Control group type	Outcome timepoint(s)	Outcome measure(s)
Allan et al. (2015) Larimer #1	USA	Quasi-experimental	Families open to in-home 'ongoing' services who were referred to any type of family meeting during the study period, and could be in either the 'high risk' or 'family assessment response' track of Larimer County's child welfare system.	289 families	287 families	Family and Safety Resource Team Meetings (FSRTs)	18-24 months from first FGC	Number of children entering out-of-home-care Referrals/ re-referrals for child maltreatment
Allan et al. (2015) Texas #2	USA	RCT	Families open for Family Based Safety Services (FBSS) in Dallas and Tarrant Counties, who were referred by their FBSS caseworker for a FGC.	196 families	272 families	NR	14-32 months from first FGC	Number of children entering out-of-home care Referrals/ re-referrals for child maltreatment
Behler (2016) #3	USA	Quasi-experimental	Families referred to in-home diversion services. These families had been the focus of a child abuse or neglect investigation, but the allegations were unsubstantiated and the families diverted to other community services by a child protection investigator.	371 families	2174 families	NR	6 and 12 months from case closure	Number of children entering out-of-home care Referrals/ re-referrals for child maltreatment



Study authors and ID number	Country	Design	Target population	Sample size (int)	Sample size (con)	Control group type	Outcome timepoint(s)	Outcome measure(s)
Berzin (2006) #4	USA	RCT	<p>Fresno County targeted children ages birth to 18 years who were assessed as being at moderate to high risk for further maltreatment, as indicated by California's Structured Decision Making Family Risk Assessment, and whose families were eligible for voluntary in-home services.</p> <p>Riverside County's program was aimed at children ages 2 to 12 years who were placed in foster family or relative care and were at-risk of placement moves or placement in a higher level of care.</p>	<p>Fresno: 103 children</p> <p>Riverside: 105 children</p>	<p>Fresno: 61 children</p> <p>Riverside: 58 children</p>	Traditional services	During the study period (up to 5 years)	<p>Number of children entering out-of-home care</p> <p>Reunification rate (children returning home)</p> <p>Referrals/ re-referrals for child maltreatment</p>



Study authors and ID number	Country	Design	Target population	Sample size (int)	Sample size (con)	Control group type	Outcome timepoint(s)	Outcome measure(s)
Berzin et al. (2008) #5	USA	RCT	<p>Fresno County targeted children ages birth to 18 years who were assessed as being at moderate to high risk for further maltreatment, as indicated by California's Structured Decision Making Family Risk Assessment, and whose families were eligible for voluntary in-home services.</p> <p>Riverside County's program was aimed at children ages 2 to 12 years who were placed in foster family or relative care and were at-risk of placement moves or placement in a higher level of care.</p>	<p>Fresno: 39 children</p> <p>Riverside: 31 children</p>	<p>Fresno: 21 children</p> <p>Riverside: 19 children</p>	Traditional services	During the study period	<p>Number of children entering out-of-home care</p> <p>Reunification rate (children returning home)</p> <p>Referrals/ re-referrals for child maltreatment</p>



Study authors and ID number	Country	Design	Target population	Sample size (int)	Sample size (con)	Control group type	Outcome timepoint(s)	Outcome measure(s)
Chambers et al. (2016) #6	USA	Quasi-experimental	Families with no prior child welfare involvement and at least one child in out-of-home care.	48 families	48 families	Standard child welfare services	One year after case closed Weeks from removal to reunification One year after case closed	Number of children re-entering out-of-home care Reunification rate (children returning home) Referrals/ re-referrals for child maltreatment
Crampton & Jackson (2007) #7	USA	Quasi-experimental	Minority group children in foster care placements or adopted in Kent County, Michigan, CO.	94 families	163 families	Withdrawn and screened-out referrals	2 years after case closure Any time after child services received	Number of children entering out-of-home care Reunification rate (children returning home) Referrals/ re-referrals for child maltreatment



Study authors and ID number	Country	Design	Target population	Sample size (int)	Sample size (con)	Control group type	Outcome timepoint(s)	Outcome measure(s)
Crampton et al. (2011) #8	USA	Quasi-experimental	Children and youth living in California and other states who entered placement from 2005 to 2007.	NR	NR	NR	Within 12 months of entry to care	Reunification rate (children returning home)
Dijsktra et al. (2016) Part IV #9	Netherlands	Quasi-experimental	Families who had a supervision order (OTS) in 2014 for one or more children.	84 families	85 families	Care as Usual	3 months	Number of children entering out-of-home care Referrals/ re-referrals for child maltreatment
Dijsktra et al. (2016) Part V #10	Netherlands	Quasi-experimental	Families with one or more children placed under supervision between 7 May 2012 and 31 December 2013.	70 families	70 families	Care as Usual	12 months	Number of children entering out-of-home care Reunification rate (children returning home) Referrals/ re-referrals for child maltreatment



Study authors and ID number	Country	Design	Target population	Sample size (int)	Sample size (con)	Control group type	Outcome timepoint(s)	Outcome measure(s)
Dijkstra et al. (2018a) #11	Netherlands	RCT	Families with complex problems across various domains, such as: child maltreatment, mental health problems, alcohol abuse and other drug problems, high-conflict divorce, and child behavioural problems. For all families, child safety is at stake and in most families, risk factors for child maltreatment are present.	242 families	104 families	Care as Usual	12 months from FGC/ care plan development Pre-test and then 3 months, 6 months, 12 months after the service	Number of children entering out-of-home care Empowerment
Dijkstra et al. (2018b) #12	Netherlands	RCT	Families using Child and Youth Care Protection Services (CYPSA) in which child safety is at stake, mostly families with multi-complex problems in domains such as child maltreatment, mental health, alcohol and other drug problems, high-conflict divorce, delinquency and school problems.	46 families	23 families	Care as Usual	Pre-test and then 6 months, 12 months after service	Empowerment Referrals/ re-referrals for child maltreatment



Study authors and ID number	Country	Design	Target population	Sample size (int)	Sample size (con)	Control group type	Outcome timepoint(s)	Outcome measure(s)
Feldman (2017) #13	USA	Quasi-experimental	Participants of the Kinship Connections Program (KCP), who were screened in if they were informal kinship caregivers and not active with the State's child welfare agency.	47 families	88 families	Case Management Services - Family Success Plan (phase 2 only)	Several months after case closure At case closure	Number of children entering out-of-home care Satisfaction
Godinet et al. (2010) #14	USA	Quasi-experimental	Families who had an Ohana with the following types of cases: (1) confirmed abuse or neglect and the family agrees to voluntary foster care, (2) confirmed abuse or neglect and a court petition has been filed and the family agrees to an Ohana conference, and (3) court jurisdiction has been established and Ohana conferencing (OC) is used as resource to assist the family.	44 cases	NR	NR	Preceding 12 months	Number of children entering out-of-home care Number of children re-entering out-of-home care Reunification rate (children returning home)
Hollinshead et al. (2017) #15	USA	RCT	Families residing in two neighbouring counties in a large western state and receiving in-home child welfare services.	248 families	255 families	Business as usual services	Tracked for up to 32 months from (average) 41 days after referral	Number of children entering out-of-home care Referrals/ re-referrals for child maltreatment



Study authors and ID number	Country	Design	Target population	Sample size (int)	Sample size (con)	Control group type	Outcome timepoint(s)	Outcome measure(s)
Huebner et al. (2012) #16	USA	Quasi-experimental	Families served by The Sobriety Treatment and Recovery Teams (START), an assistance programme for families where parental substance abuse is present.	322 families	150 families	NR	Duration of case being open	Number of children entering out-of-home care
Lambert et al. (2017) #17	USA	Quasi-experimental	Families who were investigated by Texas Department of Family and Protective Services (DFPS) Child Protective Services (CPS) between 2004 and 2009.	8682 families	604,498 families	NR	During study period	Number of children entering out-of-home care
Mason et al. (2017) #18	UK	Quasi-experimental	Families in Leeds who are experiencing problems with the care and protection of children, domestic violence, youth offending, family support needs, child contact arrangements and family breakdown, and received a FGC in 2015.	NR	NR	Local authority services (statistical neighbour to Leeds)	16 months	Number of children entering out-of-home care



Study authors and ID number	Country	Design	Target population	Sample size (int)	Sample size (con)	Control group type	Outcome timepoint(s)	Outcome measure(s)
Munro et al. (2017) #19	UK	Quasi-experimental	All families to whom a letter of intent to initiate care proceedings (Public Law Outline Letter) was issued between 1 April 2015 and 31 March 2016.	Southwark: 69 cases Wiltshire: 144 cases	NR	Local authority services (3 closest statistical neighbours to Southwark and Wiltshire)	3-12 months after FGC	Number of children entering out-of-home care
Perry et al. (2013) & Partnership for Strong Families (2012) #23	USA	RCT	Families who are involved with the child welfare system in Florida, including parents/caregivers and their children aged 18 years or younger who are victims of abuse and neglect.	Pathway 2: 266 families Pathway 3: 270 families	141 families	Family Team Conferencing (Pathway 1)	<12 months of being reunified Within two months of initial FTC	Number of children re-entering out-of-home care Reunification rate (children returning home)



Study authors and ID number	Country	Design	Target population	Sample size (int)	Sample size (con)	Control group type	Outcome timepoint(s)	Outcome measure(s)
Pennell & Burford (2000) #21	Canada	Quasi-experimental	Families consisting of a mix of two-parent and lone-parent households in St. John's and the Port au Port Peninsula and three-generational households in Nain. Although the large majority of these children were residing with their parents or relatives, a sizable minority were already in family foster care, a group home, or a custody facility.	32 families	31 families	NR	1 year pre- and 1 year post-FGDM	Number of children entering out-of-home care Referrals/ re-referrals for child maltreatment
Pennell et al. (2010) #22	USA	Quasi-experimental	Children that CFSA removed from their homes during the 2005 fiscal year in Washington D.C. - the first year of FTM implementation.	454 children	Pre-FTM group: 140 children No FTM group: 195 children	NR	8 months after FTM	Reunification rate (children returning home)



Study authors and ID number	Country	Design	Target population	Sample size (int)	Sample size (con)	Control group type	Outcome timepoint(s)	Outcome measure(s)
Sheets et al. (2009) #24	USA	Quasi-experimental	Families whose child had been removed by child welfare authorities due to abuse or neglect.	468 cases	3598 cases	Permanency Planning Team (PPT) meeting	Throughout May 2006 for all cases experiencing removal between November 2004 and July 2005 Following FGC	Reunification rate (children returning home) Empowerment Satisfaction
Sundell and Vinnerljung (2004) #25	Sweden	Quasi-experimental	Families with a first-time FGC that was held between November 1996 and October 1997 in 10 local authorities, and where the child was younger than 17 years of age at the time of the initial FGC.	97 children	142 children	Traditional CPS Investigation Procedure	3 years from case closure	Number of children entering out-of-home care Referrals/ re-referrals for child maltreatment



