Title

Adjustment to unmet parenthood goals: Systematic review of long-term adjustment after failed fertility treatment

Running title

Adjustment to unmet parenthood goals

Sofia Gameiro

Cardiff Fertility Studies Research Group, School of Psychology, Cardiff University

Amy Finnigan

School of Medicine, Cardiff University

Address for correspondence:

Sofia Gameiro, Cardiff Fertility Studies Research Group, School of Psychology, Cardiff University, Tower Building, Park Place, Cardiff, Wales CF10 3AT, UK. Email: GameiroS@cardiff.ac.uk

Table of contents

Introduction
Methods
    Systematic search
    Inclusion and exclusion criteria
    Data extraction
    Data synthesis and analysis
    Risk of bias assessment
Results
    Description of studies
    Quantitative meta-synthesis: Psychosocial adjustment after failed fertility treatment
    Qualitative meta-synthesis: Patients' perceived experiences of the period after failed treatment
Discussion
    Strengths and limitations
Conclusions

Abstract

Background: Fertility treatment is not guaranteed to succeed and around 30% of patients do not achieve parenthood. Failed treatment represents the loss of parenthood and often triggers intense and prolonged grief reactions. There is an increasing awareness of the need to support patients in the aftermath of failed treatment, however there are no effective interventions for the effect. This need for support is expected to increase as the number of people delaying parenthood increases and therefore so does the number for whom assisted reproduction will not offer a solution.

Objective and rationale: This mixed-methods review aims at investigating if patients who undergo failed fertility treatment experience significant psychosocial adjustment difficulties that warrant the provision of psychosocial support, and at developing a comprehensive model of adjustment to unmet parenthood goals that can be used to assist the design of theory led psychosocial interventions.

Search methods: Five databases were systematic searched between 1978 and December 2015. Search terms included fertility treatment AND psychosocial adjustment AND post-treatment. Quantitative studies had to include group mean comparisons on psychosocial adjustment (mental-health, wellbeing) between patients who did failed treatment and a control group. Qualitative studies had to focus on experiences of psychosocial adjustment after failed treatment. Screening, data extraction and critical appraisal were done independently by the authors using pre-defined protocols. Two meta-analyses were performed on mental-health and wellbeing with a random effect model. Primary outcome was Hedge's g. Publication bias was checked with visual inspection of funnel plots, Egger's test and the trim-and-fill method. A 3-stage thematic analysis of results reported in primary qualitative papers was implemented. First-order descriptive and second-order interpretative themes were extracted.

Outcomes: Nine quantitative (9052 individuals, 8 countries) and 9 qualitative (267 individuals, 6 countries) studies were included. Six (67%) of the quantitative studies reported on mental-health and 7 (78%) on wellbeing. The meta-analyses showed that the failed group had worse mental-health (g = -0.450, P=0.002, 95%CI [-0.734 -0.267]; I^2=85%, P<0.001) and wellbeing (g = -0.319, P<0.001, 95%CI [-0.439 -0.198], I^2=45%, P=0.001) than controls. The qualitative review resulted in 28 first-order themes that were grouped into 6 second-order themes: individual and relational adjustment, social adjustment, acceptance, pursuit of new life goals, meaning making, and fertility care perceptions and needs. The data showed that individual, relational and social adjustment tended to increase with time since treatment, and that individuals' care perceptions and needs also changed. The data also suggested that individuals who engage in the psychological tasks of accepting and making meaning of their situation and pursuing new life goals adjust better and have fewer support needs. These

predictions were articulated in the Three Tasks Model of Adjustment to Unmet Parenthood Goals.

Wider implications: Results provide compelling evidence for the provision of psychosocial care directed at helping individuals relinquishing their parenthood goals. The model developed offers comprehensive guidance on the therapeutic mechanisms that psychosocial care should target to promote adjustment. Future research should test the model with prospective cohorts studies or by developing and testing interventions based on its predictions.

Key-words: parenthood; infertility; childlessness; psychosocial adjustment; mental-health; wellbeing; psychological care; mixed-methods; systematic review; meta-analysis.

Introduction

Around 10% of individuals experience infertility and half of these undergo fertility treatments such as intra-uterine insemination (IUI) or In-Vitro Fertilization (IVF) to conceive (Boivin et al. 2007), (Datta et al. 2016). For most people fertility treatment represents their only chance to achieve their parenthood goals, that is, to have the number of children they desire. However, treatment is not guaranteed to succeed and around 30% of patients do not achieve parenthood or end up having less children than they would like to (Pinborg et al. 2009, Troude et al. 2016). For these people the remaining options are to forgo their parenthood goals or apply for adoption. Therefore, for thousands of people failed fertility treatment represents the loss of biological parenthood, or parenthood in general, and usually triggers intense and prolonged grief reactions (Daniluk 2001, Volgsten et al. 2010). There is an increasing awareness of the need to support individuals in the aftermath of failed fertility treatment and of the preventive role fertility clinics can play in this context (Gameiro et al. 2013). However, the recently published guidelines of the European Society for Human Reproduction and Embryology (ESHRE) could not list a single evidence-based intervention with this aim, suggesting that patients' support needs are not being addressed (Gameiro et al. 2015). This support need is expected to increase as the number of people delaying parenthood increases and therefore so does the number for whom assisted reproduction will not offer a solution (Leridon 2004). The present mixed-methods review aims to investigate if patients who undergo failed fertility treatment experience significant psychosocial adjustment difficulties that warrant the provision of psychosocial care and at developing a comprehensive model of adjustment to unmet parenthood goals that can be used to guide the design of psychosocial interventions.

Despite the call for psychosocial care in the aftermath of treatment (Gameiro et al. 2015), it is still unclear if individuals who do failed treatment adjust significantly worse than those for whom treatment is successful. Identity theories (Thoits 1992) would suggest this to be the case because failed treatment represents the loss of parenthood. For some it may only represent the loss of biological parenthood, for others the complete loss of parenthood, and for others yet the impossibility to have more children. Despite the significant differences in these experiences, in this article and we will only refer generically to the loss of parenthood or parenthood goals.

Research has shown that undesired childlessness is associated with poorer psychosocial adjustment (e.g., Graham 2015) and those using fertility treatment can be considered especially at risk for maladjustment because of their high commitment to become parents. Indeed, 84% of patients undergo all treatment cycles recommended by their physician (Gameiro et al. 2013) despite treatment being emotionally and physically demanding and disruptive of personal and professional routines. During the last decade several studies

investigated adjustment after failed treatment but so far the evidence is mixed. Some studies show impairments in mental-health and wellbeing for patients who experienced failed treatment (e.g., Kuivasaari-Pirinen et al. 2014, Yli-Kuha et al. 2010) but others do not (Hammarberg et al. 2001, Wischmann et al. 2012). This inconsistency may be partially explained by the conceptual and methodological heterogeneity observed.

First, different comparison groups were used. Researchers either compared patients who ended successful versus failed treatment (treatment outcome), or patients with children versus without children at the end of treatment (parenthood status). While the first approach is based on the assumption that it is the non-realization of the individuals' parenthood goals that hinders adjustment (less children than desired), the second is based on the assumption that it is their parenthood status (no children at all). These two factors (unmet parenthood goals, parenthood status) are difficult to disentangle and there is evidence to suggest that they interact, whereby the inability to have a(nother) desired child results in higher suffering for childless than non-childless individuals (Gameiro et al. 2016). Second, sometimes the failed treatment group included individuals who already had children before treatment or who adopted or conceived spontaneously after treatment and this may have attenuated group differences in adjustment. Third, the follow-up assessment period varied from immediately after treatment to up to 20 years after, which may have created mixed results because loss reactions are expected to attenuate with time (Bonanno 2004). A meta-analysis taking these potential moderators into account can clarify on the severity of adjustment difficulties experienced after failed fertility treatment.

Another important issue is to understand how to support people in the aftermath of failed fertility treatment. So far researchers focused on understanding short-term reactions to failed treatment cycles (e.g., Berghuis and Stanton 2002, Terry and Hynes 1998). The study of long-term (i.e., more than 1 year after) adjustment to unmet parenthood goals is still an emergent research topic within the field that has mostly been investigated with qualitative research methods aiming to describe patients' experiences and needs during the post-treatment period. To date only a few studies focused on identifying risk factors or exploring the mechanisms of adjustment after failed treatment (e.g., Verhaak et al. 2007), despite the existence of multiple psychological theories that can be used for the effect. In the following paragraphs we briefly describe the theoretical paradigms that have been used so far and related evidence.

Stress and coping theory states that adjustment depends on the fit between the stressor experienced and the strategies used to address it (Lazarus and Folkman 1984). Stanton and colleagues (1994) hypothesize that when stress results from a significant loss the use of emotional approach coping strategies, such as efforts to acknowledge, understand and express emotions, should promote adjustment, while emotional avoidance strategies have

been found to hinder it (Berghuis and Stanton 2002). The only two studies that investigated coping after failed treatment showed that (active and passive) avoidant emotional coping was associated with higher distress up until 5 years after treatment (Daniluk and Tench 2007, Peterson et al. 2009). In addition, meaning based coping was associated with lower personal and relational distress for women, but higher social distress for men (Peterson et al. 2009). Other risk factors for long-term maladjustment were lack of alternative non-parenting roles and social support. The latter may hinder the use of emotional approach coping. Cross sectional studies that investigated adjustment to definitive childlessness had similar findings (Kraaij et al. 2009, Lechner et al. 2007). In addition, these studies showed that cognitive coping is also important for adjustment. While catastrophizing (i.e., emphasizing the negative aspects of childlessness) was associated with worse adjustment, positive reappraisal coping (i.e., attaching a positive meaning to childlessness in terms of personal growth) was associated with better adjustment.

Self-regulation theories hypothesize that when all treatment options are exhausted and parenthood goals are highly blocked, individuals who disengage from parenthood and engage in other meaningful goals adjust better than individuals who remain engaged (Heckhausen et al. 2001). Findings from studies focusing on undesired childlessness support this view from both a behavioural and cognitive perspective (e.g., Heckhausen et al. 2001, Kraaij et al. 2009). However, research shows that people take time to disengage from parenthood: 44% of individuals still desire to have (more) children 3 to 5 years after ending treatment (Verhaak et al. 2007), 25% 10 years after (Wischmann et al. 2012), and 6% 11 to 17 years after (Gameiro et al. 2016). As hypothesized by identity theories (Thoits 1992), childlessness seems to be especially distressing when parenthood is an essential goal in life or when people have few alternative identities or goals (McQuillan et al. 2003, Moura-Ramos et al. 2012).

