Research article

The effectiveness of family group conferencing in youth care: A meta-analysis

CrossMark

Sharon Dijkstra (MSc)[a,*], Hanneke E. Creemers (PhD)[a], Jessica J. Asscher (PhD)[a], Maja Deković (PhD)[b], Geert Jan J.M. Stams (PhD)[a]

[a] *Forensic Child and Youth Care Sciences, University of Amsterdam, Nieuwe Achtergracht 127, 1018 WS, Amsterdam, The Netherlands*
[b] *Child and Adolescent Studies, Utrecht University, Heidelberglaan 1, 3584 CS, Utrecht, The Netherlands*

ARTICLE INFO

ABSTRACT

A meta-analytic study, involving 14 controlled studies ($N = 88495$ participants), was conducted to examine the effectiveness of Family Group Conferencing (FGC) in youth care. Child safety (in terms of reports of child maltreatment and out-of-home placement) and involvement of youth care were included as outcome variables; study, sample and intervention characteristics were included as moderators. Overall, FGC did not significantly reduce child maltreatment, out-of-home placements, and involvement of youth care. Study and sample characteristics moderated the effectiveness of FGC. Retrospective studies found FGC to be more effective than regular care in reducing the recurrence of maltreatment and decreasing the number and length of out-of-home placements, whereas prospective studies found FGC to be not more effective than regular care. Moreover, FGC was found to increase the number and length of out-of-home placements for families with older children and minority groups. The findings of this study showed that robust research proving effectiveness of FGC is limited. It is, therefore, crucial for the safety and protection of children in youth care that, before broadly implementing this decision making model in youth care, more robust studies examining the effectiveness of FGC be conducted.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

In the last decades, the model of Family Group Conferencing (FGC) has rapidly spread around the world. One reason for this may be that in several countries, such as New Zealand, Ireland, and the Netherlands, it is a legal requirement that all families in youth care should be offered the opportunity to make their own family group plan, for instance by carrying out a Family Group conference (FGc). However, despite its broad dissemination and implementation, the question is whether FGC is widely used because of the popularity of the philosophy of self-responsibility of citizens or because of its proven effectiveness (Maluccio, Ainsworth, & Thoburn, 2000). Studies examining effectiveness of FGC showed inconsistent results. Where some researchers reported positive results of FGC in terms of child maltreatment, out-of-home placement and involvement of professional

---

* Corresponding author.

E-mail addresses: s.dijkstra@uva.nl (S. Dijkstra), h.e.creemers@uva.nl (H.E. Creemers), j.j.asscher@uva.nl (J.J. Asscher), m.dekovic@uu.nl (M. Deković), g.j.j.m.stams@uva.nl (G.J.J.M. Stams).

youth care (Burford, Connoly, Morris, & Pennell, 2013; Crampton, 2007), other researchers found neutral or negative effects of FGC (Berzin, 2006; Lorentzen, 2009; Sundell & Vinnerljung, 2004). These inconsistent results could possibly be explained by differences in study design and sample characteristics. A *meta*-analysis, which synthesizes the results from different studies (Rothman, Greenland & Lash, 2008), seems, therefore, greatly needed in the field of FGC.

1.1. The model of FGC

FGC is a decision-making model that focuses on the family and its social network, and which aims to gather all parties with an interest in the wellbeing of a child and his or her family to make a family group plan that teaches and supports active responsibility (Merkel-Holguin, 1996). Family Group conferences (FGcs) were first used in New Zealand to resolve the overrepresentation of Maori children in out-of-home care (Levine, 2000). From New-Zealand, FGC spread out quickly to more than 20 countries, including US, Canada, Australia, Israel, United Kingdom, France, Sweden, Norway and the Netherlands (Nixon, Burford, Quinn, & Edelbaum, 2005).

The rapid growth of FGC around the world can be explained both by the philosophy underlying the model and its simplicity, which makes it attractive to apply in many different contexts and settings. FGC is based on the principle that the process of making decisions to keep children safe and healthy benefits from being democratized (Merkel-Holguin, 2004). It focuses on the strengths and resources of families that can be addressed to solve their problems and to take care of their children (Crampton, 2007; Graber, Keys, & White, 1996; Hudson, Galaway, Morris, & Maxwell, 1996). The assumption is that a plan developed and supported by the family and social network of a family is more likely to be carried out, and more successful too, than a plan developed by professionals (Burford & Hudson, 2000). An additional assumption is that families, by giving them the opportunity to make their own family group plan, feel that they have a voice in matters that concern them and are, therefore, more motivated to solve their problems (Merkel-Holguin, 2004). This is expected to result in improved collaboration with youth care agencies (Lupton & Nixon, 1999). Furthermore, FGC aim to bring family members together who have lost contact and, in this way, aim to provide new sources of support (Merkel-Holguin, 2004). It is assumed that a stronger social network reduces the need for professional care (Schuurman & Mulder, 2011). Moreover, it is assumed that when families discuss problems in private in their small inner circle the likelihood that sensitive information is brought up in the decision-making process will increase, which may improve the quality of the family plan (Crampton, 2007). Finally, because the model is thought to be culturally sensitive, the likelihood that plans will be carried out as intended is expected to increase (Merkel-Holguin, 2005). Based on all these assumptions, the model aims to improve child safety within families and to reduce the need of professional care.

