Depression and state anxiety scores during assisted reproduction treatment are associated with outcomes: a meta-analysis

Author names

S, Purewal[1,4]., S.C. E. Chapman[2]., & O.B.A van den Akker[3,] .

[1,4]. Corresponding author: Institute of Psychology, Faculty of Education, Health and Wellbeing, University of Wolverhampton, Wolverhampton, WV1 1AD, UK.

Satvinder.purewal@wlv.ac.uk

[2.] Department of Pharmacy & Pharmacology, University of Bath, Claverton Down Road, Bath, BA2 7AY UK

[3.] Department of Psychology, Faculty of Science and Technology, Middlesex University, London, NW4 4BT, UK

Abstract

This meta-analysis investigated whether state anxiety and depression scores during assisted reproductive technology (ART) treatment and changes in state anxiety and depression scores between baseline and during ART treatment are associated with treatment outcomes. PubMed, PsycInfo, Embase, ScienceDirect, Web of Science and Scopus were searched for studies to include in the meta-analysis. Meta-analytic data were analysed using random effects models to estimate standardised mean differences. 11 studies (2202 patients) were included. Women who achieved a pregnancy had significantly lower depression scores during treatment than women who did not become pregnant -0.302 (95% CI: -0.551 - -0.054, z = -2.387, p = 0.017; I^2= 77.142%, p = 0.001). State anxiety scores were also lower in women who became pregnant -0.335 (95% CI: -0.582 - -0.087: z=-2.649, p=0.008; I^2 =81.339%, p = 0.001). However, changes in state anxiety (d=-0.056; 95% CI: -0.195 - 0.082, z = -0.794; I^2= 0.00%) and depression scores (d=-0.106; 95% CI: -0.296 - 0.085, z = -1.088; I^2= 0.00%) from baseline to treatment were not associated with ART outcomes. Clinics should aim to promote better psychosocial care for patients to help them manage the psychological and physical demands ART treatment, giving realistic expectations.

Introduction

Infertility is experienced by nearly one in six couples and many of these couples go on to seek assisted reproductive technology (ART) treatment to help them conceive (Farquhar et al., 2015). However, it is estimated that only a quarter of women will get pregnant after a single ART cycle in the UK (Kushnir et al., 2017), so most couples will experience negative pregnancy results and repeat treatment. A recent meta-analysis found that rates of depression and anxiety increased after ART treatment failure, but depression decreased after successful treatment (Milazzo et al., 2016). Another recent meta-analysis also found no increased risk of depressive symptoms in women after they conceived using fertility treatment compared to those had spontaneous pregnancy (Gressier et al., 2015).

However, common psychological reactions during ART include stress, anxiety and depression (Eugster & Vingerhoets, 1999). Many women experience ART treatments as stressful, with stress and state anxiety levels increasing during oocyte retrieval and pregnancy test stages (Boivin et al., 1995; An et al., 2013). Turner et al. (2013) found that women with higher stress and anxiety levels on the day prior to oocyte retrieval had a lower chance of obtaining positive pregnancy results. It is perhaps not surprising that some women drop out of treatment due to a variety of psychological and physical burdens (Gameiro et al., 2012).

Moreover, meta-analyses have found small but significant associations between baseline (before ART treatment has started) depression, state and trait anxiety scores and reduced pregnancy chances with ART (Purewal et al., 2017a; Separate paper by Authors under review; Mattheisen et al., 2011). However, Boivin et al. (2011) found no impact of baseline psychological distress (combined depression and anxiety scores) on ART treatment success.

To date, no meta-analysis has investigated whether depression and anxiety scores during ART treatment and changes in levels of anxiety or depression from baseline (pre-treatment) to treatment are associated with ART outcomes, despite studies (e.g., Boivin et al., 1995; An et al., 2013) reporting increases in anxiety and depression rates over the course of treatment. The aims of this meta-analysis are therefore to a) investigate the impact of state anxiety and depression scores during ART treatment on ART treatment outcomes and b) investigate whether changes in levels of state anxiety and depression from baseline to during treatment predict ART treatment outcomes.

Methods

This meta-analysis is part of a larger project that also investigated whether baseline psychological distress is associated with ART outcomes (Purewal et al, 2017a) and if investigated lifestyle (smoking and alcohol use) and BMI are predictors of ART outcomes (Purewal et al., 2017b.

The systematic review and meta-analysis was performed following PRISMA and MOOSE guidelines (Stroup, et al., 2000).

Eligibility criteria

Eligible studies were considered if they reported:

a) Prospective studies which reported maternal depression and state anxiety scores during ART treatment (e.g., during oocyte retrieval or the day of embryo transfer) and ART outcomes, ideally with baseline measures (before treatment has started) of depression and state anxiety scores. Studies which reported depression and state

anxiety scores after embryo transfer were excluded. Trait anxiety scores were excluded because we were interested in the effects of treatment stage on state anxiety scores, not stable trait scores;

b) Studies were included if they used a standardised psychological measure (e.g., BDI – Beck's Depression Inventory and STAI – State-Trait Anxiety Inventory-State scale) reporting continuous or categorical (cut off score) data;

c) Studies were included if they reported only original data, reported live birth rates or pregnancy outcome data and ART treatments were included (e.g., IVF, ICSI, ZIFT). Other exclusion criteria were if it was not possible to calculate unadjusted effect sizes for predictor variables (e.g. predictor data grouped by outcome, only adjusted data reported, percentages without numbers reported c) and therefore meta-analysis of unadjusted effect sizes could not be achieved.