According to the dual process model of grief (Stroebe and Schut 1999), adjustment to loss depends on peoples' ability to process the loss and, in addition, its implications, for instance the impact it may have on the partnership. Around 42% of individuals experience symptoms of complicated grief after ending fertility treatment (Lechner et al. 2007). For some individuals these symptoms can extend until menopause and be re-experienced when other parenthood milestones (e.g., grandparenthood) are not achieved (Daniluk 2001, Wirtberg et al. 2007). Cognitive Behavioural models of grief (Boelen et al. 2006) explain the recurrence of grief symptoms by the difficulty individuals have to integrate the loss into their autobiographical knowledge. These models also claim that grief reactions are augmented by negative beliefs about the loss, for instance the idea that not having a(nother) child will threaten the partnership. Finally, as was observed in the Daniluk et al. (2007) study, avoidance seems to

be associated with grief symptoms (Boelen et al. 2006), and it is known that individuals experiencing undesired childlessness tend to avoid the 'fertile world' (Volgsten et al. 2010).

Overall these theories highlight the benefits of integrating the loss of parenthood goals and focusing on other goals, and the central role that cognition has in this process. Other theories of adjustment to stressful life events also highlight that individuals strive to find meaning in crises and to regain self-worth and a sense of control over their lives (e.g., Taylor 1983). However, none of these theories fully captures the specificities of the parenthood goal that influence how individuals adjust to its loss. For instance, the loss is never definitive until women achieve menopause, it is invisible, in the sense that there are no exterior signs of the loss (e.g., no body), it centres around a goal that was not yet achieved and it is often a loss that is shared with one (partner) or more (family) people. We argue that added value will be gained from developing a model that describes the specific psychological mechanisms that underlie adjustment to unmet parenthood goals and identifies unique risk factors for maladjustment in this context. Such model can set the basis for increasing awareness about the negative psychological impact of not realizing one's parenthood goals and to guide the provision of psychosocial care in this context.

In sum, it is still unclear if individuals who experience failed fertility treatment need psychosocial care and what such support should entail. The goals of the present mixed-methods review were threefold. First, a meta-analysis was performed to investigate if individuals who experience failed fertility treatment present worse mental-health and wellbeing than individuals for whom treatment is successful. We distinguished between mental-health, which captures the absence of psychopathological symptoms, and wellbeing, considered to be a more holistic and subjective evaluation of one's life satisfaction or happiness (Diener 2000). Second, moderation analysis were performed to investigate if group differences in adjustment vary with the type of comparisons performed, the percentage of children in the unsuccessful group, the follow-up period considered and the quality of the studies. Third, a review of the existing qualitative research was conducted based on thematic analysis (Thomas and Harden 2008) to extract information about the patients' perceived experiences of this period. Based on the qualitative data, we made inferences about the underlying mechanisms of adjustment to unmet parenthood goals and the risk factors for maladjustment, which we present in the form of a comprehensive model.

Methods

Systematic search

The electronic databases of Medline, Embase, PubMed, PsychInfo and Web of Science were searched between 1978 (the year of the first IVF treatment performed) and December

2015 (inclusively) based on search terms for fertility treatment, defined using terminology from the International Committee for Monitoring Assisted Reproductive Technology and the WHO-revised glossary of Assisted Reproductive Technologies (e.g., artificial insemination, assisted reproduction, in vitro fertilization, IVF; Zegers-Hochschild et al. 2009), AND psychosocial adjustment (e.g., adjustment, adaptation, wellbeing, mental-health, quality of life) AND post-treatment (e.g., post, following, after). The detailed search strategy is presented in Table S1 of the supplementary material. The search was adapted for each database. Medical Subject Headings (MeSH) terms were used in PubMed. No restriction was made on the type (journal, conference paper or dissertation) or language of publication. All citations were transferred to EndNote (Thomson Reuters, San Francisco, CA, USA). The results were crosschecked with five articles that had been identified previously as being eligible for the review to ensure the search was suitable for its purpose. The reference sections of all identified articles were examined to identify other relevant manuscripts.

Inclusion and exclusion criteria

This systematic review included quantitative and qualitative studies. Studies were included if they reported on the psychosocial adjustment of a group of individuals (women, men or couples) who experienced failed fertility treatment and if assessments were conducted at least one year after the last treatment cycle. This is the period recommended to ensure patients have definitely finished treatment (Gameiro et al. 2013). Further, literature on stressful life events shows that people take on average 2 years to adjust (Bonanno 2004). Therefore, a minimum period of 1 year allowed us to focus on long-term reactions to loss and not on the immediate short-term distress. To be included quantitative studies had to report on the mental-health or wellbeing of individuals who experienced failed fertility treatment (i.e., failed group) and a control group. Qualitative studies had to report on a group of individuals' stated experiences of psychosocial adjustment after failed fertility treatment.

AF screened the titles and abstracts and AF and SG independently screened the full articles. Any disagreements were resolved with discussion. Duplicate or secondary publications on the same sample reporting on the same outcome were excluded from the meta-analysis to avoid multiple publication bias. In these cases the paper with the most appropriate research questions for the review was prioritized. If this criterion did not apply, the publication reporting data for the largest sample was selected. Excluded studies were classified according to reason for exclusion (see Figure 1).

Data extraction

Quantitative studies

AF extracted data using a standardised protocol. To characterise the set of studies, we extracted data (where available) regarding the authors and country, type of fertility treatment (e.g. IUI, IVF), type and size of failed and control groups, whether the failed treatment group included individuals with children (adoption or spontaneous conception), years since the last treatment cycle, study design (cross-sectional or longitudinal) and outcomes reported (mental-health, wellbeing). AF also extracted the necessary quantitative data to include in the meta-analytical syntheses of group comparisons. SG reviewed data extraction and disagreement was resolved via discussion until consensus was reached.

Qualitative studies

SG extracted data using a standardised protocol. To characterise the set of studies, data was extracted (where available) regarding the authors, country and goal of the study, study sample and percentage with children, study design, years since treatment, and data collection and analytical method. The data extracted for analysis was the content of the results section of each study.

Data synthesis and analysis

Quantitative studies

To investigate group mean differences in mental-health and wellbeing between patients who experienced failed fertility treatment and an adequate control group two meta-analyses were carried out. Only one effect size was returned per study to create an independent set of effect sizes for each analysis. For mental-health we considered measures of anxiety, depression, mental-health, psychopathology or stress because all report on the participants' level of psychopathologic symptoms. We did not include studies reporting on the frequency of specific mental-health diagnosis (e.g., mood disorder) because frequency data cannot be pooled together with group means. For wellbeing we considered measures of general wellbeing, life satisfaction, quality-of-life, general health or vitality, because they capture participants' subjective evaluation of their positive functioning. For studies that reported on multiple measures for the same outcome, specific measures were prioritized. For mental-health, depression was prioritized because it is the measure that best captures the emotional response to the outcome of treatment (Verhaak et al. 2007), followed by mental-health, anxiety, distress and psychopathology. For wellbeing, satisfaction with life was prioritized because it is one of the most commonly used measures in the parenthood literature, followed by quality of life, wellbeing, general health and vitality. If the study reported data for couples, extraction of data took in consideration if the non-independence between male and female

data was addressed (Kenny et al. 2006). If it was not taken in consideration either the male or female data were used to ensure independence, women were prioritized because they are more negatively impacted by failed treatment (e.g., Verhaak et al. 2005).

A random effects model was adopted due to the heterogeneity of our studies (Borenstein et al. 2009). The I^2 index was used to assess the proportion of effect size variability that was due to non-chance factors (Higgins et al. 2003). I^2 values of 0%, 25%, 50% and 75% indicate no, low, moderate and high heterogeneity, respectively. The probability level of $P<0.05$ was used to indicate statistical significance. The primary outcome measure used in the meta-analyses was the standardized mean difference between the unsuccessful and control groups. Hedge's g correction was used to adjust for potential bias due to the small sample sizes found in our studies (Borenstein et al. 2009). Effect sizes were aligned across studies so that a negative standardized mean difference indicated poorer adjustment for the failed treatment group than the control group. Effect sizes of .20, .50 and .80 should be interpreted as small, medium and large, respectively (Cohen 1992).

To investigate the moderation effect of type of control group (based on treatment outcome, parenthood status), existence of children in the failed treatment group (no, yes) and years since last treatment cycle subgroup (X^2 test) and meta-regression (Z-test, unrestricted maximum likelihood) analyses were done. We used Comprehensive Meta-Analysis version 2.2.064 (Biostat, Englewood, USA).

Qualitative studies

Methods of qualitative reviews are emergent and still subject of debate. In this review we opted to follow the synthesis method described by Thomas and Harden (2008). It was designed to identify which interventions need to be developed to address a particular health issue by understanding it from the point of view of the people targeted. We conducted a 3-stage thematic analysis (Dixon-Woods et al. 2005). The first two stages were descriptive. First, SG extracted all results sections verbatim and SG and AM conducted an independent 'line by line' coding of the text to develop descriptors of post-treatment experiences. Two meetings were conducted to contrast and review the descriptors generated. Disagreement was resolved by discussion until the coders agreed if a descriptor should be changed, added or removed from the final list. Second, after a definitive list of descriptors was agreed on, SG developed first order descriptive themes to group the descriptors. The themes needed to capture something important about the data in relation to the research question, representing some level of patterned response across the studies' data (Braun and Clarke 2008). We only considered themes that appeared more than twice across all results and were present in at least two different papers. AF reviewed this process and disagreements were solved by discussion until both authors agreed on the grouping of descriptors and the labels of the second order themes. The last stage was interpretative. SG generated second order

analytical themes that were sufficiently abstract to include or explain all first order themes. This process aimed to create interpretative constructs and testable hypothesis (Thomas and Harden 2008) and constituted the crucial step to develop an explanatory model of adjustment after failed treatment. Finally, SG wrote a narrative summary and developed a graphical depiction of the data. AF reviewed SG's analysis and a similar discussion meeting was conducted to reach agreement on the themes and on their graphical and textual depiction. The percentages of agreement in the coding of the first and second order themes were calculated.