Since the introduction of FGC in New Zealand many different variations of FGC have been developed, such as Family Unity Meeting (FUM), Family Decision Meeting (FDM), Team Decision making Meeting (TDM) and Family Team Meeting (FTM). These different types of FGC have many similarities. In general, three phases can be distinguished. In the first phase, the activation phase, a (independent) coordinator contacts stakeholders (family, important people for the family and involved professional care givers) and prepares the conference. In the second phase, the actual conference takes place, consisting of 1) an information phase; 2) a private part (depending on the type of FGC); and 3) the presentation of the plan. In the information phase, professionals share information on the needs and care options and provide, if necessary, conditions for the family group plan. During the private part, no professionals and coordinator are present, and the family with its social network develops the plan. In some types of FGC, professionals and coordinator participate during this part. During the presentation phase, the plan is tested by the professional to ensure that child safety is guaranteed. If this criterion is satisfied, the coordinator places the responsibility for implementation of the plan with the family and its social network, and subsequently the third (implementation) phase starts.

1.2. Present meta-analysis

As far as we know, no previous meta-analysis has been conducted examining the effectiveness of FGC in youth care. However, recently Havnen and Christiansen (2014) published a narrative (qualitative) research review of the effectiveness of FGC. They concluded that there is no clear answer whether FGC leads to less maltreatment and abuse, shorter or less out-of-home placements and reduced need for contact with child welfare services. Because of the narrative nature of this review, no statistical data were presented to support the conclusions.

In the present study, we therefore conducted a meta-analysis to determine the effectiveness of FGC. Based on the primary aims of the FGC model, we examined if FGC leads to improved child safety in terms of less child maltreatment and reduction of out-of-home placements and less involvement of youth care compared to regular care. An advantage of meta-analytic studies is that they allow for testing moderators, given the generally larger total sample size and in particular inclusion of factors that vary between and not within studies, such as the research design. By testing moderators, it is possible to examine whether for specific subgroups FGC is more likely to be effective. In the present meta-analysis, we distinguished three types of moderators: study characteristics (i.e., study quality and publication year), sample characteristics (i.e., ethnicity, age of children and socioeconomic status) and intervention characteristics, including the approach of FGC (i.e., private time during the conference or not).

In sum, we examined if 1) FGC, when compared to regular child care, leads to improved child safety in terms of less child maltreatment and reduction of out-of-home placements and less involvement of youth care and 2) whether study, sample

and intervention characteristics influence the effectiveness of FGC. This way, we aimed to determine which families, and under what circumstances, are most likely to benefit from FGC.

2. Method

2.1. Selection of studies

For this study, we searched for all available studies performed before June 2016 addressing effectiveness of FGC in youth care, meeting the following inclusion criteria: 1) studies had to focus on FGC in youth care; 2) studies had to include outcomes in terms of child maltreatment, out-of-home placement and duration or involvement of youth care; 3) studies had to compare an experimental group (participants receiving FGC) and a control group (participants receiving regular child care) and 4) statistics suitable for meta-analyses had to be available. For the search of eligible studies, we used the following procedure. First, electronic databases, such as Psychinfo, Google Scholar, Medline, Eric and dissertation abstract, were searched using the search string "family group decision making or family group conference* or family unity meeting or family decision making or team decision making or family team meeting or family meeting" in combination with "youth care or youth protect* or youth welfare or child care or child protect* or child welfare or foster care". For studies in German and French, search terms were translated in the specific language. Furthermore, we used the international database of the American Humane Organization, a database collecting studies about FGC, to find studies that were not available in the electronic database. In addition, reference lists of reviews and other articles were checked.

This search resulted in a list of 113 studies of which 17 studies met the inclusion criteria of our meta-analysis. Two of the 17 studies were excluded because the number of participants in the experimental and control group was not described (Crampton, Usher, Wildfire, Webster, & Cuccaro-Alamin, 2011; Wheeler & Johnson, 2003) and one study was excluded because duration of follow up was substantially longer for the control group, resulting in incomparable findings between the experimental and control group (Kiely & Bussey, 2001). Finally, we excluded one article (Berzin, Cohen, Thomas, & Dawson, 2008) because the results overlapped with the results reported in another study (Berzin, 2006).

Consequently, the final sample for our meta-analysis consisted of 14 studies, reporting on N = 88495 participants (including one study reporting on 80690 participants). These studies examined the effectiveness of FGC in 12 independent samples.

Of the 14 studies, nine examined the effects of FGC on child safety in terms of reports of child maltreatment, seven studies examined the effect of FGC on out-of-home placement, and four studies examined the effect of FGC on involvement of youth care. Improved child safety in terms of reports of child maltreatment was operationalized as substantiated reports (reports of child maltreatment supported by evidence after assessment; yielding 16 effect sizes) and non-substantiated reports (reports of child maltreatment without assessment of evidence; yielding two effect sizes). Out-of-home placement was operationalized as out-of-home placement yes or no (yielding three effect sizes), duration (yielding four effect sizes) and number of different placements (yielding three effect sizes). Finally, less involvement of youth care (case closure because of family stabilization) yielded six effect sizes. An overview of the study characteristics is presented in Table 1.

2.2. Coding the studies

Two researchers coded the included studies using a coding system for registering outcome variables and moderators. Responses of the two coders were compared and, if inconsistent, were discussed with a third researcher in order to reach consensus. Three categories of moderators were distinguished, namely study characteristics, sample characteristics and characteristics of the intervention. Study characteristics were study design (prospective designs including RCT, or retrospective designs in which comparison groups were retrospectively established), quality of the study, follow-up period (average in months from date FGc until follow-up date), if the effect size was controlled for pre-treatment scores (only applicable to child safety), whether FGC was offered to families in the control group and publication year of the study. The quality of the study was determined using the Quality Assessment Tool Quantitative Studies (Thomas, Ciliska, Dobbins, & Micucci, 2004). Using this checklist we calculated a quality score based on six study characteristics, namely sample selection, study design, identification and treatment of confounders, blinding of outcome assessors and of participants, reliability and validity of data collection methods and withdrawals and drop outs (scale 0–12; mean score = 4.56). In all of the included studies, families in the control group received regular care. However, in some studies the control group consisted of families who were offered an FGc, but declined and therefore received regular care, whereas in other studies no FGc was offered to the control group.