Study authors and ID number	Country	Design	Target population	Sample size (int)	Sample size (con)	Control group type	Outcome timepoint(s)	Outcome measure(s)
Teal (2013) #26	USA	Quasi-experimental	Children who were placed in alternative care in CY 2002 and exited by 2005, and children who were placed in alternative care in CY 2005 and exited by 2008.	308 children	447 children	Care as Usual	Up to 3 years	Reunification rate (children returning home)
Titcomb & LeCroy (2005) #27	USA	Quasi-experimental	Families who worked in partnership with formal systems to create and follow-through on child safety (and permanency) plans.	291 families	249 families	NR	Within 6 months and 12 months of referral	Referrals/ re-referrals for child maltreatment
Walker (2005) #28	USA	Quasi-experimental	Cases studied in each group were randomly selected from all the CPS cases where parents voluntarily agreed to foster care for their children.	33 families	27 families	NR	NR During 6-month research period	Number of children entering out-of-home care Satisfaction



Study authors and ID number	Country	Design	Target population	Sample size (int)	Sample size (con)	Control group type	Outcome timepoint(s)	Outcome measure(s)
Wang et al. (2012) #29	USA	Quasi-experimental	Youths in the child welfare system in Texas. All had been investigated by Child Protective Services, removed from their homes, and placed in care for longer than a three-day period between 2004 and 2009.	7986 children	Placed without FGC: 52,249 children Unplaced: 20,458 children	NR	3 months, 6 months, 9 months, 12 months, 15 months and 18 months	Reunification rate (children returning home)
Weisz et al. (2006) #30	USA	Quasi-experimental	Children for whom a conference was held within 30 days of removal from the home.	33 children	33 children	Care as Usual	NR	Reunification rate (children returning home)
Wheeler & Johnson (2003) #31	USA	Quasi-experimental	Children and families involved in all phases of the child welfare system from the Santa Clara County Department of Family and Children's Services.	161 families	NR	NR	8 months from FCM and matched time period for control group	Reunification rate (children returning home) Referrals/ re-referrals for child maltreatment



Study authors and ID number	Country	Design	Target population	Sample size (int)	Sample size (con)	Control group type	Outcome timepoint(s)	Outcome measure(s)
Wi-jnen-Lunenburg et al. (2008) #32	Netherlands	RCT	Children and young people who have come into contact with the juvenile justice system.	113 families	299 families	Regular FGC group	Prior to FGC or in immediate period after	Number of children entering out-of-home-care
YMCA (2014) #33	USA	RCT	Families with substantiated CWS referrals based on the findings from the emergency response social worker.	115 children	200 children	CWS Voluntary Services 'Services as usual'	Following FGC 12 months post-enrolment	Satisfaction Referrals/ re-referrals for child maltreatment

NR= not reported



7.8 Appendix 8. Summary of shared decision-making family meetings

Author (Year)	Country	Service name	Objective	Characteristics of SDEM	Steps of family group Meetings				Who?	Where?	How much?	How well?
					(1) Referral	(2) Preparation	(3) Conference	(4) Implementation				
Allan et al. (2015) Texas site	USA	Family Group Conference	To involve families in decision making and prevent out-of-home placement for children and young people.	NR	Present	Present	Present	Present	Facilitators	39% took place in a community setting (e.g. library or community centre) and 26% took place in the family's home.	Conferences were held within 41 days of a referral (SD = 31.7 days). Conferences lasted 2 hours (SD= 40 minutes).	The 2:1 family to professional ratio was met 45% of the time; 95% of plans were approved.



Author (Year)	Country	Service name	Objective	Characteristics of SDFM	Steps of family group Meetings				Who?	Where?	How much?	How well?
Allan et al. (2015) Larimer site	USA	Family Group Conference & Family Unity Meetings	To involve families in decision making and prevent out-of-home placement for children and young people.	NR	Present	Present	Present	Present	Facilitators	Family Group Conference: 56% were in a community setting. Family Unity Meetings: 82% were at the child services agency.	Family Group Conference: 60 days (SD= 44.62) between referral and first conference. Conferences took 2 hours 12 minutes. (SD= 34 minutes) Family Unity Meetings: 68 days (SD= 90.06) between referral and first meeting. Meetings took 90 minutes (SD= 22 minutes).	Family Group Conference: The 2:1 family to professional ratio was met 49% of the time; 71% of plans were approved. Family Unity Meetings: The 2:1 family to professional ratio was met 40% of the time, 93% of plans were approved.



Author (Year)	Country	Service name	Objective	Characteristics of SDFM	Steps of family group Meetings				Who?	Where?	How much?	How well?
Beehler (2016)	USA	Family Team Conference	For families to develop a plan which addresses child safety and improves the family's circumstance.	Families were prepared for the conference beforehand. During the meeting, attendees discussed the family's strengths and needs. The family created a plan that addressed child safety and the family were referred to services that can help with specific needs (e.g. substance misuse). Professionals at the meeting developed a contingency plan that detailed the team responsibilities in supporting the family to maintain the child's safety.	Present	Present	Present	Present	Community facilitator	Neutral family friendly location.	NR	During the project fidelity improved from 88% to 94%.
Berzin (2006)	USA	Family Group Decision Making	To involve families in problem-solving and creating a plan that will improve the welfare of the family.	A blend of Family Unity and Family Group Conference models. The service incorporated a strengths assessment. During the conference there was private family time.	NR	NR	NR	NR	NR	NR	Services were provided for six months with one conference.	NR



Author (Year)	Country	Service name	Objective	Characteristics of SDFM	Steps of family group Meetings				Who?	Where?	How much?	How well?
Berzin et al. (2008) Fresno site	USA	Family Group Decision Making	To empower families to make changes that will keep children safe.	A blend of Family Unity and Family Group Conference models. The service incorporated a strengths assessment. During the conference there was private family time.	NR	NR	NR	NR	NR	NR	Services were provided for six months with one conference.	NR
Berzin et al. (2008) Riverside site	USA	Family Group Decision Making	To empower families to make changes that will keep children safe.	The service used the Family Unity Model and included a strengths assessment. All conference participants were involved in the development of a plan. Families were provided with multiple conferences after their initial conference.	NR	NR	NR	NR	Social workers	NR	Services provided until case closure.	NR



Author (Year)	Country	Service name	Objective	Characteristics of SDFM	Steps of family group Meetings				Who?	Where?	How much?	How well?
Chambers et al. (2016)	USA	Team Decision Making	To provide Family Team Decision Making Meetings to help achieve reunification.	As part of the Pomona Family First project, families were provided with Team Decision Making meetings. The project was based on a community partnership model and provided families with links to community services, an assigned case worker and weekly meetings between parents, children and caseworkers.	NR	NR	Present	NR	Caseworker	NR	NR	NR
Crampton et al. (2011)	USA	Team Decision Making	To involve community and family members in reviewing proposed removals of children or changes in their placement.	Family and friends were invited to the meeting by the family. A trained facilitator guided the decision-making process and a placement decision was made during the meeting.	NR	NR	Present	NR	Trained facilitator	A community location away from the public child welfare agency.	NR	NR



Author (Year)	Country	Service name	Objective	Characteristics of SDFM	Steps of family group Meetings				Who?	Where?	How much?	How well?
Cramp-ton and Jackson (2007)	USA	Family Group Decision Making	To involve the community in placement decisions to divert children from regular foster care services.	At the start of the meeting, the professionals presented their welfare concerns. Family members then had private time to develop a plan for placement. If the plan was agreed, the family were connected with community services that could support the placement. If there is no agreement, children were placed in foster care. After placement, meetings were held every three months with the goal of returning the child home within a year.	Present	Present	Present	Present	Family advocate	NR	NR	NR



Author (Year)	Country	Service name	Objective	Characteristics of SDFM	Steps of family group Meetings				Who?	Where?	How much?	How well?
Dijkstra et al. (2016a) Part IV	Netherlands	Family Group Conference	To establish a family group plan in which the family take control of the concerns or problems in the family.	Different forms of conferences were used. All meetings included a preparation phase before the meeting, a deliberation phase where a plan was developed and an implementation phase. Some types of conference included an independent coordinator who supported the family and helped to organise the meeting whereas others did not. In addition, some models consisted of a single conference whereas others developed a plan over several meetings.	Present	Present	Present	Present	In some cases, an independent coordinator.	NR	The meeting was held on average 18 weeks after an offer of a meeting.	NR
Dijkstra et al. (2016a) Part V	Netherlands	Family Group Conference	To establish a family group plan in which the family take control of the concerns or problems in the family.	Families were prepared for the meeting and then during the meeting a plan was developed.	NR	Present	Present	Present	Youth care worker	NR	It took 10 weeks (range 0-37) to make a plan.	NR



Author (Year)	Country	Service name	Objective	Characteristics of SDFM	Steps of family group Meetings				Who?	Where?	How much?	How well?
Dijkstra et al. (2018a)	Netherlands	Eigen Kracht-conferentie (Own Strength conference)	To give families the opportunity to make their own care plan.	An independent coordinator supported the family to organise the meeting and invite their extended family. In the conference, professionals shared information on the care options and, if necessary, provided conditions for the plan. The family then had private time to develop their plan. The plan was discussed and agreed by the family, coordinator and professionals. The care plan was then be implemented and monitored by professionals.	Present	Present	Present	Present	Independent coordinator	NR	It took 27 weeks (SD= 15.37) after referral to make a plan.	NR



Author (Year)	Country	Service name	Objective	Characteristics of SDFM	Steps of family group Meetings				Who?	Where?	How much?	How well?
Dijkstra et al. (2018b)	Netherlands	Eigen Kracht-conferentie (Own Strength conference)	To give families the opportunity to make their own care plan.	An independent coordinator supported the family to organise the meeting and invite their extended family. In the conference, professionals shared information on the care options and, if necessary, provided conditions for the plan. The family then had private time to develop their plan. The plan was discussed and agreed by the family, coordinator and professionals. The care plan was then implemented and monitored by professionals.	Present	Present	Present	Present	Independent coordinator	NR	It took 29 weeks (SD=13.13, range 8-59) to make a plan.	NR
Feldman (2017)	USA	Family Group Decision Making (also known as Family Success Conference)	To establish a plan of action "to address expressed needs and identification of responsible parties".	NR	Absent	Present	Present	NR	Kinship Connections Program worker	Grand Family Drop-in Centre or in some instances, the caregiver's home.	42.8 service hours (SD = 18.8).	Study authors concluded that most of the model requirements were met.