Risk of bias assessment

Quantitative studies

AF and SG designed a critical appraisal form to assess study quality in terms of population representativeness of the study sample (3 points), measurement validity (1 point) and reliability (1 point), and evaluation of research hypotheses (3 points). These criteria are described in detail in Table S3 of the supplemental material. Each study quality score could range from 0 to 8. Papers were classified into low [≤2], moderate [3-5] or high [≥6] quality studies. AF and SG independently assessed study quality using the form and disagreement was resolved via discussion until consensus was achieved. Subgroup analyses were performed to investigate if the pooled estimates varied as a function of study-quality. Publication bias was checked with visual inspection of funnel plots (Hedge's g against its standard error) and Egger's test (Egger et al. 1997). Trim and fill was used to adjust the standardized mean different between the failed and control groups for the presence of publication bias (Duval and Tweedie 2000).

Qualitative studies

Two possible sources of bias were considered: the quality of the primary research and the researchers. The quality of primary research was assessed using the Critical Appraisal Skills Programme Qualitative Research Checklist (10 criteria, Critical Appraisal Skills Programme (CASP) 2013), http://media.wix.com/ugd/dded87_29c5b002d99342f788c6ac670e49f274.pdf (see Table S7 of supplementary material). Based on the number of met criteria, each study was classified as being of low [≤3], moderate [4-7] or high [≥8] quality.

As in any qualitative analysis, the researchers' previous understandings may have shaped their interpretation of the data. SG has been researching the psychosocial issues of infertility for around 10 years and has in-depth knowledge of the field literature, including of the post-treatment period. AF is a medical student with no knowledge on this topic prior to the start of the study. None of the authors has regular contact with individuals from the study population. In order to increase the trustworthiness of the findings and reduce possibility of researcher bias, AF and SG independently assessed study quality and disagreement was resolved via

discussion until consensus was reached. The themes had to be closely linked to the data collected (Braun and Clarke 2008). As a final step, the findings were presented to a group of five infertility researchers and to a group of six patients and revised according to the feedback received.

Results

Description of studies

Figure 1 around here

Figure 1 presents the study decision-flow chart whereby 18 studies were included in the mixed-methods systematic review. The systematic search returned 6397 non-duplicated records. Ninety-eight per cent of these studies (n = 6265) were excluded based on their title and abstract, which reduced the number of eligible studies to 132. From these 15 were retained after inspection of the full text and three further studies were identified through the inspection of references lists. The three most common reasons for exclusion were that assessment was done during the first 12 months after treatment, the paper did not report on primary research or the paper did not focus (nor reported) on psychosocial adjustment (see Table S2 of supplementary material for reasons for exclusion). During this process we contacted seven authors to clarify about missing or inconsistent data. Only two replied with the required information.

Table 1 around here

Table 1 presents the characteristics of the nine included quantitative studies (designated in the reference list by asterisks). The studies sampled 2278 individuals with failed fertility treatment from seven countries. All participants had undergone IVF but some also underwent other treatments including ICSI and gamete intrafallopian transfer (GIFT). Two types of control group were used, based on the outcome of treatment (successful, 33%) or on parenthood status (with children, either adopted or conceived spontaneously or with treatment, 67%) at time of assessment. The time since last treatment varied from one (12 months) to 20 years. Two studies (22%) had a longitudinal design. Only three studies (33%) had a response rate of 80% or higher. Six (67%) studies reported on mental-health and six on wellbeing.

Quality ratings for the studies are also shown in Table 1. Study quality ratings indicated no low, seven moderate (78%) and two high-quality studies (22%) (See Tables S3 to S6 in

supplementary material for details on quality assessment). The percentage of agreement between SG and AF on quality ratings was 93%.

Table 2 around here

Table 2 presents the characteristics of the nine included qualitative studies (designated in the reference list by asterisks). The studies sampled 267 individuals who experienced failed treatment from six countries. In six studies (67%) the sample included people with children. Eight studies (88%) were cross-sectional and one (12%) was longitudinal with three follow-up assessments occurring more than one year after treatment (Daniluk 2001). Data was collected via interviews and the most common analytical technique was phenomenological analysis (n = 4, 44%), but thematic, content and discourse analysis and grounded theory were also used.

Quality ratings for the qualitative studies are also shown in Table 2. Study quality ratings indicated no low, two moderate (22%) and seven high-quality studies (78%) (See Table S7 in supplementary material for details on quality assessment). The percentage of agreement between SG and AF on quality ratings was 83%.

Quantitative meta-synthesis: Psychosocial adjustment after failed fertility treatment

Figure 2 around here

Mental-health

Six studies reported data on group mean differences for at least one mental-health measure that could be included in the meta-analysis. Depression scores were considered for four studies, mental-health for one (Gameiro et al. 2016) and anxiety for one (Leiblum et al. 1998). One study reported on frequency of diagnosis (Yli-Kuha et al. 2010) and could not be included.

Figure 2 shows the pooled standardized mean differences in mental-health and wellbeing between the failed and control groups for the random effects model. The meta-analysis showed that the failed treatment group had significantly worse mental-health than the control group (g = -0.496, P=0.001, 95%CI [-0.791 -0.200]), with significant heterogeneity between studies (I^2=86%, P<0.001).

The Egger's test indicated the presence of publication bias (intercept= -3.192, t = 4.85, P = 0.004), which was confirmed through visual inspection of the funnel plot. The trim and fill method also confirmed the presence of publication bias. This method identified one missing study, estimating an adjusted point estimate of -0.458 that was still significant (95% CI -0.726 -0.189).

Table 3 around here

Table 3 presents the standardized mean differences for mental-health and wellbeing between the failed and control groups for the pre-defined moderators. Subgroup analysis for the type of group comparisons done showed that mean group differences were bigger when the control group is based on the outcome of treatment than when it is based on parenthood status. The remaining subgroup and meta-regression analyses were statistically non-significant.

Wellbeing

Six studies reported data on group mean differences for at least one wellbeing measure that could be included in the meta-analysis. Life satisfaction scores were considered for three studies, quality of life for two (Weaver et al. 1997, Wischmann et al. 2012) and wellbeing (Johansson et al. 2010) for one. The meta-analysis showed that the failed treatment group had significantly worse wellbeing than the control group (g = -0.415, P<0.001, 95%CI [-0.627 -0.203]), with non-significant heterogeneity between studies (I^2=53%, P=0.059). Visual inspection of the forest plot (see Figure 1) suggests that the study by Leiblum and colleagues (1998) can be considered an outlier (its effect size was more than twice the pooled standard deviation distant from the pooled effect size). Therefore we repeated the meta-analysis excluding this study. The meta-analysis still showed that the failed treatment group had significantly worse wellbeing than the control (g = -0.324, P<0.001, 95%CI [-0.454 -0.193]), with no heterogeneity between studies (I^2=0%, P=0.552). Subsequent analyses were run without including this study.

The Egger's test indicated no presence of publication bias for wellbeing (intercept = -0.523, t= 0.428, P = 0.349). The trim and fill, however, identified one missing study, estimating an adjusted point estimate of -0.298 that was still significant (95% CI -0.422 -0.175).

Subgroup and meta-regression analyses could only be performed for the moderators existence of children in the failed treatment group and years since treatment and none were statistically significant.

Qualitative meta-synthesis: Patients' perceived experiences of the period after failed treatment

To illustrate the analysis process, please consider the following lines of text presented in the study by Boden (2007): "On a positive note, many couples expressed their belief that the experience of treatment had enhanced their relationship with their partner, bringing them, as participants described, 'closer together'".

To summarize these data the following descriptor was developed: 'treatment strengthens the relationship'. After, this and other similar descriptors, for instance 'stronger partnership as result of infertility and treatment experience' and 'treatment strengthens the marital relationship and leads to marital satisfaction', were grouped into the first order theme 'treatment strengthens partnership'. Finally, this and other first order themes, for instance, 'intense grief', 'anger directed at everything and everyone', 'sexual difficulties' and 'men stoic and supportive of their partner' were classified into the second order theme 'Individual and relational adjustment'. Table S8 of the supplemental material presents, for each study included, the list of descriptors of post-treatment experiences and their classification into first and second order themes. Inter-coder agreement for 1st and 2nd order themes was 83.3% and 92.2%, respectively.

A total of 28 first order themes and six second order themes were developed and are presented in Table 4. Table S9 presents all first order themes developed, the number of times they were cited in the results sections of the studies included, the number of studies citing it, and its association with the second order themes. In total 33 first order themes were developed, however, as can be seen in table S9, five were only cited once and were therefore not considered.

Table 4 around here

A comprehensive description of the themes is provided below. The themes Individual and relational adjustment, Social adjustment and Fertility care perceptions and needs capture changes across time in experiences and adjustment. The themes Acceptance, Meaning making and Pursuit of new life-goal capture the psychological tasks patients engaged with that were associated with positive adjustment in the aftermath of failed treatment.

Individual and relational adjustment. Participants described the period immediately after ending treatment as of intense grief (pain, sadness and emptiness) and anger felt towards 'everything and everyone'. Grief endured in time and was lived in isolation. The grief experience was different for women and men. For women it was very intense and characterized by a sense of emptiness, meaningless, loss of control and self-blame. Women tended to be open about their grief with their family and friends. Men felt pain, tried to overcome grief by denying it and tended to be private in their suffering. Men were surprised with the intense grief reactions of their (female) partners and felt it was their duty to be stoic and supportive. However, they felt they lacked the knowledge to do this. In addition, most participants reported that the treatment experience strengthened their partnership. However, the sexual desire of many couples was affected during treatment and it did not recover once treatment was over. Despite the initial difficulties and grief, as time passed participants

reported a sense of survival and personal and spiritual growth that was associated with having undergone the treatment and the childlessness experience.

Social adjustment. Most participants experienced social isolation. Isolation was the result of not meeting parenthood social norms and therefore being unable to share parenthood experiences with peers. Men also felt isolated from family members, with whom they felt they communicated less due to not having children. The absence of children was a loss experienced by the whole family and raised concerns about not leaving a legacy to continue the family bloodline and about loneliness in later life. Many women talked about how social constructions and pressure for parenthood made it harder for them to accept their situation. For instance, women's identities relied so strongly on parenthood that not having the desired children made them doubt their femininity or feel abnormal. Some women felt the need to show others they did everything they could to become parents to avoid social scrutiny. Others talked about how it was difficult to escape society's expectations. Overall participants felt that not meeting their parenthood goals hindered their relationship with their social network. Despite the overall isolation, participants also described choosing to share their struggles with a selected few close friends and work colleagues who provided them with good emotional and instrumental support. Other people with infertility also provided them with useful information. Finally, for women social isolation also resulted from active efforts to avoid situations where children were present in order to escape suffering.