The following sample characteristics were coded: country where the study was conducted, sample size (small = <200; large = >200), mean-age of the child(ren) and proportion of minorities (i.e., African-America, Hispanic, Anglo, Moroccan, Turkish, Surinam, Antillean and Asian). Country was coded as a dichotomous moderator, divided into North America and other countries. For only four studies information on the number of single parent families was available, and for only two studies information on socioeconomic status. We therefore decided to exclude these characteristics as moderators.

The first characteristic of the intervention coded was availability of private time during the conference. Second, it was coded if FGC was used specifically for decision-making about out-of-home placement.

Table 1
Studies included in meta-analyses.

Study, sample and intervention characteristics								Outcome variables
Study	N	Type of FGC	Design	Quality	Year	Country	Placement	1; 2; 3
Berzin (2006)	164	FCG	RCT	3	2006	US	No	1, 3
Berzin (2006)	163	FU	RCT	3	2006	US	Yes	1, 2
Crampton (2007)	153	FGDM	QE – retrospective	5	2007	US	Yes	1, 2
Edwards, Tinworth, Burford and Pennell (2007)	680	FTM	QE – retrospective	2	2007	US	Yes	1
Pennell and Burford (2000)	63	FGDM	QE – retrospective	2	2000	CA	No	1, 2
Pennell et al. (2010)	594	FTM	QE – retrospective	2	2010	VS	Yes	2
Sheets et al. (2009)	4066	FGC	QE – retrospective	4	2009	VS	Yes	3
Sundell and Vinnerljung (2004)	239	FGC	QE – prospective	7	2004	SWE	No	1, 2, 3
Teal (2013)	755	TDM	QE – retrospective	6	2013	US	Yes	2
Titcomb and LeCroy (2005)	540	FGDM	QE – retrospective	2	2005	US	No	1
Wang et al. (2012)	80690	FGC	QE – retrospective	4	2012	US	Yes	2
Wingrove and Weisz (2005)	66	FGC	QE – retrospective	2	2005	US	Yes	3
Dijkstra, Creemers, Asscher and Stams (2016)	169	FGC	QE – prospective	7	2016	NL	No	1
Dijkstra et al. (2016)	140	FGC	QE – retrospective	7	2016	NL	No	1

Notes: Berzin-F = Berzin with sample of Fresno county; Berzin-R = Berzin with sample of Riverside county; FGC = Family Group Conferencing; FU = Family Unity; FGDM = Family Group Decision Making; FTM = Family Team Meeting; TDM = Team Decision Making; RCT = Randomized Controlled Trial; QE – retrospective = Quasi experimental, retrospective design; QE – prospective = Quasi experimental, prospective design; Quality = Quality score on scale 0–12; Placement = Using FGC for decision-making for out-of-home placement; 1 = reports of maltreatment; 2 = out-of-home placements; 3 = Involvement of youth care.

2.3. Calculating and analysis of effect sizes

Cohen's *d* was calculated, using formulas from Lipsey and Wilson (2001) and Mullen (1989) to establish the effectiveness of FGC in youth care on the basis of differences between families receiving FGC and families receiving regular care. Following the criteria of Cohen (1988), an effect size of *d* = 0.20 was considered small, an effect size of *d* = 0.50 was considered medium and an effect size of *d* = 0.80 was considered large. In most instances, Cohen's *d* was computed based on means, standard deviations, *t,* F, χ^2 or a one-tailed *p*-value. When studies did not provide one of these values, but only gave a proportion, χ^2 values were calculated to compute Cohen's *d.* Group differences were, where relevant, calculated both pre- and post-treatment. To account for initial group differences that may influence the post-treatment effects, pre-treatment scores were – when available – subtracted from post-treatment scores. Dichotomous dummy codes were made for the categorical variables and continuous moderators were centered.

In the present study, a multilevel approach to meta-analysis was used to account for dependency of effect sizes retrieved from the same primary study. Specifically, a three-level random effects model was used for calculating an overall effect as well as for performing moderator analyses (Van den Noortgate, López-López, Marín-Martínez, & Sánchez-Meca, 2013). This three-level meta-analytic model allows for three levels of variance, including the sampling variance of each effect size (level 1), the variance between effect sizes retrieved from the same study (level 2), and variance between studies (level 3) (Assink & Wibbelink, 2016). First, an overall effect was estimated using a model without moderators (i.e., an intercept-only model) for each of the outcome variables. To determine whether any variation in effect sizes within (level 2) and between (level 3) studies was significant, we conducted a separate log-likelihood test for each of these two levels. When there was significant variance on at least one of these levels, the distribution of effect sizes was considered to be heterogeneous. In that case, the three-level intercept-only models were extended by including (possible) moderators as covariates, so that moderating effects could be examined. An omnibus test was performed to determine the significance of included moderators (Houben, Van Den Noortgate, & Kuppens, 2015; Assink & Wibbelink, 2016). In each moderator analysis, moderators were only included in the model if 1) a minimum of three effect sizes for each moderator was available and 2) each category of the moderator was filled with at least one effect size.

2.4. Publication bias

When conducting a meta-analysis, it is important to consider publication bias, i.e. studies reporting no significant effects or negative effects are less likely to be published than studies reporting positive effects. This phenomenon was referred to by Rosenthal (1979) as the "file drawer problem". We examined publication bias in two ways. First, we calculated the fail-safe number. There is no indication of publication bias if the fail-safe number exceeds the critical value obtained with Rosenthal's

Table 2
Results for the overall mean effect size for each of the outcome variables.