Author (Year)	Country	Service name	Objective	Characteristics of SDFM	Steps of family group Meetings				Who?	Where?	How much?	How well?
Godinet et al. (2010)	USA	Ohana Conference	To offer a respectful, collaborative, solution-oriented process that protects children, strengthens families, and enhances the health of the community.	<p>The central values of the approach were; family centred, strengths-focused, community-based and culturally relevant.</p> <p>Families were encouraged to invite their family members to the meeting.</p> <p>The first part of the meeting was information sharing. Then the family had private time to develop a plan for the child. The professionals returned to the room to hear the plan and the plan was negotiated and agreed by all parties. If no agreement was reached, the judge who presided over the case made a final decision. A social worker continued to monitor the case including compliance with the plan.</p>	NR	NR	Present	Present	Two trained counsellors	Accessible community location	NR	NR



Author (Year)	Country	Service name	Objective	Characteristics of SDFM	Steps of family group Meetings				Who?	Where?	How much?	How well?
Hollinshead et al. (2017)	USA	Family Group Conference	To give families receiving child welfare services the opportunity to lead decision making.	The conferences were led by an independent coordinator and included private family time. Solution-focused practices were incorporated into the meeting.	Present	NR	NR	NR	Independent coordinator	NR	The conference occurred on average 41 days after the referral (range 1-279 days)	NR
Huebner et al. (2012)	USA	Family Team Meeting, as part of Sobriety Treatment and Recovery Teams (START)	For family, relatives and professionals to participate in developing a plan to keep children safe.	Within 1-2 days of the referral, the START program began with a Family Team Meeting. Family, relatives and professionals attended the meeting to establish a plan for keeping children safe. The family's strengths and needs were discussed during the meeting and resources for meeting the family's needs were identified.	Present	NR	Present	Present	Trained social service worker and family mentor	NR	START cases were typically open for 14.2 months	Study authors conclude that fidelity to the model improved over time.



Author (Year)	Country	Service name	Objective	Characteristics of SDFM	Steps of family group Meetings				Who?	Where?	How much?	How well?
Lambert et al. (2017)	USA	Family Team Meeting	To quickly respond to child safety concerns by engaging family, community members and other caregivers in critical decisions.	Family Team Meetings were attended by family, community members and other caregivers. Attendees were involved in making decisions about protecting children.	NR	NR	Present	NR	NR	NR	NR	NR
Mason et al. (2017)	UK	Family Group Conference (as part of Leeds Family Valued)	To ensure the child's family has "the opportunity to lead the planning needed to resolve their difficulties; and within this, that the needs and wishes of vulnerable family members (both children and adults) are heard and respected"	The service followed the principles of restorative approaches. Conferences were family-led and incorporated private family time. It was optional for families to engage in the process. Family plans and resources were negotiated and agreed unless they could cause harm.	Present	Present	Present	Present	Independent coordinator	Venues were selected with the aim of putting family members at ease.	NR	NR



Author (Year)	Country	Service name	Objective	Characteristics of SDFM	Steps of family group Meetings				Who?	Where?	How much?	How well?
Munro et al. (2017)	UK	Family Group Conferences	For "families and friends to make decisions and plans for resolving problems around a child, young person or vulnerable adult."	Families were prepared for the conference by a coordinator beforehand. In the conference, professionals started by sharing their information and families were then given private time to develop a plan. Professionals returned to the room to agree the plan. The family then arranged a time to meet again to review the plan and make any changes.	Present	Present	Present	NR	Independent coordinator	Neutral location e.g. community centre	Southwark: 42.7 days from referral to Family Group Conference (range 9-142 days) Wiltshire: 23.7 days from referral to Family Group Conference (range 5-76 days)	NR



Author (Year)	Country	Service name	Objective	Characteristics of SDFM	Steps of family group Meetings				Who?	Where?	How much?	How well?
Partnership for Strong Families (2012) & Perry et al. (2013)	USA	Family Team Conference	<p>“To respectfully engage families in decision making and case planning through a strength based, family-centered, culturally appropriate system of care that includes initial and ongoing Family Team Conferences”</p>	<p>Pathway 2: Families were prepared for the conference. The extended family were invited to attend along with key individuals including the facilitator, service providers and family supports that are invited by the family. Follow-up conferences were held at key points throughout the case.</p> <p>Pathway 3: The same process as pathway 2 with the addition of private family time.</p>	Present	Present	Present	Present	Family service facilitator	A neutral, family friendly location e.g. family visitation centre	Families typically receive services for a period of < 10 months	It was concluded from observations that facilitators delivered the conferences with high fidelity.



Author (Year)	Country	Service name	Objective	Characteristics of SDFM	Steps of family group Meetings				Who?	Where?	How much?	How well?
Pennell and Burford (2000)	Canada	Family Group Decision Making	"To advance child and safety and strengthen family unity while expanding its meaning"	A coordinator organised the conference with the involvement of the family members. The conference started with an opening in the culture of that family group (e.g. a prayer). The coordinator reviewed the purpose of the meeting and the services provider shared information (including reports from authorities or any concerns). The family were then given private time to develop a plan. The plan was then approved by the referring agency.	Present	Present	Present	Present	Coordinator	NR	On average conferences lasted 5.5 hours	NR



Author (Year)	Country	Service name	Objective	Characteristics of SDFM	Steps of family group Meetings				Who?	Where?	How much?	How well?
Pennell et al. (2010)	USA	Family Team Meeting	To engage families, family supports and professionals in creating plans for children's safety.	Family Team Meetings were convened within a 72-hour period in advance of the hearing on where children were to live. A coordinator prepared the family before the meeting. The meeting was focused on child placement and did not formally incorporate private family time.	Present	Present	Present	NR	Coordinator prepares the family and a separate facilitator conducts the meeting.	NR	The process takes three or more weeks. The meeting lasted between 1 – 2.5 hours.	Study authors concluded the model was delivered as planned. Agreement was highest for the involvement of family members and lowest for the consistent follow up of plans.
Sheets et al. (2009)	USA	Family Group Decision Making	To involve extended family members in the development of a plan during the first 30-45 days of a child coming into foster care due to abuse or neglect.	Extended family members were invited to the conference where the family's strengths and wishes were reviewed. The family was given private time to develop a plan.	NR	NR	NR	NR	NR	NR	NR	NR



Author (Year)	Country	Service name	Objective	Characteristics of SDEM	Steps of family group Meetings				Who?	Where?	How much?	How well?
Sundell and Vinnerljung (2004)	Sweden	Family Group Conference	To involve families in decision making and problem solving.	NR	NR	Present	NR	NR	NR	NR	Services were provided for 514 days on average.	NR
Teal (2013)	USA	Team Decision Making	To involve families in making placement decisions for children.	NR	NR	NR	NR	NR	NR	NR	NR	NR
Titcomb and LeCroy (2003)	USA	Family Group Decision Making	To encourage families to work collaboratively with professionals to develop and implement child safety plans.		Present	Present	Present	NR	NR	NR	NR	NR
Wang et al. (2012)	USA	Family Group Conference	To offer families the opportunity to develop a plan "that ensures children are cared for and protected from future harm".	NR	NR	NR	NR	NR	NR	NR	NR	NR



Author (Year)	Country	Service name	Objective	Characteristics of SDFM	Steps of family group Meetings				Who?	Where?	How much?	How well?
Walker (2005)	USA	Ohana Conference	For families to participate in the development of a plan "to deal with the report of child abuse and neglect"	Families and professionals attended a conference to discuss the child abuse and neglect problems. Families were provided with private time to formulate a plan. The professionals then listened to the plan and the social worker decided if it could be accepted. If the plan was not accepted, the facilitator led negotiation between the social worker and family.	Present	NR	Present	NR	Community facilitator	NR	NR	NR



Author (Year)	Country	Service name	Objective	Characteristics of SDFM	Steps of family group Meetings				Who?	Where?	How much?	How well?
Weisz et al. (2006)	USA	Family Group Conference	To involve families in creating plans for child safety.	<p>The coordinator met with the family and arranged to invite the extended family members to the meeting. In the meeting, participants shared their concerns relating to the child's welfare and discussed the family strengths.</p> <p>The family had private time to develop a permanency plan for the child. The plan was then presented to professionals and the group clarified and adjusted the plan until agreement was reached.</p>	Present	NR	Present	NR	Coordinator	NR	NR	NR



Author (Year)	Country	Service name	Objective	Characteristics of SDFM	Steps of family group Meetings				Who?	Where?	How much?	How well?
Wheeler and Johnson (2003)	USA	Family Group Conference	To support families to make decisions that ensure the safety of their children.	The model was based on the models used for Family Group Conferences and Family Unity Meetings. Families were prepared for the meeting by a facilitator who was matched to their needs and primary language. The conference consisted of a discussion about the family strengths, concerns and possible options for the family to consider. The family had private time to develop a plan which was then reviewed by all the meeting participants.	NR	Present	Present	NR	Facilitator	NR	The entire process takes approximately 45 days. The conferences lasts 4 hours on average.	NR



Author (Year)	Country	Service name	Objective	Characteristics of SDFM	Steps of family group Meetings				Who?	Where?	How much?	How well?
Wijenjn-Lunenburg et al. (2008)	Netherlands	Family Group Conference	To assist families in making decisions and solving problems that improve the wellbeing and safety of their child.	Independent coordinators prepared the family for a meeting which was attended by family members, friends and professionals. Families were given private time to develop a plan. The plan was agreed by professionals if it was safe for the child.	Present	NR	Present	Present	Independent coordinator	NR	NR	Nearly three quarters of plans were not fully implemented.
YMCA Families United (2014)	USA	Family Group Conference	To bring "families and invited participants together to discuss how to implement and execute a plan for safety and stable placement for a child".	A coordinator prepared the family for the meeting. The meeting incorporated the family's cultural practices and offered them private time to develop a plan. Coordinators facilitated two to five meetings and provided ongoing support to help families achieve their goals.	Present	Present	Present	Present	Coordinator	NR	Planning time was 4.2 months (range 2-7 months). Meetings lasted 4.6 hours (range 2-7 hours).	100% on the fidelity checklist. Study authors concluded practice was in line with the model at all times.