Fertility care perceptions and needs. Participants said that when all treatment options were exhausted they felt abandoned by their clinic. They perceived a lack of support from the clinic in both deciding if and when ending treatment and once this decision was made. They expressed a need for psychosocial care or counselling to help them cope with the intense grief they experienced, and for sexual counselling to help them overcome the sexual difficulties that developed during treatment. Participants acknowledged the limitations of reproductive medicine and this facilitated acceptance of their situation. Reaching acceptance was associated with developing a positive outlook on ending treatment, for instance as allowing participants to let go of their medical experience, as actively managing their child-desire and as allowing others to also have an opportunity to achieve parenthood.

Meaning making. In the process of coming to terms with their unmet parenthood goals, participants engaged in an intense re-evaluation effort whereby they tried to make sense of their past efforts to achieve parenthood and they re-evaluated their life values and priorities in search for a new identity and future in the absence of the desired children. Traditional notions of marriage (e.g., marriages do not survive without children), femininity (e.g., women are born to be mothers) and family (e.g., children are part of family life) hindered this meaning making process. Therefore many participants questioned such beliefs and reflected on the meaning of marriage and family without children. This critical approach to social

constructions of parenthood facilitated acceptance. Overall participants felt that meaning making was associated with higher acceptance of one's unmet parenthood goals.

Acceptance. Immediately after finishing treatment participants maintained a strong hope to have (more) biological children and felt unwilling or unable to accept their situation. Acceptance was a process that unfolded over time because of the need to 'move on' with life. Many patients differentiated rational from emotional acceptance and stated that although rationally they accepted their situation, this was not synonymous with or followed by emotional acceptance. Indeed, the hope to have the desired children could persist for many years after treatment and for some women it only ended when they reached menopause, as menstruation meant that a pregnancy was still possible. Acceptance was easier when patients were convinced they did their best and explored all available treatment options, when their efforts were acknowledge by their family and friends and met with empathy instead of pressure for them to become parents. Acceptance was also easier when participants felt they received good medical advice during treatment, when their prognosis was clear and poor and when the reasons for treatment failure were identified. Acceptance was associated with a change in the participants' view of the world. This translated in recognizing their limits to control reality, in the development of a more realistic view of the world or in increasingly valuing the 'good things' in life. Acceptance was also associated with better adjustment. It translated in a sense of freedom from the infertility experience and hope towards the future. This new state was described as of restored strength and agency, and equilibrium with both the self and in the partnership. In contrast, the inability to accept their situation was associated with maladjustment, in particular with continuous grief, feeling 'stuck in life', low self-esteem and different avoidance strategies.

Pursuit of new life-goals. When ending treatment participants felt incapable of envisioning themselves as anything else than parents. For most participants the need to move on with their lives resulted from a desire to overcome the pain of grief and to reject self-pity and continuous mourning. Coherently, the willingness to 'move on' was lower when there was a stronger desire for children. Overall participants perceived that the needs to accept their situation and 'move on' were interrelated. Participants' typical strategy was to invest in new goals and activities that distracted them from their pain, compensated for their childlessness status and boosted their self-esteem. Despite desiring to 'move on', many participants did not know how to do this and expressed that the process required an active effort to look into the future, which implied a confrontation with later life stages and their lack of legacy. Participants talked about many different ways of finding new fulfilling life goals and activities and these included caring for other people (e.g., family members) and pets, supporting friends with infertility and advocating for the rights of other infertile people, travelling and focusing on their career and, when full acceptance was achieved, some individuals even

refocused on caring for other children (e.g., siblings). Overall participants felt that the pursuit of new life-goals was associated with higher acceptance of their unmet parenthood goals.

Discussion

Findings from this mixed-methods systematic review show that people who do not manage to conceive with fertility treatment experience worse mental-health (i.e., more psychopathologic symptoms) and wellbeing (lower positive functioning and life satisfaction) than those who conceive. Positive adjustment in this context seems to be associated with a progressive acceptance of the loss; making meaning of past efforts to conceive and the present situation; and pursuing new life goals or activities. Individuals who engage with these three psychological tasks tend to report a sense of personal growth and to develop a positive outlook on ending treatment and renewed views of the world. The results reported suggest that the risk for maladjustment varies according to individual, social and treatment related factors. Overall the data indicate that the post treatment period is worthy of clinical attention and reinforce advice for health professionals to support individuals during this period.

Long-term adjustment after failed fertility treatment is an emergent research topic within reproductive psychology. Only 18 studies could be identified for inclusion with an even split between quantitative and qualitative research. An examination of the papers excluded (see Figure 1) indicates that researchers tended to focus on the first 12 months following treatment (only capturing short term reactions), did not report on time since treatment, focused on topics either than adjustment, or, if they focused on adjustment, did not use a control group. From the included studies only two reported on correlates or predictors of adjustment (Gameiro et al. 2016, Verhaak et al. 2007) making it impossible to quantitatively investigate risk factors. In sum, research on this topic is still at a descriptive state that needs to be moved towards an explanatory theory-driven approach. This mixed-methods review is timely in that it provides the evidence base for the development of testable hypotheses organized in a comprehensive theoretical model.

Results from the meta-analysis show that individuals who do not manage to conceive with fertility treatment report significantly worse mental-health (medium effect sizes) and wellbeing (small to medium effect sizes) than those who do. Moderation analyses focusing on the type of control group and percentage of children in the failed treatment group suggest that the poorer adjustment of the failed group is better explained by their inability to meet their parenthood goals than their childless status. Indeed, group differences in mental-health were larger when based on the treatment outcome than on parenthood status (-0.604 versus -0.108) and the existence of children in the failed treatment group did not prove to be a significant moderator for mental-health nor wellbeing. These data suggests that what is associated with worse adjustment is having less children than desired or, in other words, not

meeting ones parenthood goals. Recent research shows that both the population of people with undesired childlessness and fewer children than desired are increasing as a result of parenthood delay (Schmidt et al. 2012).

Contrary to our hypothesis, group differences in mental-health and wellbeing did not subside within the 20 years time period surveyed. However, this result needs to be carefully considered because it may be explained by a methodological limitation in the operationalization of time since treatment. Indeed, seven out of the nine studies included provided a range for time since last treatment (instead of a specific time or average) and this could span several years. In these cases we had to use the mean and this is likely to have obscured associations.

Findings from the qualitative review indicate that positive adjustment is dependent on individuals accepting and making meaning of their unmet parenthood goals and pursuing new life-goals. Individuals perceived these as three separate but inter-dependent psychological tasks. Acceptance was portrayed as a process (and not an outcome). It appears to involve the emotional processing of one's unmet parenthood goals. According to psychological theories, this implies the activation and restructuring of maladaptive fear structures, exposure and habituation to the situation and regulation of the negative affect it elicits (Park 2010). Therefore, acceptance may be fostered by emotional approach coping to the loss (Boelen et al. 2006, Stanton et al. 1994) so that, in accordance with grief models, its integration in autobiographical knowledge is achieved (Boelen et al. 2006). However, unmet parenthood goals can be an ambiguous loss because the chances of achieving parenthood are rarely completely null. This may preclude the integration of the loss. Indeed, our qualitative data showed that acceptance was easier when chances to conceive were lower (poor prognosis, clear advice to stop treatment from fertility team, reasons for failure identified), or in other words, when the loss was perceived as definitive.

Meaning making seems to reflect an adaptive cognitive processing of the situation. The literature about adaptation to stressful life events or trauma (e.g.; Park 2010, Taylor 1983) considers meaning making as a critical task for adjustment and states that individuals will always engage in it in order to restore a sense of control and predictability over their lives (Kelley and Michela 1980). Our results seem to be consistent with this literature in that patients perceived meaning making and acceptance to be related and in that acceptance was associated with new views of the world and redefinition of priorities in life (Park 2010). Grief theorists also claim that such cognitive processing needs to address dysfunctional beliefs associated with the loss (Boelen et al. 2006). For instance, individuals talked about guilt related with ending treatment and concerns about their role in their family. Previous research also highlighted concerns about the impact of not achieving parenthood on the partnership and one's ability to endure the related grief (Boivin et al. 2005). In this context,

self-blaming and catastrophizing were shown to be negatively associated with adjustment (Kraaij et al. 2009). On the contrary, positive reappraisal coping may assist meaning making and was associated with positive adjustment (Kraaij et al. 2009, Peterson et al. 2009).

Finally, results suggest that the pursuit of new life goals is associated with higher acceptance of the unmet parenthood goals and, in the long-term, better adjustment. Individuals claimed that pursuing new life goals required an active effort from their part. According to the dual process model of grief, this can be regarded as an effort to process the implications of the loss (Stroebe and Schut 1999) by finding new purpose(s) for living and regaining a sense of agency. A recently published meta-analysis showed that, in the context of a parenthood blockage, reengagement in new life goals is associated with better wellbeing (i.e., higher positive mood; Mesquita da Silva et al. 2016). Our results support these data, but also show that engaging with new life goals is harder for individuals who highly value parenthood or perceive a lack of other meaningful goals to pursuit (Thoits 1992). Our data also show that individuals start by pursuing other goals just as a distraction from the suffering of not achieving their main goal of parenthood, but that they end up perceiving benefits from this. These results are consistent with literature showing that pursuing new goals is adaptive even when individuals have not completely disengaged from parenthood (Wrosch et al. 2003).

Another important finding was how the unmet parenthood goals affected individuals' social relationships and how their social context influenced their adjustment process. This is not surprising as it is well known that the socio-cultural context has a strong influence in shaping the experience of infertility and childlessness of both women and men (Greil et al. 2011, Martins et al. 2016). Social isolation was a common theme in the data, either resulting from less empathic reactions from the social network or from self-avoidance of that social network, when it implied contact with children. Social support is a consistent protective factor in the context of infertility and its treatment (Martins et al. 2016, Rockliff et al. 2014). Our data was no different in this regard, but also showed that the social context could influence adjustment via other routes than support, more specifically via social norms (i.e., social pressure for parenthood) and prevalent representations of parenthood, marriage and femininity. Overall, the pervasiveness of social and cultural first order themes suggests that psychosocial care needs to include a thorough assessment of the social context and provide individuals with strategies to cope with it.