Overall effect	NStudies[a]	NES	Nexp	Ncon	Mean d	t_0	F[b]
Reports of child maltreatment	9	18	815	1334	0.096	0.634	Level 2: 0.005 Level 3: 0.183***
Out-of-home placement	7	10	13282	69400	0.112	0.814	Level 2: 0.000 Level 3: 0.116**
Involvement of youth care	4	6	653	3795	−0.094	−0.535	Level 2: 0.000 Level 3: 0.098

Note: NStudies = number of studies; NES = number of effect sizes; Nexp = number of participants in experimental group; Ncon = number of participants in control group; Mean d = Mean effect size (d); t_0 = difference in mean d with zero; F = omnibus test.

[a] Because the three outcome variables were not examined in all the studies, the number of studies is reported in every table.

[b] Level 2 describe the variances of effect sizes within the same study; Level 3 describe the variances of effect sizes between studies.

** $p < 0.01$.

*** $p < 0.001$.

(1994) formula of $5 \times k + 10$ (k is the number of studies included in a multilevel meta-analysis). Second, as the Rosenthal's formula has recently been criticized (Assink et al., 2015), we also conducted the funnel-plot-bases trim and fill method (Duval & Tweedie, 2000a, 2000b), by using the "trimfill" package of Viechtbauer (2010) in R (R Development Core Team, 2016). With the funnel-plots, we tested if effect sizes were missing on both sides of the distribution. By imputing missing effects sizes, the trim and fill method restores the symmetry of an asymmetric funnel plot. After imputing the missing effect sizes, the meta-analyses were repeated to compute an overall effect size that would take the influence of publication bias into account.

3. Results

3.1. Overall effect sizes

The overall mean effect sizes were not significant, which indicates that FGC is not more effective than regular care with regard to less child maltreatment ($d = 0.096$), reduction of out-of-home placement ($d = 0.112$) and less involvement of youth care ($d = -0.094$) (Table 2). To correct the meta-analysis on out-of-home placement for the large number of participants in the study by Wang et al. (2012), it was repeated excluding this study, yielding a comparable result ($d = 0.138$, $p > 0.05$).

The omnibus test revealed that there was no heterogeneity for involvement of youth care (Level 2: $t = 0.000$, $p = 1.000$; Level 3: $t = 0.098$, $p = 0.144$). Therefore, no moderator analyses were conducted for this outcome measure. For reports of child maltreatment and out-of-home placements, heterogeneity tests revealed that there was heterogeneity in the effect sizes (respectively Level 3: $t = 0.183$, $p < 0.001$; Level 3: $t = 0.116$, $p < 0.01$). For these outcome variables, moderator analyses were conducted to determine if effects for FGC depended on study, sample or intervention characteristics.

3.2. Publication bias

For reports of child maltreatment, the fail-safe number was 426, which exceeds the critical value of 55 ($5 \times 9 + 10$). For out-of-home placement, the fail-safe number was 227, which exceeds the critical value of 45 ($5 \times 7 + 10$). Finally, the fail-safe number for involvement of youth care was 63, which exceeds the critical value of 30 ($5 \times 4 + 10$). Based on Rosenthal's formula, publication bias is, therefore, unlikely. Possible publication bias was also examined with funnel-plots. These plots and corresponding analyses showed that publication bias was present in the outcome factors 'reports of child maltreatment' and 'involvement of youth care', as indicated by an asymmetrical distribution of effect sizes on the right side of the funnel plot, which suggests that the overall effect size is underestimated. Therefore, 'corrected' overall effects were calculated for these outcome factors. Results of the new analyses were comparable to the original results for both reports of child maltreatment ($d = 0.144$, $p > 0.05$) as well as involvement of youth care ($d = 0.125$, $p > 0.05$). Therefore, it seems that publication bias did not influence the results of the meta-analyses. The funnel-plots are depicted in Figs. 1 and 2.

3.3. Moderator analyses for child maltreatment

Moderator analyses showed that the effect of FGC was moderated by one study characteristic, namely study design. Studies using a retrospective design showed positive results of FGC in terms of less reports of child maltreatment, whereas studies with a prospective design showed no differences in effectiveness. No moderating effect was found for substantiated versus non-substantiated reports, which means that type of report (reports with evidence after assessment versus reports without assessment of evidence) did not affect the effectiveness of FGC when compared to regular care. No sample or intervention characteristics moderated the effectiveness of FGC in terms of child maltreatment (Table 3).
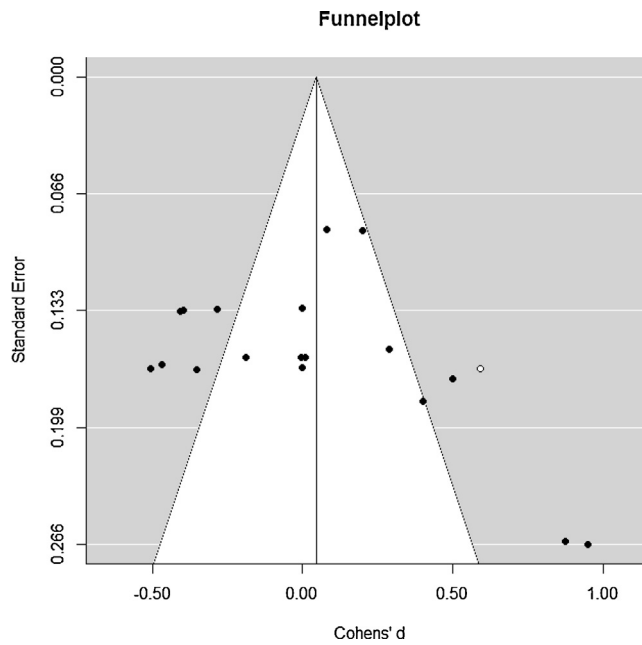
Funnelplot