Note: NR= not reported, SD= standard deviation



7.9 Appendix 9: Glossary table of economic terms

Term	Definition
Economic evaluation	An evaluation that compares both the costs and outcomes of a service of interest against a suitable comparator.
Cost-analysis	A partial economic evaluation comparing the costs of a service and a comparator.
Cost-benefit analysis	A full economic evaluation where both costs and outcomes are measured in monetary terms.
Cost-consequence analysis	A full economic evaluation where a list of disaggregated costs and a range of appropriate outcomes are reported for both the service and comparator.
Cost-effectiveness analysis	A full economic evaluation where costs are measured in monetary terms and outcomes are measured in units directly related to the service e.g. number of children who avoided care.
Cost-minimisation analysis	A full economic evaluation that is used when outcomes across the service and comparator arms of a study are known to be equivalent. In this type of economic evaluation, costs only are compared with the aim of deciding on the least costly service to implement to achieve the same outcome.
Cost-cost offset analysis	A partial economic evaluation measuring the cost savings as a result of the service.
Cost-utility analysis	A full economic evaluation where costs are measured in monetary terms and outcomes are measured using quality-adjusted life years that capture the effects on both the extension and the quality of life in a single metric.
Discounting	The adjustment of the value of future costs and benefits to reflect their current value. This is generally necessary where studies adopt a time horizon longer than 1-year.
Incremental cost-effectiveness ratio	The main result of a full economic evaluation calculated by dividing the difference in costs by the difference in outcomes to provide a ratio of the incremental cost per extra unit of benefit.
Models	Model based economic evaluations apply mathematical techniques using computer software to estimate cost-effectiveness.
Preference based outcomes	These are used in the evaluation of healthcare interventions to estimate quality-adjusted life years. They use questions with more than one domain for patients to describe their health, these are converted to utility values using algorithms that are based on general public preferences for each health state.
Study perspective	The point of view adopted when deciding the types of cost and outcomes to be included in an economic evaluation. The perspective of a study can be narrow, reflecting social care costs and outcomes only, or it can be broad, capturing all costs and outcomes to society.

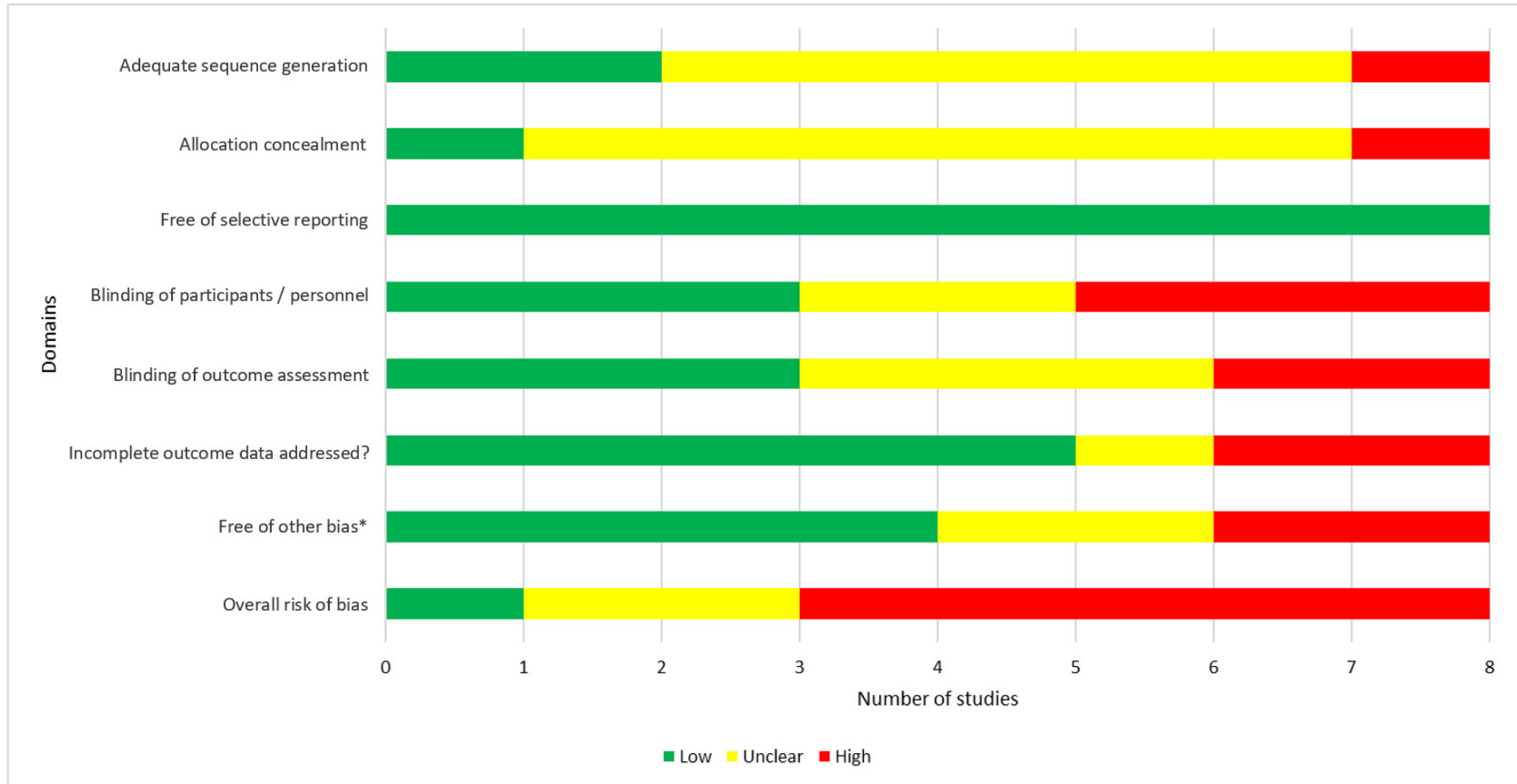


7.10 Appendix 10a. Critical appraisal of included RCTs assessed by the Cochrane Risk of Bias tool

Study (Author, year, country)	Adequate sequence generation	Allocation concealment	Selective reporting	Blinding of participants/ personnel	Blinding of outcome assessment	Incomplete data	Other bias	Overall risk of bias judgement
Allan et al. 2015 (Texas), USA	Unclear	High	Low	High	Unclear	Low	High	High
Berzin 2006, USA	Unclear	Unclear	Low	Low	Low	Unclear	Low	Unclear
Berzin et al. 2008, USA	High	Unclear	Low	Low	Unclear	Low	Unclear	High
Dijkstra et al. 2018a, Netherlands	Unclear	Unclear	Low	Unclear	Unclear	Low	Unclear	Unclear
Dijkstra et al. 2018b, Netherlands	Unclear	Unclear	Low	High	High	High	Low	High
Hollinshead et al. 2017, USA	Unclear	Unclear	Low	Unclear	Low	Low	High	High
Perry et al. 2013, USA and Partnership for Strong Families 2012, USA	Low	Unclear	Low	High	High	High	Low	High
YMCA 2014, USA	Low	Low	Low	Low	Low	Low	Low	Low



7.10 Appendix 10b. Risk of bias graph: review authors' judgments about each risk of bias item presented a percentage across all included RCTs





7.10 Appendix 10c. Risk of bias summary: review authors judgments about each risk of bias item for each RCT

Study (Author, year, country)	Adequate sequence generation	Allocation concealment	Free of selective reporting	Blinding of participants/	Blinding of outcome assessment	Incomplete outcome data	Free of other bias*	Overall risk of bias
Allan et al. 2015 (Texas), USA								
<u>Berzin</u> 2006, USA								
<u>Berzin</u> et al. 2008, USA								
Dijkstra et al. 2018a, <u>Netherlands</u>								
Dijkstra et al. 2018b, <u>Netherlands</u>								
<u>Hollinshead</u> et al. 2017, USA								
Perry et al. 2013, USA and Partnership for Strong Families, 2012, USA								
YMCA 2014, USA								



7.11 Appendix 11. Critical appraisal of included quasi-experimental studies assessed by the ROBINS-I tool

Study (Author, year, country)	Domain							Overall risk of bias
	Bias due to confounding	Bias in selection	Bias in classification	Bias due to deviations	Bias due to missing data	Bias in measurement of outcomes	Bias in selection of the reported result	
Allan et al. 2015 (Larimer), USA	Low	No information	Moderate	Low	Low	Moderate	Moderate	Moderate
Beehler 2016, USA	No information	No information	Serious	No information	No information	No information	Moderate	Serious
Chambers et al. 2016, USA	Moderate	Moderate	Low	No information	Low	Low	Moderate	Moderate
Crampton et al. 2011, USA	Serious	Moderate	Moderate	No information	No information	Serious	Moderate	Serious
Crampton & Jackson 2007, USA	Serious	Serious	Moderate	Serious	Serious	Serious	Serious	Serious
Dijkstra et al. 2016a (Part IV), Netherlands	Moderate	Moderate	Low	No information	Moderate	Moderate	Moderate	Moderate
Dijkstra et al. 2016a (Part V), Netherlands	Moderate	Moderate	Low	No information	Moderate	Low	Moderate	Moderate
Feldman 2017, USA	No information	Serious	Low	Moderate	No information	Moderate	Moderate	Serious
Godinet et al. 2010, USA	No information	No information	Serious	No information	Serious	Low	Moderate	Serious



Study (Author, year, country)	Domain							Overall risk of bias
	1	2	3	4	5	6	7	
Huebner et al. 2012, USA	No information	Moderate	No information	No information	Low	Serious	Moderate	Serious
Lambert et al. 2017, USA	Low	Serious	Moderate	No information	No information	Moderate	Moderate	Serious
Mason et al. 2017, UK	Serious	Moderate	No information	Serious	Serious	Moderate	Moderate	Serious
Munro et al. 2017, UK	Serious	No information	No information	No information	No information	Serious	Moderate	Serious
Onrust et al. 2015, Netherlands	Moderate	Serious	Low	No information	Moderate	Low	Low	Serious
Pennell & Burford 2000, Canada	Serious	Low	Low	No information	Low	Low	Moderate	Serious
Pennell et al. 2010, USA	Serious	Low	Moderate	No information	Low	Serious	Moderate	Serious
Sheets et al. 2009, USA	Serious	Low	Moderate	No information	No information	Moderate	Moderate	Serious
Sundell & Vinnerljung 2004, Sweden	Moderate	Low	Low	Moderate	Low	Low	Moderate	Moderate
Teal 2013, USA	No information	Low	Moderate	No information	Moderate	Low	Moderate	Moderate
Titcomb & LeCroy 2005, USA	Moderate	No information	Serious	No information	No information	Serious	Serious	Serious



Study (Author, year, country)	Domain							Overall risk of bias
Wang et al. 2012, USA	Serious	Low	Moderate	No information	Low	Low	Moderate	Serious
Walker 2005, USA	Moderate	Serious	Moderate	No information	Serious	Moderate	Serious	Serious
Weisz et al. 2006, USA	Serious	Low	Moderate	No information	Moderate	Moderate	No information	Serious
Wheeler & Johnson 2003, USA	Serious	Low	Moderate	No information	Low	Moderate	Serious	Serious
Wijnen-Lunenburg et al. 2008, Netherlands	Moderate	Serious	Moderate	Serious	No information	Moderate	Moderate	Serious



7.12 Appendix 12. GRADE summary tables

Table A: Effectiveness of shared decision-making family meetings on care entry

Outcome: Out-of-home care							
Patient or population: Children and young people ≤18 years old							
Setting: Any setting							
Service: Shared decision-making family meetings							
Comparison: Usual care							
Certainty assessment						Relative effect	Certainty
No of participants (studies)	Risk of bias	Inconsistency	Indirectness	Imprecision	Other consideration		
620,711 ¹ (20 ²)	Serious ³	Serious ⁴	Not serious	Serious ⁵	Not serious ⁶	Unable to calculate	Low
GRADE Working Group grades of evidence							
High quality: We are very confident that the true effect lies close to that of the estimate of the effect.							
Moderate quality: We are moderately confident in the effect estimate. The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.							
Low quality: Our confidence in the effect estimate is limited. The true effect may be substantially different from the estimate of the effect.							
Very low quality: We have very little confidence in the effect estimate. The true effect is likely to be substantially different from the estimate of effect.							