Strengths and limitations

The strengths of this mixed-methods review were the systematic review of almost 40 years of research on adjustment after failed fertility treatment, which yielded 18 studies from ten countries sampling adjustment experiences of 2545 patients. Data were independently

extracted and quality evaluations and analyses were made according to a-priori defined and rigorous protocols in order to minimize the risk of bias. Another strength was the inclusion of quantitative and qualitative research to answer different research questions, allowing for a more comprehensive picture of adjustment to emerge. The primary research was of moderate to high quality. In quantitative studies quality was due to the fact that most studies had a representative sample of the population, used valid and reliable outcome measures and had enough power to detect significant differences. In qualitative studies the quality was due to the adequacy of the design, recruitment strategy and data collection method used. However, the quantitative studies did not focus on investigating risk factors for maladjustment and used heterogeneous samples of individuals in terms of time since treatment and parenthood status. It was also not totally clear if those individuals who did not reach parenthood via treatment were still pursuing it, for instance by doing more treatment at different clinics or by choosing to adopt. These data, together with previous research (e.g., Gameiro et al. 2014, Wischmann et al. 2012), suggest that the larger the proportion of the sample still pursuing parenthood the worse adjustment will be. The main problem of the qualitative data was that, due to the nature of qualitative research, the data extracted from the papers were interpretations that other researchers made of patients' interpretations of their experiences. Another problem was that the goals of the qualitative studies varied slightly, for instance one study focused only on the perceived gains of doing fertility treatment (Lee et al. 2009). This might have affected the prevalence of the emergent themes.

Despite these limitations, overall the results from the quantitative and qualitative review were consistent and its integration provided a coherent picture of adjustment after failed treatment. The only inconsistency detected was that qualitative data provided a more positive perspective on adjustment over time than the quantitative data. However, we explained that the quantitative data might not be reliable. Furthermore, it should be noted that the quantitative data taped into hedonic wellbeing, which is the extent to which individuals are 'feeling well', and the qualitative data also captured eudaimonic wellbeing, which is the extent to which individuals are 'doing well' in terms of purpose, meaning and fulfilment in life (Ryan and Deci 2001). By analysing both types of data we achieved a more nuanced view of wellbeing, which can explain the discrepancy observed.

Conclusions

Based on the data reported we developed a comprehensive model to explain adjustment to unmet parenthood goals, labelled the Three Tasks Model of Adjustment to Unmet Parenthood Goals and presented in Figure 3. The model predicts that, in the face of unmet parenthood goals, individuals able to accept their situation, make meaning of their experience and pursue new life goals will adjust better than those who do not engage with

these psychological tasks. These three main predictions are consistent with knowledge from other theories previously used to explain adjustment after failed treatment and general theories of adjustment to stressful life events. The Three Tasks Model of Adjustment to Unmet Parenthood Goals also predicts that specific variables may influence adjustment, either directly (e.g., women are more likely to present worse adjustment), by impacting on the individual's ability to perform the three psychosocial tasks (e.g., one is less likely to pursue new goals if they do not find alternative meaningful goals), or by moderating the relationship between the three psychological tasks and adjustment (e.g., pursuit of new life goals is more likely to lead to better adjustment when one is supported by their network than when is not supported). The variables identified were, at the individual level, gender, importance of parenthood and availability of alternative life goals; at the social level, social support, social representations of parenthood and parenthood social norms; and at the treatment level, prognosis, perceived quality of medical advice, reasons for failure identified and treatment options exhausted.

In this review we considered two adjustment outcomes: mental health and wellbeing. The meta-analysis showed that these are sensitive measures in this context and should therefore continue to be used in future research. The qualitative data suggests that patients assess their adjustment not only in terms of how they feel (hedonic wellbeing, e.g. intense grief, anger) but also in terms of how they think they are realizing their human potential (eudaimonic wellbeing, e.g. resiience and personal and spiritual growth; Ryff 1995). Future research could consider using outcomes that tap into eudaimonic wellbeing. This seems particularly important in this context because the pursuit and realisation of intrinsic goals seems to be more strongly related with eudaimonic than hedonic wellbeing (Ryan and Deci 2001). Examples of eudaimonic outcomes would be self-acceptance (positive evaluations on oneself and one's past life, personal growth (sense of continued growth and development as a person) or purpose in life (the belief that one's life is purposeful and meaningful (Ryff and Keyes 1995).

In conclusion, this mixed-methods review offers compelling evidence for the need to provide psychosocial care directed at helping individuals relinquishing their parenthood goals. Patients themselves recognize the need for psychosocial care during this period and express frustration at feeling abandoned by their clinic. The Three Tasks Model of Adjustment to Unmet Parenthood Goals offers comprehensive guidance on the therapeutic mechanisms that seem to underlie positive adjustment. The model was developed based on the experiences of patients who experienced failed fertility treatment, but there is nothing to suggest that it cannot be applied to everyone who is confronted with the loss of their parenthood goals, regardless of the pathway that led to it. Future research should now test the model in these different contexts. This can be done by conducting prospective cohort

studies to test the model predictions. Another way to test the model is by developing and testing psychosocial interventions that are based on its three main hypothesis, in other words, interventions designed to engage individuals with acceptance, meaning making and pursuit of new goals. An important question in this context is how such specialized care can reach individuals who are already outside the reproductive healthcare system. The ESHRE guideline argues that this is the responsibility of fertility clinics (Gameiro et al. 2015). We concur that specialized expertise is needed to develop and test effective interventions. A sensible approach would be to develop self-help interventions that individuals could access via multiple outlets (the most obvious one being the internet) instead of having to go back to their fertility clinics. Another crucial approach is to promote knowledge transfer about the topic so that primary care physicians and mental-health professionals are aware of the deleterious impact of unmet parenthood goals on wellbeing and of how it can be addressed.

Author's roles

S.G. did data extraction, critical appraisal, data analysis and interpretation and writing of the report. A.F. did data extraction, critical appraisal, data analysis and interpretation. All authors approved the final version for submission.

Acknowledgements

We would like to thank our colleagues Professors Jacky Boivin and Christianne Verhaak and Dr Emily Koert for their careful reviews of draft versions of the present paper.

Funding

This work was supported by the UK Economic and Social Research Council [grant number 508860].

Conflict of interest

There are no conflicts of interest.

(Hammarberg et al. 2001, Johansson and Berg 2004, Lee et al. 2009, McCarthy 2008, Su and Chen 2006, Sydsjö et al. 2015, Throsby 2001)

References

Berghuis, J. P. and A. L. Stanton. Adjustment to a dyadic stressor: A longitudinal study of coping and depressive symptoms in infertile couples over an insemination attempt. *Journal of Consulting and Clinical Psychology*. 2002**: 2**; 433-438.

Boden, J. When IVF treatment fails. *Human Fertility*. 2007**: 2**; 93-98.

Boelen, P. A., M. A. van den Hout and J. van den Bout. A cognitive-behavioral conceptualization of Complicated Grief. *Clinical Psychology: Science and Practice*. 2006**: 2**; 109-128.

Boivin, J., J. Takefman and A. Braverman. Giving bad news: It's time to stop. In N. Macklon Assisted Reproduction in the Complicated Patients. 2005. Taylor & Francis, London, 233-240.

Boivin, J., L. Bunting, J. A. Collins and K. G. Nygren. International estimates of infertility prevalence and treatment-seeking: Potential need and demand for infertility medical care. *Human Reproduction*. 2007**: 10**; 1506-1512.

Bonanno, G. A. Loss, trauma, and human resilience: Have we underestimated the human capacity to thrive after extremely aversive events? *American Psychologist*. 2004**: 1**; 20-28.

Borenstein, M., L. V. Hedges, J. P. Higgins and H. R. Rothstein. Introduction to Meta-Analysis. 2009. John Wiley & Sons, Ltd., UK.

Braun, V. and V. Clarke. Using thematic analysis in psychology. *Qualitative Research in Psychology*. 2008**: 2**; 77-101.

Cohen, J. A power primer. *Psychological Bulletin*. 1992**: 1**; 155-159.

Critical Appraisal Skills Programme (CASP) (2013). CASP Qualitative Research Checklist.

Daniluk, J. C. Reconstructing their lives: A longitudinal, qualitative analyses of the transition to biological childlessness for infertile couples. *Journal of Counselling and Development*. 2001**: 4**; 439-449.

Daniluk, J. C. and E. Tench. Long-term adjustment of infertile couples following unsuccessful medical intervention. *Journal of Counselling and Development*. 2007**: 1**; 89-100.

Datta, J., M. J. Palmer, C. Tanton, L. J. Gibson, K. G. Jones, W. Macdowall, A. Glaiser, P. Sonnenberg, N. Field, C. H. Mercer, A. M. Johnson and K. Wellings. Prevalence of infertility and help seeking among 15000 women and men. *Human Reproduction*. 2016**: 9**; 2108-2118.

Diener, E. Subjective well-being: The science of happiness and a proposal for a national index. *Americal Psychologist*. 2000**: 1**; 34-43.

Dixon-Woods, M., S. Agarwal, D. Jones, B. Young and A. Sutton. Synthesising qualitative and quantitative evidence: a review of possible methods. *Journal of Health Services Research & Policy*. 2005**: 1**; 45-53.

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

Egger, M., G. D. Smith, M. Schneider and C. Minder. Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*. 1997**: 7109**; 629-634.

Gameiro, S., J. Boivin and A. D. Domar. Optimal IVF for 2020 should reduce treatment burden and enhance care delivery for patients and staff. *Fertility and Sterility*. 2013**: 2**; 302-309.

Gameiro, S., C. M. Verhaak, J. A. M. Kremer and J. Boivin. Why we should talk about compliance with assisted reproductive technologies (ART): a systematic review and meta-analysis of ART compliance rates. *Human Reproduction Update*. 2013**: 2**; 124-135.

Gameiro, S., A. W. van den Belt-Dusebout, E. Bleiker, D. Braat, F. E. van Leeuwen and C. M. Verhaak. Do children make you happier? Sustained child-wish and mental health in women 11-17 years after fertility treatment. *Human Reproduction*. 2014**: 10**; 2238-2246.