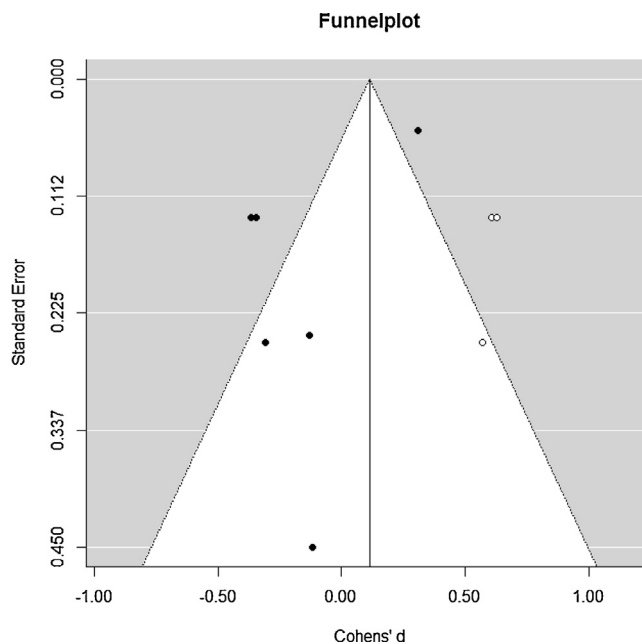Fig. 1. Funnel plot of the outcome 'Reports of child maltreatment'.

Funnelplot

Fig. 2. Funnel plot of the outcome 'Involvement of youth care'.

3.4. Moderator analyses for out-of-home placement

Moderator analyses showed that the effect of FGC on reduction of out-of-home placements was moderated by one study characteristic, namely study design. Studies with a retrospective design showed that FGC leads to reduction of out-of-home placements, whereas studies with a prospective design showed no differences in effects between the two groups. Two sample characteristics, namely mean age of the child and proportion of minorities, moderated the effect of FGC. These moderator effects indicate that FGC leads to more out-of-home placements for families with older children and for minorities. No moderating effect was found for measure of out-of-home placements, distinguishing between prevalence of out-of-home placements (yes or no) and duration or number of placements, indicating that, irrespective of measure, FGC did not

Table 3
Results of the moderator analyses for reports of child maltreatment.

Moderator Variables	NStudies	EES	β_0,Mean d	t_0	β_1, RC	t_1	F(df$_1$, df$_2$)
Outcome variable							
Reports of child maltreatment							F(1,16) = 0.639
Substantiated report (RC)	9	16	0.106	0.677			
Non-substantiated report	2	2	−0.015	0.070	−0.120	0.799	
Study characteristics							
Study design							F(1,16) = 5.675*
Retrospective (RC)	5	9	0.348	2.146*			
Prospective	4	9	−0.231	1.275	−0.579	2.382*	
Study quality	9	18	0.102	0.624	0.019	0.266	F(1,16) = 0.071
Follow up period	6	15	0.183	1.031	−0.012	0.889	F(1,13) = 0.791
Corrected for pre-treatment effects							F(1,16) = 0.012
Corrected (RC)	8	15	0.085	0.458			
Not-corrected	2	3	0.103	0.623	0.017	0.110	
FGC offered in control group							F(1,11) = 0.999
Offered (RC)	1	1	0.400	0.953			
Not-offered	6	12	−0.050	0.307	−0.450	1.000	
Publication year	9	18	0.083	0.489	0.010	0.293	F(1,16) = 0.086
Sample characteristics							
Country							F(1,16) = 0.082
North America (RC)	6	10	0.064	0.322			
Other countries	3	8	0.163	0.582	0.099	0.287	
Sample size							F(1,16) = 0.285
Small <200 (RC)	7	10	0.146	0.791			
Large >200	2	8	−0.054	0.166	−0.200	0.534	
Mean age child	5	10	0.030	0.148	0.129	1.439	F(1,8) = 2.070
% Minorities	6	11	0.007	0.032	−0.002	0.217	F(1,9) = 0.047
Intervention characteristics							
Availability of private time							F(1,8) = 2.365
Private time (RC)	4	5	0.321	1.523			
No private time	2	4	−0.274	0.844	−0.595	1.538	
FGC as decision-making for out-of-home placement							F(1,16) = 0.487
Yes (RC)	3	5	−0.060	0.218			
No	6	13	0.174	0.904	0.234	0.698	

Notes. NStudies = number of studies; NES = number of effect sizes; β_0 Mean d = Mean effect size (d); t_0 = difference in mean d with zero; β_1, RC = Regression coefficient; t_1 = difference in mean d with regression category; F(df$_1$, df$_2$) = omnibus test; (RC) = reference category.
  * $p < 0.05$.

outperform regular care in terms of out-of-home placements. No other study and sample characteristics and no intervention characteristics moderated the effectiveness of FGC in terms of out-of-home placement (Table 4).

4. Discussion

4.1. Summary and discussion of results

The aim of the present study was to examine the effectiveness of FGC in youth care through meta-analyses. A total of 14 controlled studies (consisting of N = 88495 participants) and 34 effect sizes were included. The findings of the present study indicate that, overall, FGC does not outperform regular care in terms of less child maltreatment, reduction of out-of-home placements and less involvement of youth care. However, moderator-analyses indicate that in studies using a retrospective design, FGC leads to less reports of child maltreatment and reduction of out-of-home placements when compared to regular care, whereas this effect was not present in studies using prospective research designs. Furthermore, the percentage of minorities moderated the effect of FGC on out-of-home placements; FGC resulted in more and longer out-of-home placements in minority groups.

An explanation for the fact that FGC leads to more out-of-home placements in minority groups may be found in cultural differences. It has been argued that shame about problems in the family context within minority groups decreases the likelihood to share sensitive information during a conference (Futa, Hsu, & Hansen, 2001), which may result in a lower quality of the conference and the family group plan (Crampton, 2007), resulting in more or longer out-of-home placements. However, one may argue that this would also result in more reports of child maltreatment, which was not found in this study. More research is needed to examine differences in the effectiveness of FGC in families with various ethnic backgrounds.