¹ Sample size reported for 18 of the studies. It was not reported by Mason et al., 2017; Munro et al., 2017; (trend data).



² Five studies were RCTs (Berzin et al. 2008; Allan et al. 2015 - Texas; Hollinshead et al. 2017; Dijkstra et al. 2018a) and 15 were quasi-experimental studies (Pennell and Burford 2000; Sundell and Vinnerljung 2004; Walker 2005; Crompton and Jackson 2007; Wijnen-Lunenburg et al. 2008; Godinet et al. 2010; Huebner et al. 2012; Allan et al. 2015 - Larimer; Beehler 2016; Dijkstra et al. 2016a - Parts IV and V; Feldman 2017; Lambert et al. 2017; Mason et al. 2017; Munro et al. 2017).

³ This was judged as serious as 14 of the studies were rated as serious risk of bias (Beehler, 2016; Crompton & Jackson, 2007; Feldman, 2017; Godinet et al., 2010; Huebner et al., 2012; Lambert et al., 2017; Mason et al., 2017; Munro et al., 2017, Pennell & Burford, 2000; Wijnen-Lunenburg et al., 2008; Walker, 2005) or high risk of bias (Allan et al., 2015 - Texas; Berzin et al., 2008; Hollinshead et al., 2017). Six studies had an unclear (Berzin, 2006; Dijkstra et al., 2018) or moderate (Allan et al., 2015 - Larimer; Dijkstra et al., 2016a - Parts IV and V; Sundell & Vinnerljung, 2004) risk of bias.

⁴ This domain was judged as “serious” as there is a moderate degree of inconsistency in the results, such as effects in opposite directions (i.e. benefit and harm).

⁵ This domain was judged as “serious” as the studies do not include sample size calculations or power analysis in their design.

⁶ This domain was judged as “not serious” as publication bias has been minimised by a systematic search of multiple databases, “grey literature”, no restriction to language or geography, supplementary searches, contacting a panel of international experts etc.

Table B: Effectiveness of shared decision-making family meetings on care re-entry

Outcome: Out-of-home care re-entry							
Patient or population: Children and young people ≤18 years old							
Setting: Any setting							
Service: Shared decision-making family meetings							
Comparison: Usual care							
Certainty assessment						Relative effect	Certainty
No of participants (studies)	Risk of bias	Inconsistency	Indirectness	Imprecision	Other consideration		
1074 (3 ¹)	Serious ²	Very serious ³	Not serious	Serious ⁴	Serious ⁵	Unable to calculate	Very low



GRADE Working Group grades of evidence

High quality: We are very confident that the true effect lies close to that of the estimate of the effect.

Moderate quality: We are moderately confident in the effect estimate. The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.

Low quality: Our confidence in the effect estimate is limited. The true effect may be substantially different from the estimate of the effect.

Very low quality: We have very little confidence in the effect estimate. The true effect is likely to be substantially different from the estimate of effect.

¹ One RCT (Partnership for Strong Families 2012) and two quasi-experimental studies (Godinet et al. 2010; Chambers et al. 2016).

² This domain was judged as serious as two of the studies were rated as serious (Chambers et al. 2016) or high risk of bias (Partnership for Strong Families 2012). Godinet et al. (2010) was rated as moderate risk of bias.

³ This domain was judged as "very serious" as there is a high degree of inconsistency in the results, such as effects in opposite directions (i.e. benefit and harm).

⁴ This domain was judged as "serious" as the studies do not include sample size calculations or power analysis in their design.

⁵ This domain was judged as "serious" as the results are inconclusive due to the small number of studies included.

Table C: Effectiveness of shared decision-making family meetings on reunification

Outcome: Reunification with family following a period in care.

Patient or population: Children and young people ≤18 years old

Setting: Any setting

Service: Shared decision-making family meetings

Comparison: Usual care



Certainty assessment						Relative effect	Certainty
No of participants (studies)	Risk of bias	Inconsistency	Indirectness	Imprecision	Other consideration		
77930 ¹ (14 ²)	Serious ³	Serious ⁴	Not serious	Serious ⁵	Not serious ⁶	Unable to calculate	Low

GRADE Working Group grades of evidence

High quality: We are very confident that the true effect lies close to that of the estimate of the effect.

Moderate quality: We are moderately confident in the effect estimate. The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.

Low quality: Our confidence in the effect estimate is limited. The true effect may be substantially different from the estimate of the effect.

Very low quality: We have very little confidence in the effect estimate. The true effect is likely to be substantially different from the estimate of effect.

¹ Sample size not available for Crampton, Usher, Wildire & Cuccaro-Alamin, 2011 or the control group in Wheeler & Johnson 2003.

² Three studies were RCTs (Berzin 2006; Berzin et al. 2008; Partnership for Strong Families 2012) and 11 were quasi-experimental studies (Wheeler and Johnson 2003; Weisz et al. 2006; Crampton and Jackson 2007; Sheets et al. 2009; Godinet et al. 2010; Pennell et al. 2010; Crampton et al. 2011; Wang et al. 2012; Teal 2013; Chambers et al. 2016; Dijkstra et al. 2016a - Part IV).

³ Nine of the studies were considered to be at high risk of bias (Berzin et al. 2008; Partnership for Strong Families 2012) or serious risk of bias (Wheeler and Johnson 2003; Weisz et al. 2006; Crampton and Jackson 2007; Berzin et al. 2008; Sheets et al. 2009; Godinet et al. 2010; Pennell et al. 2010; Crampton et al. 2011). Four studies had a moderate risk of bias (Wang et al. 2012; Teal 2013; Chambers et al. 2016; Dijkstra et al. 2016a - Part IV) and one was unclear (Berzin 2006).

⁴ This domain was judged as "serious" as there is heterogeneity in study methodology and sample sizes, inconsistency in the results, not a strong direction of effect.

⁵ This domain was judged as "serious" as the studies do not include sample size calculations or power analysis in their design.

⁶ This domain was judged as "not serious" as publication bias has been minimised by a systematic search of multiple databases, "grey literature", no restriction to language or geography, supplementary searches, contacting a panel of international experts etc.



Table D: Effectiveness of shared decision-making family meetings on family empowerment

Outcome: Family empowerment							
Patient or population: Children and young people ≤18 years old							
Setting: Any setting							
Service: Shared decision-making family meetings							
Comparison: Usual care							
Certainty assessment						Relative effect	Certainty
No of participants (studies)	Risk of bias	Inconsistency	Indirectness	Imprecision	Other consideration		
2415(4 ¹)	Serious ²	Not serious ³	Serious ⁴	Serious ⁵	Not serious ⁶	Unable to calculate	Low
GRADE Working Group grades of evidence							
High quality: We are very confident that the true effect lies close to that of the estimate of the effect.							
Moderate quality: We are moderately confident in the effect estimate. The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.							
Low quality: Our confidence in the effect estimate is limited. The true effect may be substantially different from the estimate of the effect.							
Very low quality: We have very little confidence in the effect estimate. The true effect is likely to be substantially different from the estimate of effect.							

¹ Three studies were RCTs (Partnership for Strong Families 2012; Dijkstra et al. 2018a; Dijkstra et al. 2018b) and one was a quasi-experimental study (Sheets et al. 2009).

² This domain was judged as “serious” as two of the RCTs had a high risk of bias (Partnership for Strong Families 2012; Dijkstra et al. 2018a) and the comparative observational study had a serious risk of bias (Partnership for Strong Families 2012). The third RCT had an unclear risk of bias (Dijkstra et al. 2018b).

³ This domain was judged as “not serious” as overall the studies suggest no difference in empowerment between family group meetings and care as usual services.

⁴ This domain was judged as “serious” due to the differences in the population and how empowerment was measured.



⁵ This domain was judged as “serious” as the studies do not include sample size calculations or power analysis in their design.

⁶ This domain was judged as “not serious” as publication bias has been minimised by a systematic search of multiple databases, “grey literature”, no restriction to language or geography, supplementary searches, contacting a panel of international experts etc.

Table E: Effectiveness of shared decision-making family meetings on family satisfaction

Outcome: Family Satisfaction							
Patient or population: Children and young people ≤18 years old							
Setting: Any setting							
Service: Shared decision-making family meetings							
Comparison: Usual care							
Certainty assessment						Relative effect	Certainty
No of participants (studies)	Risk of bias	Inconsistency	Indirectness	Imprecision	Other consideration		
1509 (4 ¹)	Very serious ²	Serious ³	Serious ⁴	Serious ⁵	Not serious ⁶	Unable to calculate	Very low
GRADE Working Group grades of evidence							
High quality: We are very confident that the true effect lies close to that of the estimate of the effect.							
Moderate quality: We are moderately confident in the effect estimate. The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.							
Low quality: Our confidence in the effect estimate is limited. The true effect may be substantially different from the estimate of the effect.							
Very low quality: We have very little confidence in the effect estimate. The true effect is likely to be substantially different from the estimate of effect.							

¹ One RCT (Partnership for Strong Families 2012) and four quasi-experimental studies (Walker 2005; Sheets et al. 2009; Beehler 2016; Feldman 2017).



² The studies were high (Partnership for Strong Families 2012) or serious risk of bias (Walker 2005; Sheets et al. 2009; Beehler 2016; Feldman 2017).

³ Even though 3 CS2 studies favoured towards the service group for this outcome, however, the included RCT found no difference.

⁴ This domain was judged as "serious" due to the differences in how satisfaction was measured.