Gameiro, S., J. Boivin, E. A. F. Dancet, C. de Klerk, M. Emery, C. Lewis-Jones, P. Thorn, U. Van den Broeck, C. A. Venetis, C. M. Verhaak, T. Wischmann and N. Vermeulen. ESHRE Guideline: Routine psychosocial care in infertility and medically assisted reproduction - A guide for fertility staff. *Human Reproduction*. 2015**: 11**; 2476-2485.

Gameiro, S., A. W. van den Belt-Dusebout, J. Smeenk, D. Braat, F. E. van Leeuwen and C. M. Verhaak. Women's adjustment trajectories during IVF and impact on mental health 11-17 years later. *Human Reproduction*. 2016**: 8**; 1788-1798.

Graham, M. Is being childless detrimental to a woman's health and well-being across her life course? *Women's Health Issues*. 2015**: 2**; 176-184.

Greil, A., J. McQuillan and K. Slauson-Blevins. The social constructions of infertility *Sociology Compass*. 2011**: 8**; 736-746.

Hammarberg, K., J. Astbury and H. W. G. Baker. Women's experience of IVF: A follow-up study. *Human Reproduction*. 2001**: 2**; 374-383.

Heckhausen, J., C. Wrosh and W. Fleeson. Developmental regulation before and after a developmental deadline: The sample case of 'Biological clock' for childbearing. *Psychology and Aging*. 2001**: 3**; 400-413.

Higgins, J. P. T., S. G. Thompson, J. J. Deeks and D. G. Altman. Measuring inconsistency in meta-analysis. *British Medical Journal*. 2003**:** 557-560.

928 Johansson, M. and M. Berg. Women's experiences of childlessness 2 years after the end
929 of in vitro fertilization treatment. *Scandinavian Journal of Caring Sciences*. 2004**: 1**; 58-
930 63.
931
932 Johansson, M., A. Adolfsson, M. Berg, J. Francis, L. Hogström, P. O. Janson, J. Sogn and A. L.
933 Hellström. Gender perspective on quality of life, comparisons between groups 4-5.5
934 years after unsuccessful or successful IVF treatment. *Acta Obstericia et Gynecologica*
935 *Scandinavica*. 2010**: 5**; 683-691.
936
937 Kelley, H. H. and J. L. Michela. Attribution theory and research. *Annual Review of*
938 *Psychology*. 1980**:** 457-501.
939
940 Kenny, D. A., D. A. Kashy and W. L. Cook. Dyadic Data Analysis. 2006. The Guilford Press,
941 New York, London.
942
943 Kraaij, V., N. Garnefski and M. J. Schroevers. Coping, goal adjustment and positive and
944 negative affect in definitive infertility. *Journal of Health Psychology*. 2009**: 1**; 18-26.
945
946 Kuivasaari-Pirinen, P., H. Koivumaa-Honkanen, M. Hippeläinen, K. Raatikainen and S.
947 Heinonen. Outcome of Assisted Reproductive Technology (ART) and subsequent self-
948 reported life satisfaction. *Plos One*. 2014**: 11**; e112540.
949
950 Lazarus, R. S. and S. Folkman. Stress, appraisal and coping. 1984. Springer, New York.
951
952 Lechner, L., C. Bolman and A. van Dalen. Definitive involuntary childlessness:
953 associations between coping, social support and psychological distress. *Human*
954 *Reproduction*. 2007**: 1**; 288-294.
955
956 Lee, G. L., W. H. H. Choi, C. C.H.Y., C. L. W. Chan and E. H. Y. Ng. Life after unsuccessful IVF
957 treatment in an assisted reproduction unit: A qualitative analysis of gains through loss
958 among Chinese persons in Hong Kong. *Human Reproduction*. 2009**: 8**; 1920-1929.
959
960 Leiblum, S. R., A. Aviv and R. Hamer. Life after infertility treatment: a long-term
961 investigation of marital and sexual function. *Human Reproduction*. 1998**: 12**; 3569-3574.
962
963 Leridon, H. Can assisted reproduction technology compensate for the natural decline in
964 fertility with age? A model assessment. *Human Reproduction*. 2004**: 7**; 1548-1553.
965
966 Martins, M. V., M. Basto-Pereira, J. Pedro, B. Peterson, V. Almeida, L. Schmidt and M. E.
967 Costa. Male psychological adaptation to unsuccessful medically assisted reproduction
968 treatments: a systematic review. *Human Reproduction Update*. 2016**: 4**; 466-478.
969
970 McCarthy, M. P. Women's lived experience of infertility after unsuccessful medical
971 intervention. *Journal of Midwifery & Women's Health*. 2008**: 4**; 319-324.
972
973 McQuillan, J., A. Greil, L. White and M. Jacob. Frustated fertility: Infertility and
974 psychological distress among women. *Journal of Marriage and the Family*. 2003**: 4**; 1007-
975 1018.
976

977 Mesquita da Silva, S., J. Boivin and S. Gameiro. Self-regulation and wellbeing when facing
978 a blocked parenthood goal: A systematic review and meta-analysis. *PlosOne*. 2016**: 6**;
979 e0157649.
980
981 Moura-Ramos, M., S. Gameiro, M. C. Canavarro, I. Soares and A. T. Santos. The indirect
982 effect of contextual factors on the emotional distress of infertile couples. *Psychology and*
983 *Health*. 2012**: 5**; 533-549.
984
985 Park, C. Making sense of the meaning literature: An intergrative review of meaning
986 making and its effects on adjustment to stressful life events. *Psychological Bulletin*. 2010**:**
987 **2**; 257-301.
988
989 Peterson, B. D., M. Pirritano, U. Christensen, J. Boivin, J. Block and L. Schmidt. The
990 longitudinal impact of partner coping in couples following 5 yeards of unsuccessful
991 fertility treatments. *Human Reproduction*. 2009**: 7**; 1656-1664.
992
993 Pinborg, A., C. O. Hougaard, A. N. Andersen, D. Molbo and L. Schmidt. Prospective
994 longitudinal cohort study on comulative 5-year delivery and adoption rates among 1338
995 couples initiating infertility treatment. *Human Reproduction*. 2009**: 4**; 991-999.
996
997 Rockliff, H. E., S. L. Lightman, E. Rhidian, H. Buchanan, U. Gordon and K. Vedhara. A
998 systematic review of psychosocial factors associated with emotional adjustment in In
999 Vitro Fertilization patients. *Human Reproduction Update*. 2014**: 4**; 594-613.
1000
1001 Ryan, R. M. and E. L. Deci. On hapiness and human potentials: A review of research on
1002 hedonic and eudaimonic well-being. *Annual Review of Psychology*. 2001**:** 141-166.
1003
1004 Ryff, C. D. Psychological wellbeing in adult life. *Current Direction in Psychological Science*.
1005 1995**: 4**; 99-104.
1006
1007 Ryff, C. D. and C. L. M. Keyes. The structure of psychological wellbeing revisited. *Journal*
1008 *of Personality and Social Psychology*. 1995**: 4**; 719-727.
1009
1010 Schmidt, L., T. Sobotka, J. G. Bentzen and A. Nyboe Andersen. Demographic and medical
1011 consequences of the postponement of parenthood. *Human Reproduction Update*. 2012**:**
1012 **1**; 29-43.
1013
1014 Stanton, A. L., S. Danoff-Burg, C. L. Cameron and A. P. Ellis. Coping throught emotional
1015 approach: Problems of conceptualization and confounding. *Jornal of Personality and*
1016 *Social Psychology*. 1994**: 2**; 350-362.
1017
1018 Stroebe, M. and H. Schut. The dual process model of coping with bereavement Rationale
1019 and description. *Death Studies*. 1999**: 3**; 197-224.
1020
1021 Su, T.-J. and Y.-C. Chen. Transforming hope: The lived experience of infertile women who
1022 terminated treatment after in Vitro Fertilization failure. *Journal of Nursing Research*.
1023 2006**: 1**; 46-53.
1024

Sydsjö, G., J. Vikström, M. Bladh, B. Jablonowska and A. S. Svanberg. Men report good mental health 20 to 23 years after in vitro fertilisation treatment. *BMC Public Health*. 2015**:** 1175.

Taylor, S. E. Adjustment to threatening events: A theory of cognitive adaptation. *American Psychologist*. 1983**: 11**; 1161-1173.

Terry, D. J. and G. J. Hynes. Adjustment to a low-control situation: reexamining the role of coping responses. *Journal of Personality and Social Psychology*. 1998**: 4**; 1078-1092.

Thoits, P. A. Identity structures and psychological well-being: Gender and marital status comparisons. *Social Psychology Quarterly*. 1992**: 3**; 236-256.

Thomas, J. and A. Harden. Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*. 2008**: 45**;

Throsby, K. No one will ever call me Mummy: Making sense of the end of IVF treatment. 2001. London School of Economics - Gender Institute, London.

Troude, P., G. Santin, J. Guibert and E. de la Rochebrochard. Seven out of 10 couples treated by IVF achieve parenthood following either treatment, natural conception or adoption. *Reproductive BioMedicine Online*. 2016**: 5**; 560-567.

Verhaak, C. M., J. M. Smeenk, A. van Minnen, J. M. Kremer and F. W. Kraaimaat. Predicting emotional response to unsucessful fertility treatment: A prospective study. *Journal of Behavior Medicine*. 2005**: 2**; 181-190.

Verhaak, C. M., J. M. Smeenk, A. W. M. Evers, J. M. Kremer, F. W. Kraaimaat and D. M. Braat. Women's emotional adjustment to IVF: A systematic review of 25 years of research. *Human Reproduction Update*. 2007**: 1**; 27-36.

Verhaak, C. M., J. M. Smeenk, M. J. Nahuis, J. A. M. Kremer and D. D. Braat. Long-term psychological adjustment to IVF/ICSI treatment in women. *Human Reproduction*. 2007**: 1**; 305-308.

Volgsten, H., A. Skoog-Svanberg and P. Olsson. Unresolved grief in women and men in Sweden three years after undergoing unsuccessful in vitro fertilization treatment. *Acta Obstericia et Gynecologica Scandinavica*. 2010**: 10**; 1290-1297.

Weaver, S. M., E. Clifford, D. M. Hay and J. Robinson. Psychosocial adjustment to unsuccessful IVF and GIFT treatment. *Patient Education and Counseling*. 1997**: 1**; 7-18.

Wirtberg, I., A. Möller, L. Hogström, S.-E. Tronstad and A. Lalos. Life 20 years after unsuccessful infertility treatment. *Human Reproduction*. 2007**: 2**; 598-604.