FGC resulted in more and longer out-of-home placements in families with older children. Differences in the nature of problems experienced by families with younger versus older children may explain this result. That is, families in youth care with older children often are referred because of child behavioral problems, whereas in families with younger children the main reason for involvement with youth care is often severe child rearing problems (Knorth, 1995; Vinnerljung & Sallnäs,

Table 4
Results of the moderator analyses for out-of-home placement.

Moderator variables	N Studies	NES	β_0, Mean d	t_0	β_1	t_1	$F(df_1, df_2)$
Outcome variable							
Out-of-home placements							$F(2,7) = 0.574$
Yes or no (RC)	2	3	0.020	0.110			
Duration	4	4	0.160	1.118	0.140	0.938	
Number of placements	3	3	0.103	0.675	0.083	0.476	
Study characteristics							
Study design							$F(1,8) = 7.299^*$
Retrospective (RC)	5	6	0.268	2.321*			
Prospective	2	4	−0.314	1.727	−0.058	2.702*	
Study quality	7	10	0.046	0.355	−1.040	1.540	$F(1,8) = 2.370$
Follow up period	5	8	0.100	0.586	−0.001	−0.176	$F(1,6) = 0.031$
Corrected for pre-treatment effects							$F(1,8) = 0.207$
Not corrected (RC)	6	9	0.090	0.588			
Corrected	1	1	0.303	0.687	0.212	0.455	
FGC offered in control group							$F(1,6) = 1.472$
Not offered (RC)	3	6	−0.148	0.821			
Offered	2	2	0.202	0.897	0.350	1.213	
Publication year	7	10	0.105	0.713	0.015	0.432	$F(1,8) = 0.187$
Sample characteristics							
Country							$F(1,8) = 4.438$
North America (RC)	6	7	0.206	1.729			
Other countries	1	3	−0.418	1.541	−0.624	2.107	
Sample size							$F(1,8) = 0.195$
Small < 200 (RC)	4	7	0.198	0.821			
Large > 200	3	3	0.064	0.341	−0.135	0.442	
Mean age child	4	7	−0.103	0,955	0.001	3.171*	$F(1,8) = 10.057^*$
% Minorities	6	9	−0.018	0.310	0.012	5.287**	$F(1,7) = 27.948^{**}$
Intervention characteristics							
Availability of private time							$F(1,5) = 0.019$
No private time (RC)	3	4	0.191	1.043			
Private time	3	3	0.228	1.157	0.037	0.136	
FGC as decision-making for out-of-home placement							$F(1,8) = 0.300$
No (RC)	2	4	−0.134	0.522			
Yes	5	6	0.197	1.293	0.330	1.108	

Notes: NStudies = number of studies; NES = number of effect sizes; β_0, Mean d = Mean effect size (d); t_0 = difference in mean d with zero; β_1, RC = Regression coefficient; t_1 = difference in mean d with regression category; $F(df_1, df_2)$ = omnibus test; (RC) = reference category.

  * $p < 0.05$.
  ** $p < 0.01$.

2008; Wouters, 2005). Possibly FGC, developed to increase family functioning by focusing on the strengths and resources of families and parents (Crampton, 2007; Graber et al., 1996; Hudson et al., 1996), is less effective than regular care in terms of reducing family problems associated with adolescent behavioral problems. Moreover, during adolescence, children become more autonomous and independent from their parents (Steinberg, 1990). Perhaps, a reduced influence of parents and network during this developmental period makes it more difficult to reduce adolescent behavior problems through FGC. This may result in more out-of-home placements following FGC in families with older children.

4.2. Number and quality of the included studies

The current study provided also insights in the research into the effectiveness of FGC. First, the number of studies included in the present meta-analyses is limited. Only 14 controlled studies, providing sufficient information to conduct the meta-analyses, were available when focusing on the effectiveness of FGC in youth care in terms of child maltreatment, out-of-home placement and involvement of youth care. To control for publication bias, analyses were conducted. Results showed that in none of the outcomes, an indication of publication bias was present. Thus, despite the widespread implementation of FGC and the large number of studies examining FGC, a limited number of (controlled) effectiveness studies has been conducted. The large majority of studies into FGC has focused on implementation and participant satisfaction, or evaluated FGC without a control group. In such studies, over time changes cannot be attributed to FGC.

Second, the weak quality of the included studies, as determined by the *Quality Assessment Tool for Quantitative Studies* (Thomas et al., 2004), was remarkable. The main reason for the low quality scores is the retrospective design that was used in most of the studies. In these studies, all information was obtained from file data, which were often incomplete, and obtained without valid instruments. In addition, a great risk of retrospective studies is (self) selection bias because assignment to the experimental and control group is determined retrospectively. Therefore, comparability of the experimental and control group is extremely difficult to establish, and in most cases remains unclear.

Furthermore, generally, limited information was available on the fidelity of FGC in the studies included in the meta-analysis. More information on program fidelity would have enabled the assessment of the degree to which the level of program fidelity may influence the effectiveness of FGC. In addition, it is important to have information on program fidelity, as this helps determine whether lack of positive effects of FGC on child safety, out-of-home placements and involvement of youth care is due to low levels of program fidelity.

4.3. Limitations of the study

There are some limitations of the current study that should be mentioned. First, as a consequence of the small number of included studies and effect sizes, and a lack of information on sample characteristics, we were restricted in the possibilities to perform moderator analyses. Second, the lack of information on program fidelity of FGC in the included studies makes it difficult to compare the effects across studies, since it is unclear to what extent the studies examined a comparable approach. However, although it is unclear which exact method was used to develop a family group plan, it is clear that a plan was made by the family and the extended network, which imposes active responsibility on them to carry out the plan as intended (Merkel-Holguin, 1996).