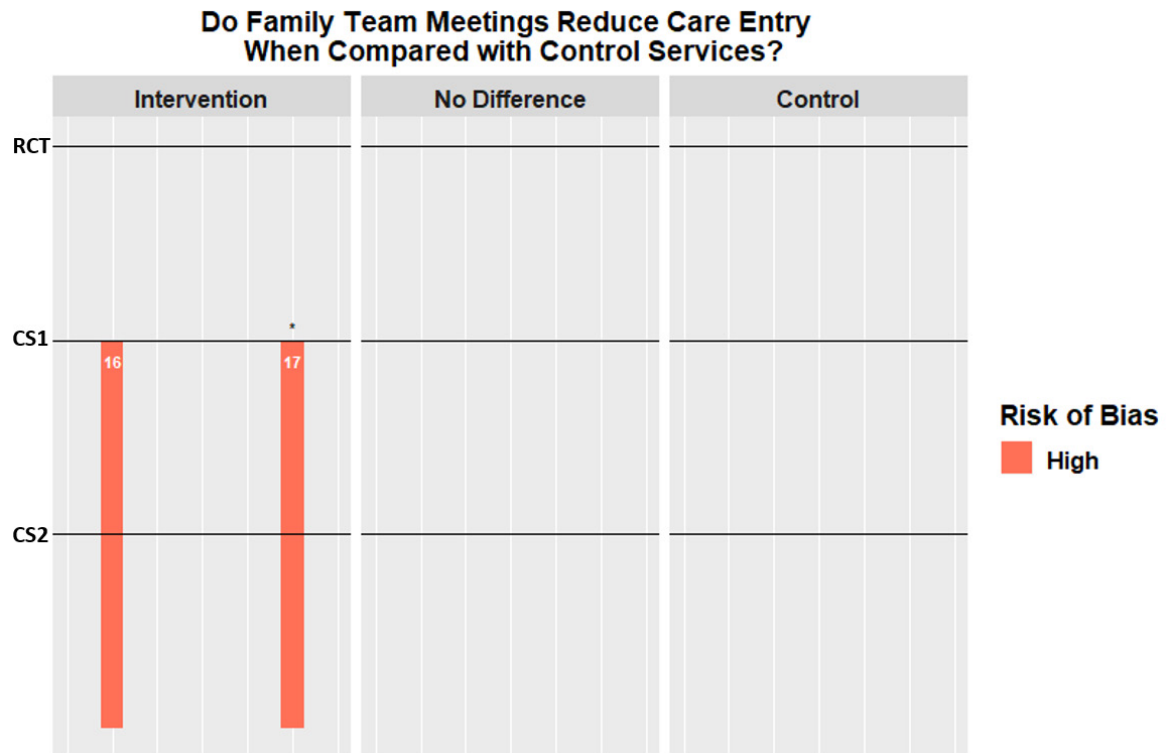
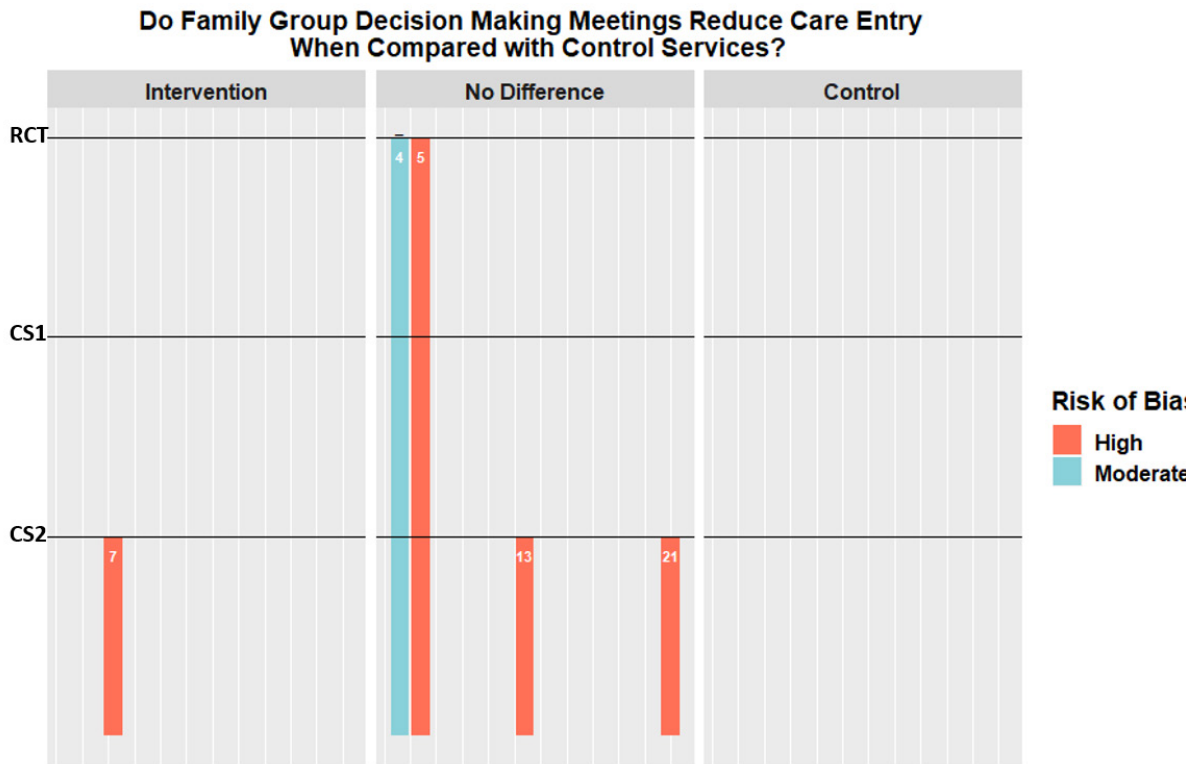
⁵ This domain was judged as "serious" as the studies do not include sample size calculations or power analysis in their design.

⁶ This domain was judged as "not serious" as publication bias has been minimised by a systematic search of multiple databases, "grey literature", no restriction to language or geography, supplementary searches, contacting a panel of international experts etc.



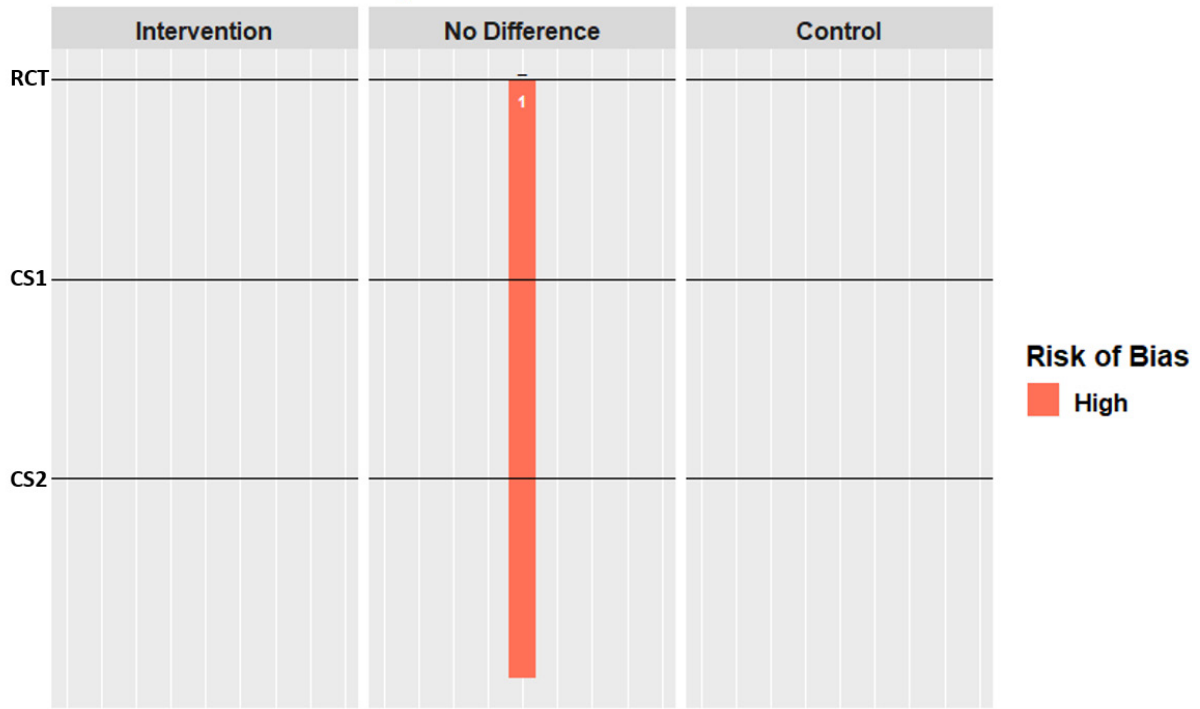
7.13 Appendix 13. Harvest plots for non-FGC meeting types

7.13 Appendix 13a. Care Entry

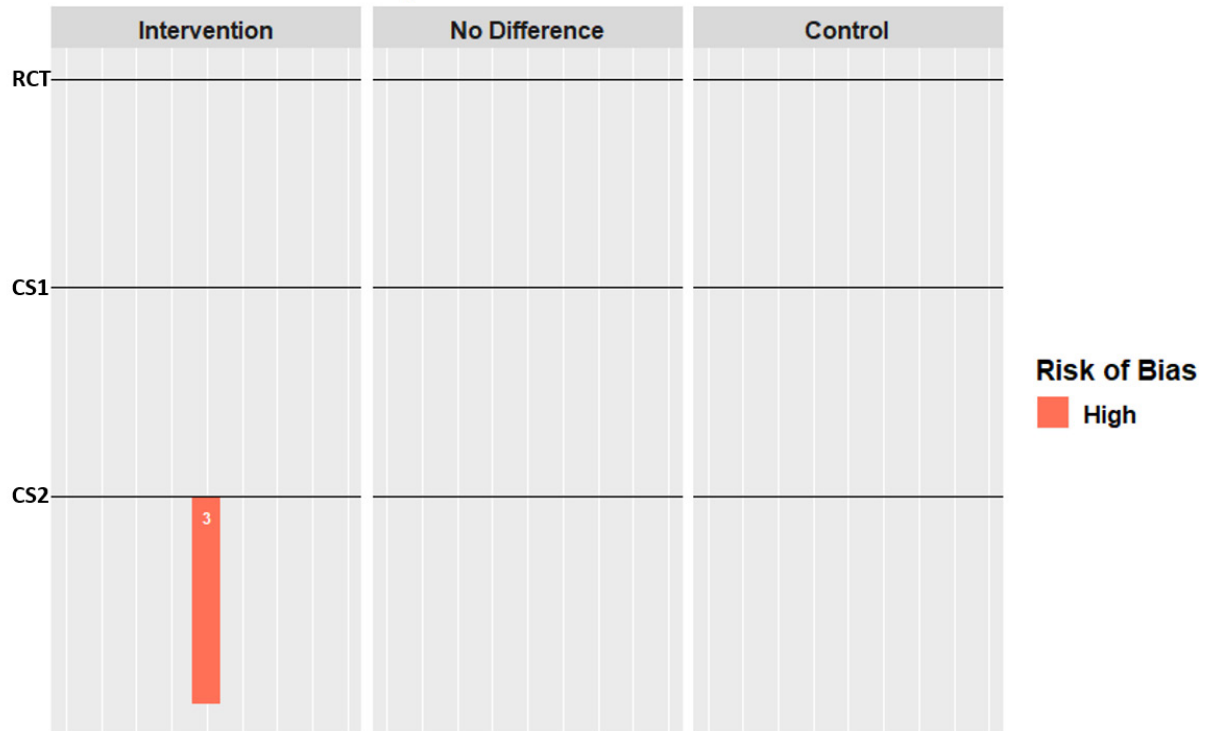




Do Family Unity Meetings Reduce Care Entry When Compared with Control Services?