Wischmann, T., K. Korge, H. Scherg, T. Strowitzki and R. Verres. A 10-year follow up study of psychosocial factors affecting couples after infertility treatment. *Human Reproduction*. 2012**: 11**; 3226-3232.

Wrosch, C., M. F. Scheier, G. E. Miller, R. Schulz and C. S. Carver. Adaptive self-regulation of unattainable goals: Goal disengagement, goal reengagement, and subjective wellbeing. *Personality and Social Psychology Bulletin*. 2003**: 12**; 1494-1508.

Yli-Kuha, A.-N., M. Gissler, R. Klemetti, R. Luoto, E. Koivisto and E. Hemminki. Psychiatric disorders leading to hospitalization before and after infertility treatments. *Human Reproduction*. 2010**: 8**; 2018-2023.

Zegers-Hochschild, F., G. D. Adamson, J. de Mouzon, O. Ishihara, R. Mansour, K. Nygren, E. Sullivan and S. Vanderpoel. International Committee for Monitoring Assisted Reproductive Technology (ICMART) and the World Health Organization (WHO) revised glossary of ART terminology, 2009. *Fertility and Sterility*. 2009**: 5**; 1520-1524.

Table 1. Characteristics of the nine included quantitative studies

Authors, year Country	**Fertility treatment**	**Failed treatment group**	**Control group**	**Existence of children in failed treatment group (**YES, NO**)**[a]	**Years since last treatment cycle**	**Design**	**Response /attrition rate**	**Quality**	**Mental-health**[b]	**Wellbeing**[b]
Gameiro *et al* 2014 The Netherlands	Fertility drugs/ IUI/ AI/ IVF/ ICSI/ gamete donation	1493 women	5655 women with children[b]	NO	range 11-17	Cross-sectional	60.4%	High	✓	✕
Hammarberg *et al* 2001 Australia	IVF/ ET/ GIFT	65 women	51 women successful treatment	YES	range 2.5-3.5	Cross-sectional	55%	Moderate	✕	✓
Johansson *et al* 2010 Sweden	IVF/ ET/ ICSI	149 women + 121 men	118 women + 93 men with children	YES	range 4-5.5	Cross-sectional	67.5%	Moderate	✓	✓
Kuivasaari-Pirinen *et al* 2014 Finland	IVF/ ICSI	209 women	296 women successful treatment	YES	range 1.3-11.7	Longitudinal	100% / 54.7%	Moderate	✕	✓
Leiblum *et al* 1998 USA	IVF	18 women	41 women successful treatment	NO	range 2-13	Cross-sectional	31%	Moderate	✓	✓
Sydsjö et al 2015 Sweden	IVF	62 men	230 men with children	NO	20	Cross-sectional	66.5%	Moderate	✓	✕
Verhaak *et al* 2007 The Netherlands	IVF/ ICSI	39 women	68 women successful treatment	NO	range 3-5	Longitudinal	84% / 78%	High	✓	✕
Weaver *et al* 1997 UK	IVF/ GIFT	21 couples	20 couples successful treatment	YES	range 1.25-2.25	Cross-sectional	95%	Moderate	✓	✓
Wischmann *et al* 2012 Germany	IUI / IVF/ ICSI	40 couples	91 couples successful treatment	NO	10 (since first consultation)	Cross-sectional	35.7%	Moderate	✕	✓

Legend. SC = spontaneous conception; IVF = in-vitro fertilisation; ICSI = intra-cytoplasmic sperm injection; FET = frozen thawed embryo transfer cycle; IUI = intrauterine insemination; GIFT = gamete intrafallopian transfer; AI = artificial insemination; ET = embryo transfer; NR = not reported; [a] children could be adopted or conceived spontaneously or with previous fertility treatment; [b]✓= reported; ✕ = not reported. Note: 180 (2.5%) patients who participated in the Gameiro et al 2014 study also participated in the Verhaak et al 2007. Due to the extremely low sample overlap we still included the two studies.

Table 2. Characteristics of the nine included qualitative studies

Author, year Country	**Goal**	**Failed treatment group**	**% with children**	**Design**	**Years since treatment**	**Data collection method**	**Analytical method**	**Quality**
Boden 2007 UK	Extend the understanding of how women and their partners experience infertility and come to terms with their childlessness	18 couples + 15 women	NR	Cross-sectional	5	Interviews	Thematic analysis	High
Daniluk 2001 Canada	Understand how couples make sense of their infertility and reconstruct their lives when faced with the permanence of their biological childlessness	37 couples	46	Longitudinal	1, 1.83 & 2.66	Semi-structured interviews	Phenomenological analysis	High
Johansson & Berg 2004 Sweden	Describe women's experiences of ongoing childlessness 2 years after the end of IVF treatment	8 women	13	Cross-sectional	2	Unstructured interviews	Phenomenological analysis	Moderate
Lee *et al* 2009 China	Provide an in-depth description of the gains perceived by Chinese men and women and how they re-constructed their lives after unsuccessful IVF treatment	4 couples (4 women + 4 men) + 6 women	36	Cross-sectional	range 0.5 – 3	Semi-structured interviews (joint and individual)	Grounded Theory	High
McCarthy 2008 USA	Understand the lived experience of infertility for women in the aftermath of unsuccessful medical treatment	22 women	32	Cross-sectional	3.9, SD 3.2	Interviews	Hermeneutic–phenomenology	High
Su & Chen 2006 Taiwan	Explore the lived experiences of infertile women who terminated treatment after IVF failure	24 women	NR	Cross-sectional	1.33, SD 0.33 range 1-1.92	Telephone interviews	Phenomenological analysis	High
Throsby 2001 UK	Identify the factors which inform the decision to end treatment; consider the ways in which the unsuccessful engagement with IVF marked both the experience of infertility and the participants' own perceptions of that technology; explore the ways in which those feelings might change over time and the influence they might have on subsequent choices	13 couples + 15 women	54	Cross-sectional	2	Interviews	Discourse analysis	High
Volgsten *et al* 2010 Sweden	Explore the experience of undergoing unsuccessful in vitro fertilization (IVF) treatment and of remaining childless 3 years after IVF in both women and men	7 couples + 3 women + 2 men	21	Cross-sectional	3.17	Semi-structured interviews	Content analysis	High
Wirtberg *et al* 2007 Sweden	Obtain an increased knowledge and deeper insight into the long-lasting effects of, and coping with, involuntary childlessness for a group of women who had sought help for infertility and had completed infertility treatment over 20 years before	14 women	0	Cross-sectional	20	Semi-structured interviews	NR	Moderate

Legend. IVF = In Vitro Fertilization; NR = not reported.

Table 3. Subgroup and meta-regression analyses of mean differences between the failed treatment and control groups for mental-health and wellbeing.

	Mental-health					Wellbeing				
Subgroup analysis	***k***	**Effect size**	**95% CI LL**	**95% CI UL**	**χ^2**	***k***	**Effect size**	**95% CI LL**	**95% CI UL**	**χ^2**
Comparison groups					**12.534*** **					NA
Based on treatment outcome (failed, successful)	4	-0.648	-0.828	-0.467		5	-0.324	-0.454	-0.193	
Based on parenthood status (no children, children)	2	-0.108	-0.168	-0.048		0				
Existence of children in failed treatment group					0.050					0.679
No	5	-0.484	-0.803	-0.165		2	-0.228	-0.491	0.036	
Yes	1	-0.589	-1.457	0.279		3	-0.355	-0.505	-0.205	
Quality					1.132					NA
Moderate	4	-0.608	-0.953	-0.263		5	-0.324	-0.454	-0.193	
High	2	-0.323	-0.719	0.073		0				
Meta-regression		**Coefficient**	**95% CI LL**	**95% CI UL**	**z**		**Coefficient**	**95% CI LL**	**95% CI UL**	**z**
Years since treatment	6	0.025	-0.003	0.052	1.777	5	0.055	-0.008	.117	1.712

Legend. *P<0.05, **P<0.01, ***P<0.001, ART = assisted reproductive techniques, k=number of studies, CI = confidence intervals, LL = lower limit UL = upper limit, NA = not applicable as there were not enough studies in the group, bold indicates P<0.05

Table 4. First and second order themes developed to describe patients' experiences after failed fertility treatment. Second order themes are presented in bold with a grey background. For each of the second order themes the associated first order themes are listed below.

Individual and relational adjustment	**Social adjustment**	**Fertility care perceptions and needs**
<u>Immediately after ending treatment</u> • Intense grief (especially for women) • Anger directed at 'everything & everyone' • Treatment strengthens partnership • Men stoic & supportive of their partner • Sexual difficulties <u>As time since treatment passed</u> • Resilience & personal & spiritual growth	• Social isolation • Parenthood loss hinders partner and social relationships • Social representations of parenthood hinder adjustment • Support from selected family members & friends • Social avoidance	<u>Immediately after ending treatment</u> • Feeling abandoned\lack of support from clinic • Need for psychosocial & sexual care/counselling • Acknowledgement of reproductive medicine limitations <u>As time since treatment passed</u> • Positive outlook on ending treatment
Meaning making	**Acceptance**	**Pursuit of new life-goals**
• Need to find meaning in life (past efforts, new identity & future) • Questioning of traditional notions of marriage, family… in light of parenthood loss • Re-evaluation of life values & priorities • Acceptance and meaning making are interdependent	<u>Immediately after ending treatment</u> • Inability/unwillingness to accept childlessness <u>As time since treatment passed</u> • Acceptance easier when… [all options explored, efforts acknowledged, no pressure for parenthood, good medical advice, poor prognosis, reasons for failure identified] • Acceptance associated with new views of the world • Non-acceptance associated with worse adjustment • Acceptance associated with better adjustment	• New fulfilling goals/activities include… [adoption, caring for other children (e.g., siblings) or pets, supporting & advocating for the rights of the infertile, travelling, career/professional development] • 'Moving on' requires active effort • Need to 'move on' with life • Acceptance and pursuit of new life-goals are interdependent.