Furthermore, as described earlier, the current meta-analysis included one study with a considerably larger sample size (Wang et al., 2012) compared to the other included studies. This study, focusing on out-of-home placement, may therefore have unduly influenced the overall effect for this outcome measure. To correct for the impact of this study on the effectiveness in terms of out-of-home placement, we repeated this meta-analysis excluding the study by Wang et al. (2012). This yielded comparable results, indicating that this particular study did not excessively influence the results. In addition, both the study by Wang et al. (2012) and this meta-analysis resulted in a non-significant effect of FGC on out-of-home placement.

Finally, although a three-level meta-analysis is an innovative and appealing approach, it still has some limitations. For example, it has been argued that studies yielding relatively many effect sizes may excessively contribute to the meta-analytic results. However, multi-level meta-analysis produces virtually the same overall effect sizes as standard meta-analysis (Hox, 2010), while both within and between study variance is accounted for (Assink & Wibbelink, 2016). Furthermore, as a consequence of the limited research on the effectiveness of FGC, the number of studies and included effect sizes is low. It has been stated that the statistical power in the analyses increases as the number of studies and effect sizes increases (Assink & Wibbelink, 2016). The low number of studies may lead to the problem of a biased estimate of the between-study variance and standard error (Van den Noortgate et al., 2013). However, since methods for determining the exact power in multi-level meta-analytic models are not yet available (Assink & Wibbelink, 2016), it was not possible to examine if the current meta-analysis was underpowered. Therefore, when interpreting the results, we should be aware of this limitation.

4.4. Practical implications

Despite these limitations, this study is, to our knowledge, the first meta-analysis examining the effectiveness of FGC in youth care and provides, therefore, a first overview of the effectiveness of FGC. The results of this study have some practical and policy implications. First, FGC does not outperform regular care in terms of reducing child maltreatment, out-of-home placements and involvement of youth care. For the organizers of FGCs, it is important to determine how FGC can add to the existing youth care. Most important implication for policy makers is that they need to determine whether the costs of FGC outweigh its benefits. In New Zealand, Ireland, and most recently the Netherlands, it is imposed by law that all families in youth care should be offered the possibility to make a family group plan, for instance by carrying out a family group conference. However, given the absence of evidence on the superiority of FGC in youth care, it seems premature to impose its implementation by law.

5. Conclusion

This study showed that there were no significant overall effects of FGC on child maltreatment, out-of-home placements and involvement of youth care. However, studies using a less robust design showed positive outcomes for FGC on child maltreatment and out-of-home placement, whereas these effects were not present in studies with a more robust design. Additionally, the FGC model was less effective for minority groups and families with older children in reducing out-of-home placements.

Robust research into the effectiveness of FGC is so limited that it does not support its widespread use. Only 14 controlled studies could be included in the meta-analysis, most of which were retrospective and provided limited information about the composition of the sample. It is crucial for the safety and protection of children in youth care that more effectiveness studies into FGC will be conducted, meeting the quality criteria, including model fidelity, and with sufficient power to be able to draw conclusions. Only empirical research that allows for conclusions on causality can determine the effectiveness of FGC and can identify which families, and under what circumstances, benefit most from FGC.

Funding

Acknowledgements

References

Assink, M., & Wibbelink, C. J. M. (2016). Fitting three-level meta-analytic models in R: A step-by-step tutorial. *The Quantitative Methods for Psychology, 12,* 154–174.

Assink, M., van der Put, C. E., Hoeve, M., de Vries, S. L., Stams, G. J. J., & Oort, F. J. (2015). Risk factors for persistent delinquent behavior among juveniles: A meta-analytic review. *Clinical Psychology Review, 42,* 47–61.

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. *Child Welfare, 87*(4), 35–54.

Berzin, S. C. (2006). Using sibling data to understand the impact of family group decision-making on child welfare outcomes. *Children and Youth Services Review, 28*(12), 1449–1458.

Burford, G., & Hudson, J. (2000). General introduction: Family group conference programming. In *Family group conferencing: New directions in community-centered child & family practice.* pp. xix–xxvii. New York: Aldine de Gruyter.

Burford, G., Connolly, M., Morris, K., & Pennell, J. (2013). *Introduction, principles, and processes. Part of a larger review of family group conferencing research.* [Retrieved from:]. http://www.eigen-kracht.nl

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Crampton, D. S., Usher, C. L., Wildfire, 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. *Child Welfare, 90*(4), 61–78.

Crampton, D. (2007). Research review: Family group decision-making: A promising practice in need of more programme theory and research. *Child & Family Social Work, 12*(2), 202–209.

Duval, S., & Tweedie, R. (2000a). A nonparametric trim and fill method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95*(449), 89–98.

Duval, S., & Tweedie, R. (2000b). Trim and fill: A simple funnel-plot?based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2), 455–463.

Futa, K. T., Hsu, E., & Hansen, D. J. (2001). Child sexual abuse in asian american families: An examination of cultural factors that influence prevalence, identification, and treatment. *Clinical Psychology: Science and Practice, 8*(2), 189–209.

Graber, L., Keys, T., & White, J. (1996). Family group decision-making in the united states: The case of oregon. In *Family group conferences: Perspectives on policy and practice.* pp. 180–194.

Havnen, K. J. S., & Christiansen, Ø. (2014). *Knowledge review on Family Group Conferencing: Experiences and outcomes.* Norway: Regional Centre for Child and Youth Mental Health and Child Welfare, Uni Research Health.

Houben, M., Van den Noortgate, W., & Kuppens, P. (2015). The relation between short-term emotion dynamics and psychological well-being: A meta-analysis. *Psychological Bulletin, 141,* 901.

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications.* New York, NY: Routledge.

Hudson, J., Galaway, B., Morris, A., & Maxwell, G. (1996). Research on family group conferencing in child welfare in new zealand. In *Family group conferences: Perspectives on policy and practice.* pp. 533–552.