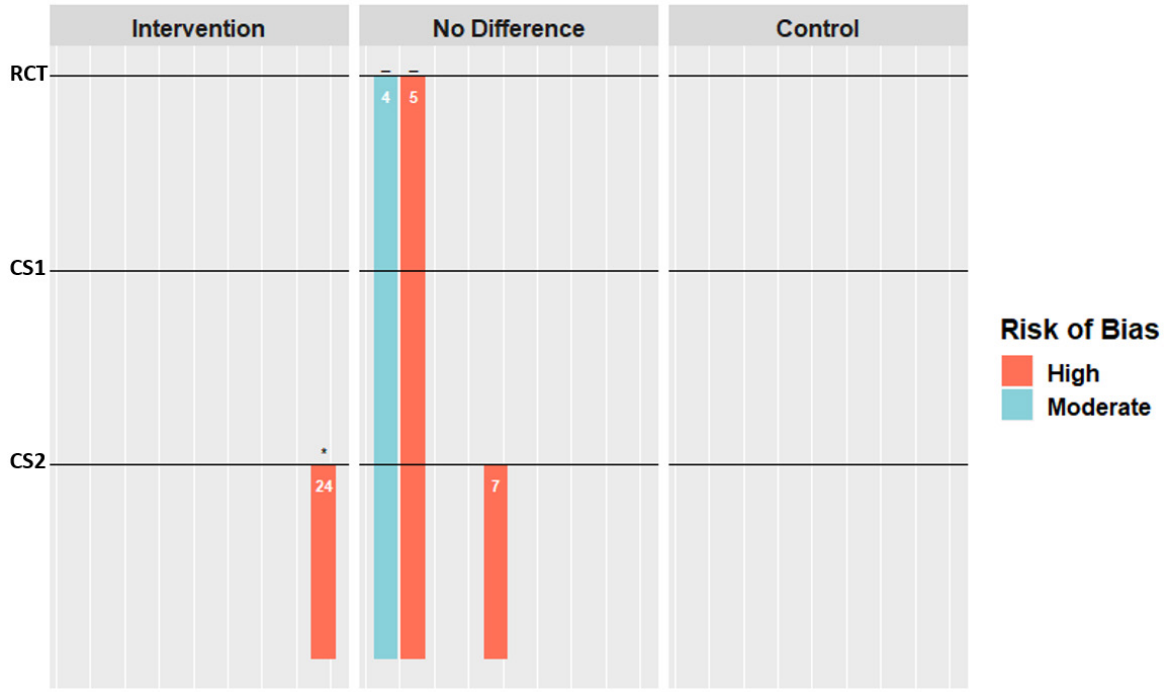
Do Family Team Conferences Reduce Care Entry When Compared with Control Services?





7.13 Appendix 13b. Reunification

Do Family Group Decision Making Meetings Result in More Children Returning Home When Compared with Control Services?

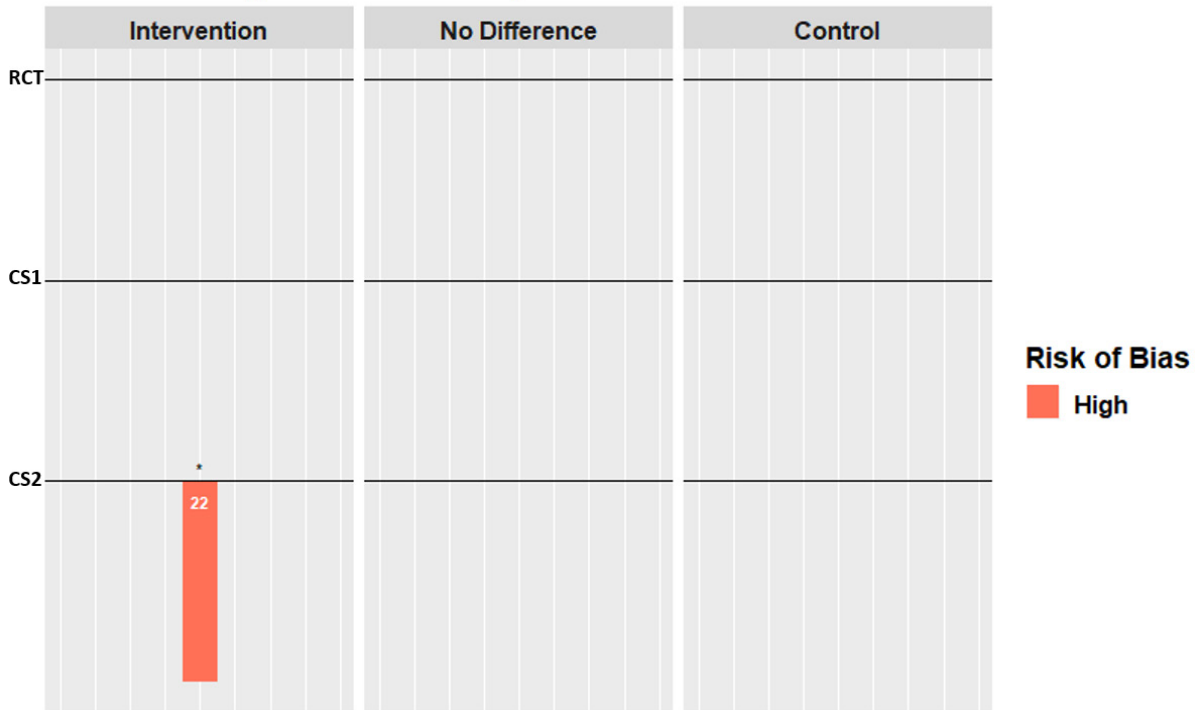


Do Team Decision Making Meetings Result in More Children Returning Home When Compared with Control Services?

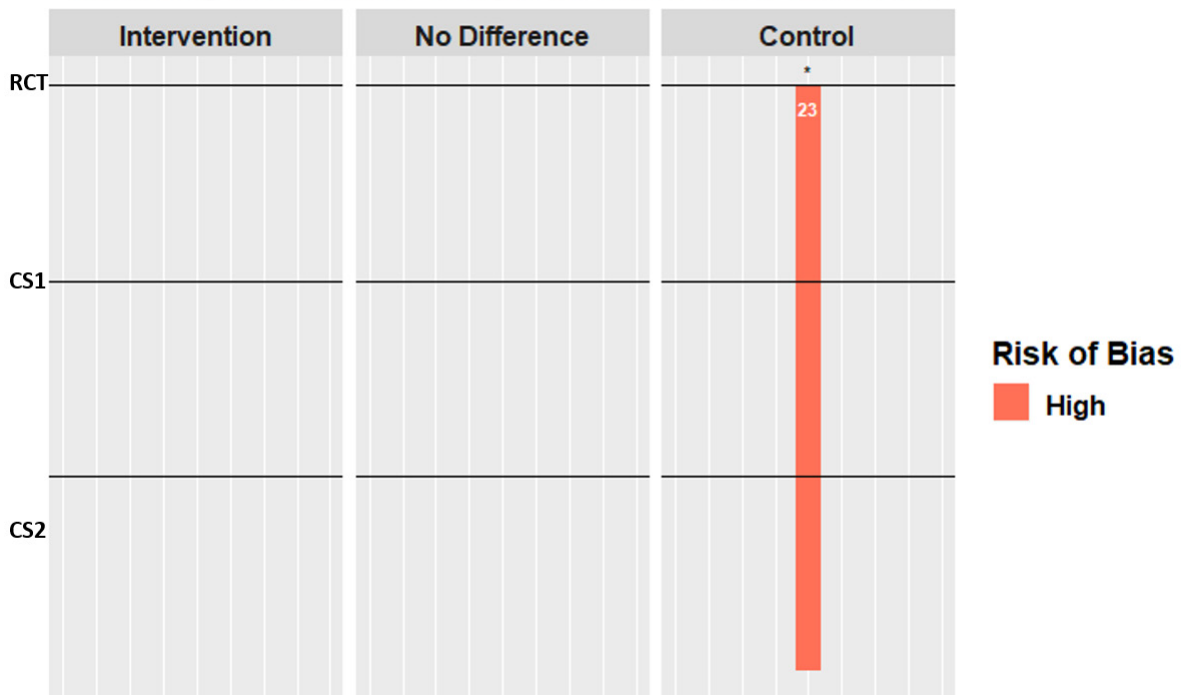




Do Family Team Meetings Result in More Children Returning Home When Compared with Control Services?