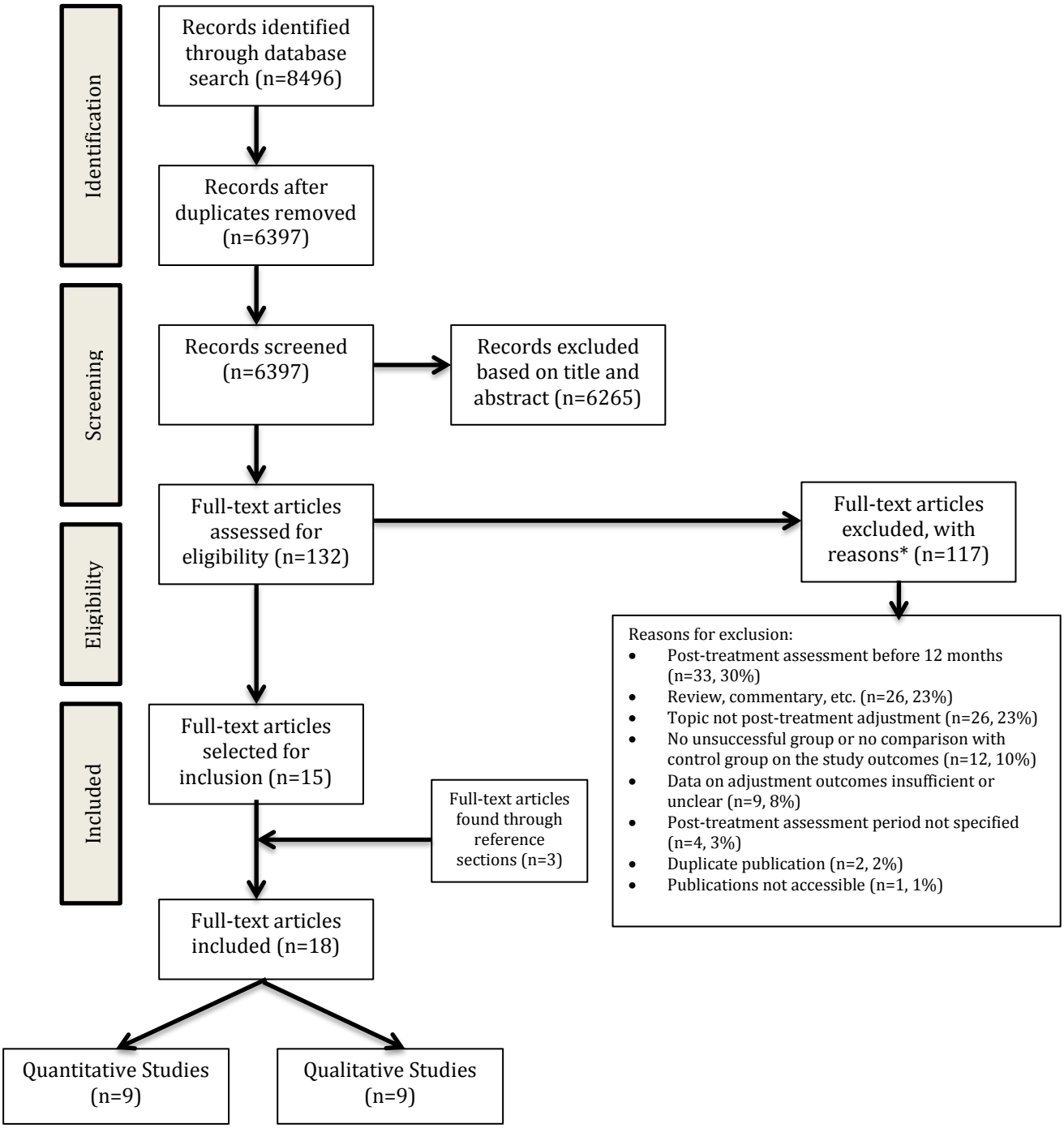

Figure 1. Decision flow chart for identified studies. * For detailed exclusion strategy see Table S2 of the supplementary material.

Mental-health

Study names	Hedges's g	Standard error	Lower limit	Upper limit	Z-Value	p-Value
Gameiro et al 2014	-0.093	0.031	-0.155	-0.032	-2.968	0.003
Johanssen et al 2010	-0.742	0.186	-1.107	-0.377	-3.985	0.000
Leiblum et al 1998	-0.753	0.287	-1.317	-0.190	-2.620	0.009
Syydsjo et al 2011	-0.415	0.144	-0.696	-0.133	-2.884	0.004
Verhaak et al 2007	-0.596	0.123	-0.838	-0.355	-4.833	0.000
Weaver et al 1997	-0.589	0.299	-1.175	-0.004	-1.972	0.049
	-0.496	0.151	-0.791	-0.200	-3.283	0.001

Wellbeing

Study name	Hedges's g	Standard error	Lower limit	Upper limit	Z-Value	p-Value
Hammarberg et al 2001	-0.502	0.189	-0.872	-0.132	-2.658	0.008
Johansson et al 2010	-0.355	0.183	-0.715	0.004	-1.939	0.052
Kuivasaari-Pirinen et al	-0.300	0.091	-0.478	-0.122	-3.305	0.001
Leiblum et al 1998	-1.168	0.299	-1.754	-0.582	-3.907	0.000
Weaver et al 1997	-0.484	0.222	-0.919	-0.048	-2.177	0.029
Wischmann et al 2012	-0.080	0.197	-0.466	0.307	-0.404	0.686
	-0.415	0.108	-0.627	-0.203	-3.840	0.000

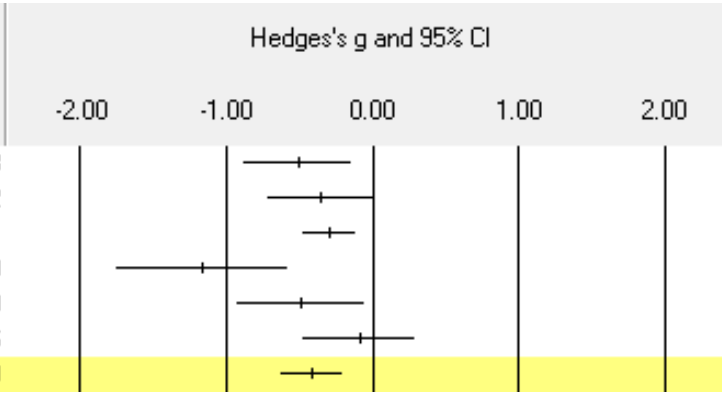

Figure 2. Standardized mean differences (Hedge's g, Standard Error (SE), 95% Confidence Intervals (CIs), Z value and P value) between failed treatment and control groups for mental-health and wellbeing. Negative Hedge's values indicate worse adjustment for the failed treatment group than the control group.

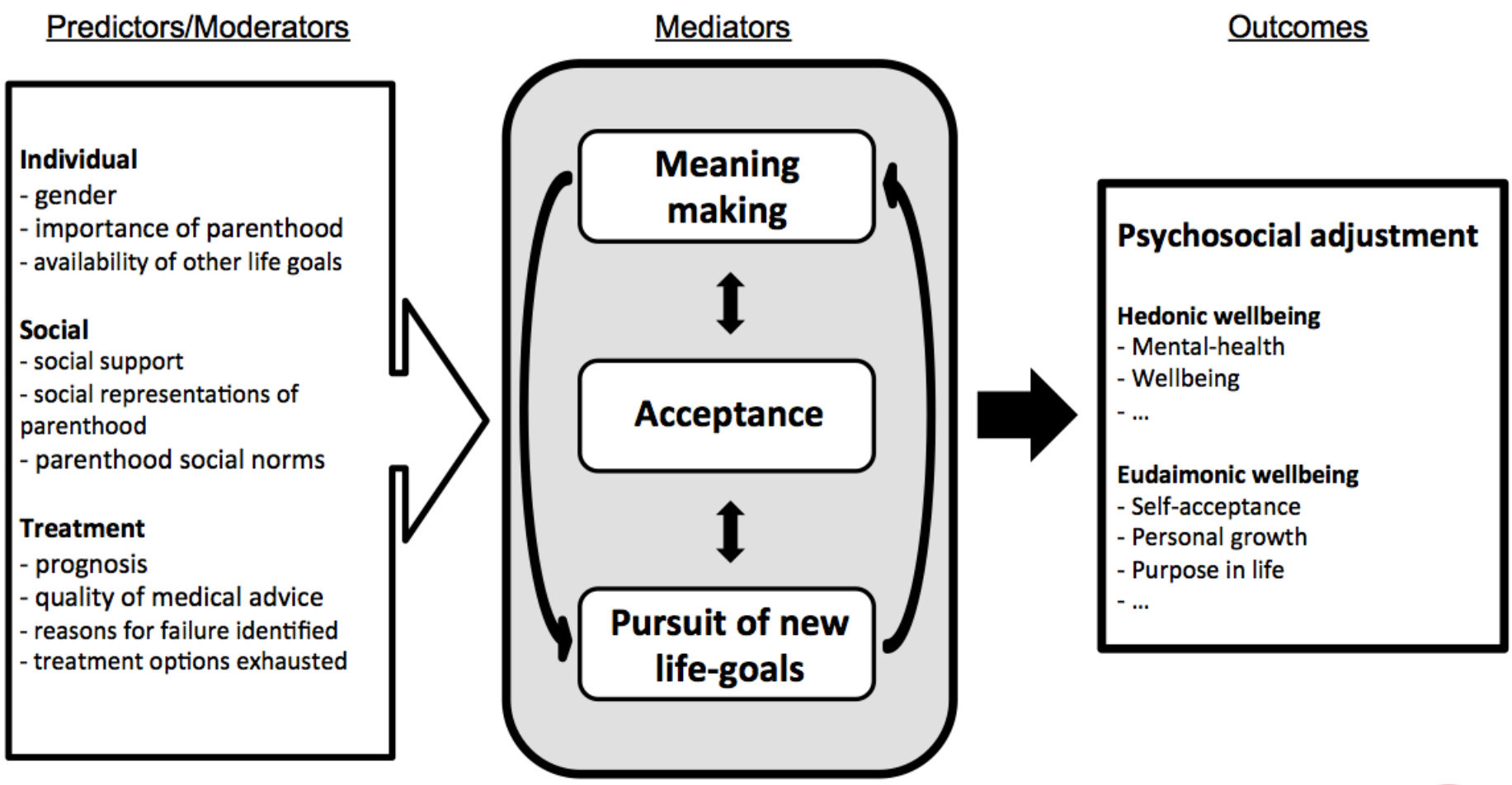

Figure 3. A schematic representation of the Three Tasks Model of Adjustment to Unmet Parenthood Goals