Kiely, P., & Bussey, K. (2001). *Family group conferencing: A longitudinal evaluation.* Sydney, Australia: Macquarie university.

Knorth, E. J. (1995). Besluitvorming over uithuisplaatsing in de jeugdzorg. *Kind En Adolescent, 16*(2), 45–59.

Levine, M. (2000). The family group conference in the new zealand children, young persons, and their families act of 1989 (CYP&F): Review and evaluation. *Behavioral Sciences & the Law, 18*(4), 517–556.

Lipsey, M. W., & Wilson, D. B. (2001). The way in which intervention studies have personality and why it is important to meta-analysis. *Evaluation & the Health Professions, 24*(3), 236–254.

Lorentzen, B.L., 2009. Dissertation Abstracts International PhD Thesis. Effects of Family Group Decision Making in a Voluntary Family Maintenance Program, [Retrieved from:] http://www.americanhumane.org/children/programs/family-group-decision-making/bibliographies/research-and-evaluation/effects-of-family-group.html.

Lupton, C., & Nixon, P. (1999). *Empowering practice?: A critical appraisal of the family group conference approach.* Policy Press.

Maluccio, A. N., Ainsworth, F., & Thoburn, J. (2000). *Child welfare outcome research in the United States, the United Kingdom, and Australia.* Child Welfare League of America.

Merkel-Holguin, L. (1996). Putting families back into the child protection partnership: Family group decision making. *Protecting Children, 12*(3), 4–7.

Merkel-Holguin, L. (2004). Sharing power with the people: Family group conferencing as a democratic experiment. *Journal of Sociology & Social Welfare, 31,* 155.

Merkel-Holguin, L. (2005). *The intersection between family group decision making and systems of care.* [Retrieved from American Humane Association Website]. http://www.americanhuman.org/site/DocServer/FGDM_Brief_8.pdf?docID=32212005

Mullen, B. (1989). *Advanced basic meta-analysis.* Hillsdale, NJ: Erlbaum.

Nixon, P., Burford, G., Quinn, A., & Edelbaum, J. (2005). *A survey of international practices, policy & research on family group conferencing and related practices.* London, England: Family Rights Group [http://Frg.Org.Uk/Pdfs/Family%20Group%20practices%20report.Pdf (accessed 06.01.10)]

R Development Core Team. (2016). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing [Retrieved from]. http://www.R-project.org/

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*(3), 638.

Rosenthal, R. (1994). Statistically describing and combining studies. In *The handbook of research synthesis.* pp. 231–244.

Rothman, K. J., Greenland, S., & Lash, T. L. (Eds.). (2008). Modern epidemiology. In. Lippincott Williams & Wilkins.

Schuurman, M., & Mulder, C. (2011). *Eigen kracht-conferenties bij gezinnen in de regio amsterdam wat levert het op?*

Steinberg, L. (1990). Autonomy, conflict and harmony in the family relationship. In S. S. Feldman, & G. R. Elliott (Eds.), *At the threshold: The developing adolescent* (pp. 255–276). Cambridge, MA, US: Harvard University Press.

Sundell, K., & Vinnerljung, B. (2004). Outcomes of family group conferencing in sweden: A 3-year follow-up. *Child Abuse & Neglect, 28*(3), 267–287.

Thomas, B., Ciliska, D., Dobbins, M., & Micucci, S. (2004). A process for systematically reviewing the literature: Providing the research evidence for public health nursing interventions. *Worldviews on Evidence Based Nursing, 1*(3), 176–184.

Titcomb, A., & LeCroy, C. (2005). Outcomes of arizona's family group decision making program. *Protecting Children*, *19*(4), 47–53.
Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2013). Three-level meta-analysis of dependent effect sizes. *Behavior Research Methods*, *45*(2), 576–594.
Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48.
Vinnerljung, B., & Sallnäs, M. (2008). Into adulthood: A follow-up study of 718 young people who were placed in out of home care during their teens. *Child & Family Social Work*, *13*(2), 144–155.
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. *Children and Youth Services Review*, *34*(4), 845–850.
Wheeler, C., & Johnson, S. (2003). Evaluating family group decision making: The santa clara example. *Protecting Children*, *18*(1–2), 65–69.
Wouters, J. (2005). Uithuisplaatsing in Nederland en Zweden. In *Een vergelijkend onderzoek naar overwegingen in de jeugdzorg*. Utrecht: Universiteit Utrecht.

References to articles solely included in meta-analyses

Dijkstra, S., Creemers, H. E., Asscher, J. J., & Stams, G. J. J. M. (2016). *De inzet van familienetwerkberaden in de jeugdbescherming*. Den Haag: WODC.
Edwards, M., Tinworth, K., Burford, G., & Pennell, J. (2007). *Family team meeting (FTM) process, outcome, and impact evaluation: Phase II report*. Englewood, CO: American Humane Association.
Pennell, J., & Burford, G. (2000). Family group decision making: Protecting children and women. *Child Welfare-New York*, *79*(2), 131–158.
Pennell, J., Edwards, M., & Burford, G. (2010). Expedited family group engagement and child permanency. *Children and Youth Services Review*, *32*(7), 1012–1019.
Sheets, J., Wittenstrom, K., Fong, R., James, J., Tecci, M., Baumann, D. J., et al. (2009). Evidence-based practice in family group decision-making for anglo ispani, ispanic and ispanic families. *Children and Youth Services Review*, *31*(11), 1187–1191.
Teal, B. L. (2013). *The effectiveness of Team Decision Making in children's division on child welfare outcomes. (Docteral dissertation)*. Capella University.
Wingrove, T. A., & Weisz, V. (2005). An evaluation of expedited family group conferences. In *Presented at american law-Psychology society annual conference*.