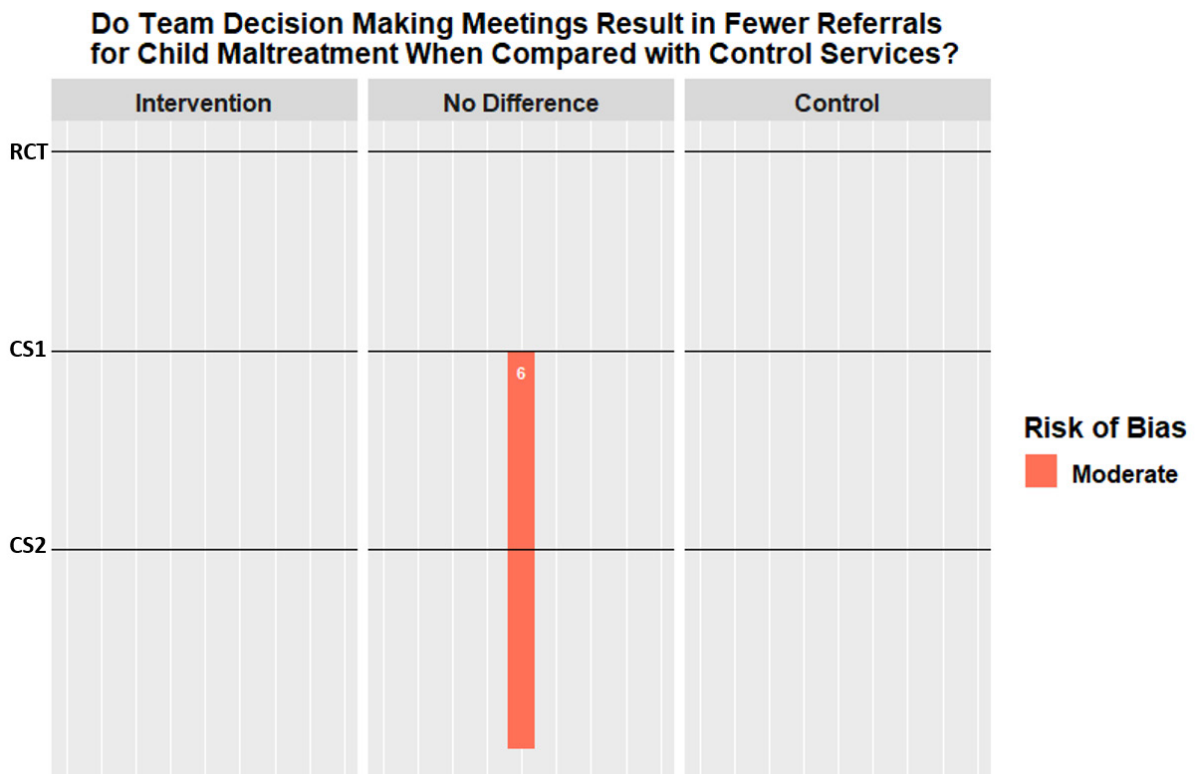
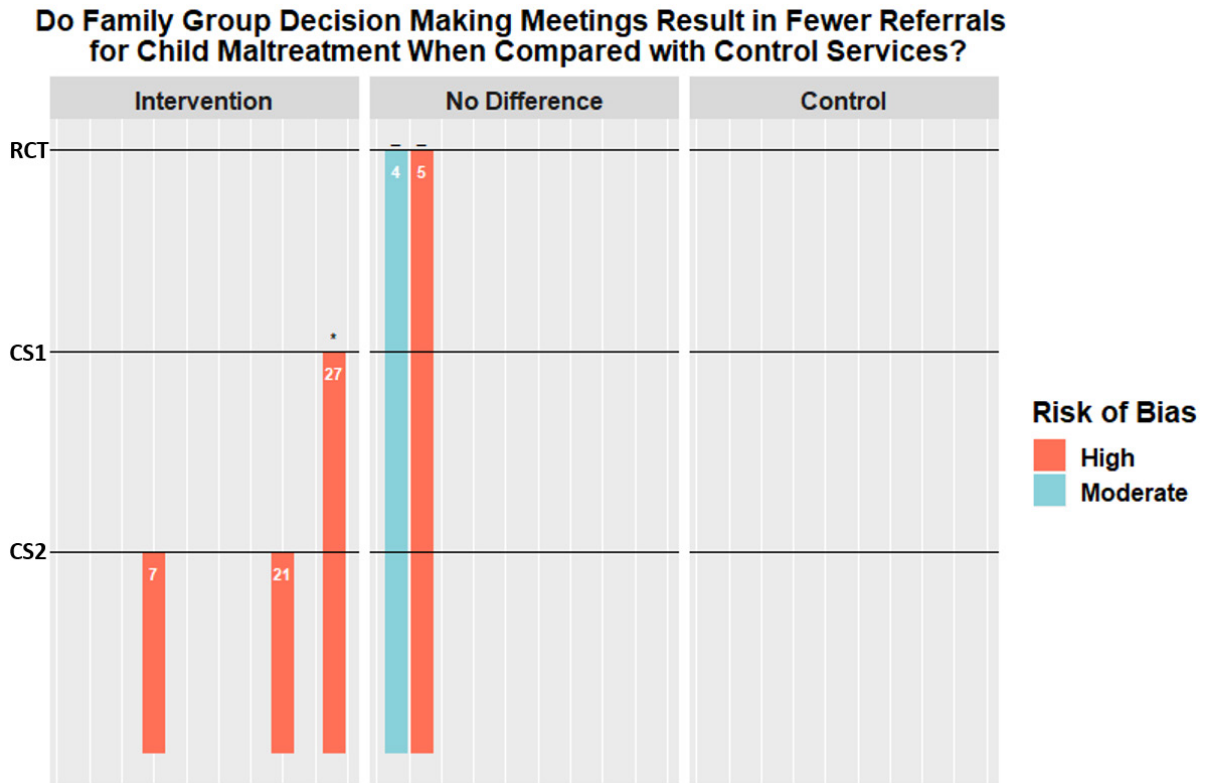


Do Family Team Conferences Result in More Children Returning Home When Compared with Control Services?



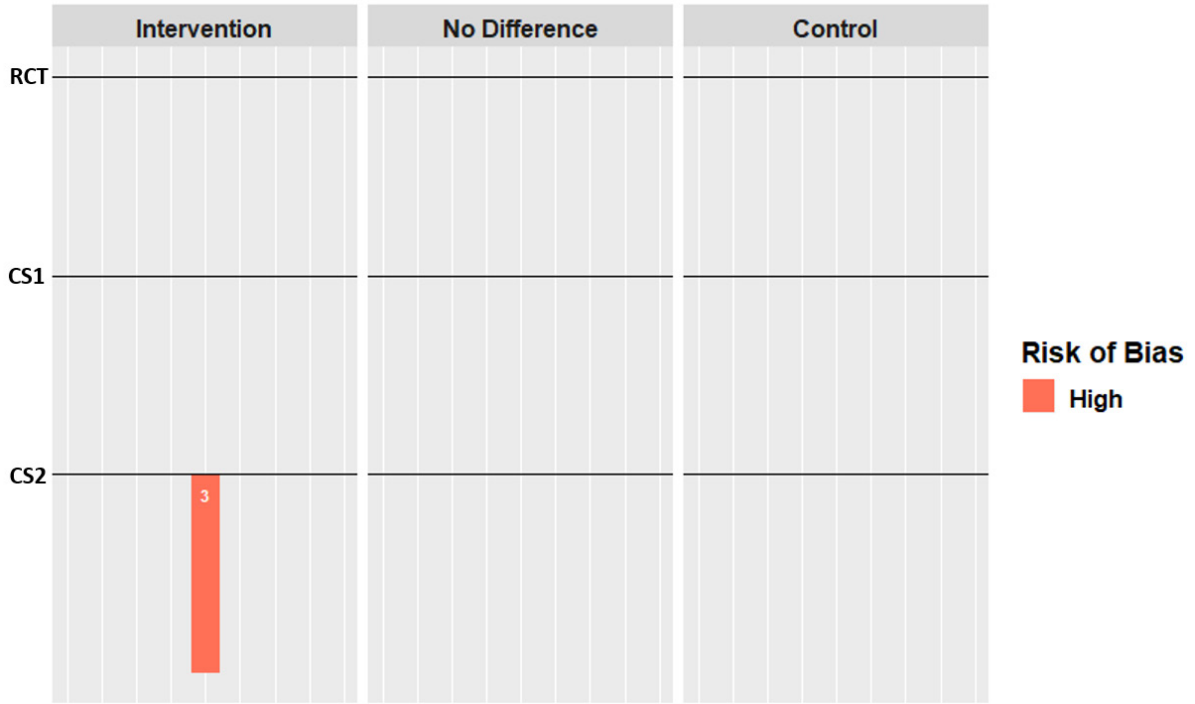


7.13 Appendix 13c. Referrals for maltreatment

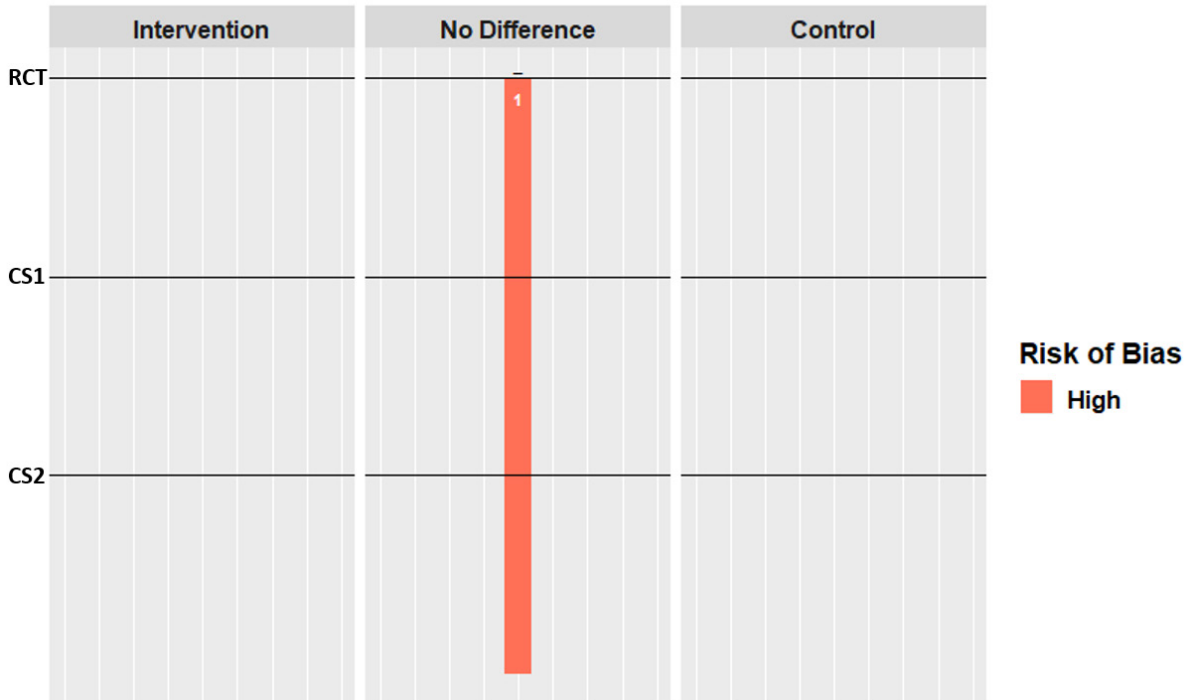




Do Family Team Conferences Result in Fewer Referrals for Child Maltreatment When Compared with Control Services?



Do Family Unity Meetings Result in Fewer Referrals for Child Maltreatment When Compared with Control Services?





What Works *for*
**Children's
Social Care**

wwcsc@nesta.org.uk

[@whatworksCSC](https://twitter.com/whatworksCSC)
whatworks-csc.org.uk