Information sources and search

Six bibliographic databases were searched: PubMed, PsycInfo, Embase, ScienceDirect, Web of Science and Scopus. In PubMed, we used the following keywords in keywords and abstracts: ("Pregnancy"[Mesh] OR "Pregnancy" OR "pregnant" OR "live birth" OR "birth rate") AND ("IVF" OR "intracytoplasmic" OR "intracytoplasmic sperm injection" OR "in vitro fertilization" OR "ICSI" OR "assisted reproductive technology" OR "in vitro fertilisation") AND ("psychological stress" OR "depressive disorder" OR "anxiety " OR "anxiety disorder " OR "adjustment disorder" OR "emotions" OR "psychosomatic medicine "OR "psychological adaption" OR "distress" OR "depression" OR "stress" OR "occupation stress" OR "stressful life events" OR "major life events" OR "stressors"). The searches were limited to 1979/01/01-

present (November 2016) and humans. Hand searches of references cited in relevant papers were also conducted.

Study selection, data collection process and data items

Using PRISMA guidelines (Moher et al., 2009) all authors independently screened titles, abstracts and full-text reports and disagreements were resolved by discussion between all authors. Data were extracted and independent (depression and state anxiety scores at baseline and during ART treatment) and dependent variables (live birth or pregnancy) and sample sizes were recorded. When two or more dependent variables were reported (e.g., serum pregnancy, clinical pregnancy and live birth), the data considered 'gold standard' (Maheshwari et al., 2008) was recorded (in this case, live birth; however, no study reported live birth data, so clinical pregnancy rates were used). Other data were also extracted, such as patient characteristics (e.g., average female age, whether they are first time ART users or previously used ART, number of oocytes retrieved, percentage with primary infertility); treatment characteristics (e.g., treatment location, ICSI use (all/some vs none used ICSI), average number of embryos transferred, single or multiple cycle recorded, pregnancy verification (pregnancy test vs ultrasound scan) and study characteristics (e.g., publication date; design of study).

Summary measures and synthesis of results

The meta-analyses were performed on Comprehensive Meta-Analysis V2 (Borenstein et al., 2005) using weighted effect sizes with a random effects model. The extracted data (e.g., state anxiety and depression mean scores, standard deviations and sample sizes for pregnant and not pregnant groups at two time points; Time 1) baseline and Time 2) during ART treatment were inputted. These data were converted into standardised mean differences and used to

compare women who became pregnant and women who did not. Outliers were identified as studies with residuals greater than 1.96 and they were removed from the analysis as recommended.

Heterogeneity

The I^2 statistics were used to quantify heterogeneity, with 50-90% representing potentially substantial heterogeneity (Deeks et al. 2009). As recommended by Deeks et al. (2009), we intended to conduct moderator analyses to investigate significant heterogeneity if we had 10 studies or more which provided data on potential moderators (e.g., average female age, duration of subfertility, bFSH and number of oocytes , see Van Loendersloot et al. 2010). However, apart from mean maternal age for state anxiety during ART (where we had 10 studies per confounder variable), there were insufficient numbers of studies to analyse moderators hence these analyses could not be performed.

Risk of bias

To assess the quality of studies, the Newcastle-Ottawa Scale (NOS) (Wells et al., 2009) was used. SP and OvdA independently assessed the quality of each selected study and cross-checked with each other to reach 100% consensus. The scale awarded a maximum of nine stars to each study: four stars for the adequate selection of cases and controls, two stars for comparability of cases and controls, and three stars for the adequate ascertainment of the exposure in both the case and control groups. High quality was defined as scoring at least seven stars; medium quality as scoring five or six stars and low quality as four or less.

Risk of bias analyses were conducted to examine whether effects were robust under different methodological assumptions. These included when ultrasound was used to diagnosis pregnancy and NOT pregnancy test; when only first time ART user's data is included; when results from a single cycle are used (not multiple cycles); when only IVF treatments are used; when only psychological data is collected before or during the oocyte retrieval period and not after; when only high quality studies were included; and when studies were recent (studies published from 2010 onwards were considered recent).

Publication bias

We tested for publication bias by examining funnel plots for evidence of asymmetry. Asymmetric funnel plots can occur because of biased publication strategies e.g. if small, imprecise studies are only published if they support a particular hypothesis, but are not published if they do not support the hypothesis (Sterne & Egger, 2001). We used Duval and Tweedie's trim and fill method (Duval & Tweedie, 2000), which imputes studies where evidence of asymmetry is present and tested for the significance of these effects using Egger's t-test.

Results

Study selection

The screening process is summarised in the study PRISMA flow chart (Fig 1). A total of 11 studies (An et al., 2011; Gourounti et al., 2011; Gurhan et al., 2009; Hashemi et al., 2012; Karlidere et al., 2008: Li et al., 2011: Lintsen et al., 2009; Saravelos et al., 2016; Taguchi et al., 2015; Turner et al., 2013; Zaig et al., 2012) were included in the meta-analysis.

Study characteristics

An overview of the study characteristics is presented in Table 1. Data from 2202 4 patients were included in the meta-analyses. Most studies collected psychological data before or during the oocyte retrieval period, (An et al., 2011; Gourounti et al., 2011; Gurhan et al., 2009; Hashemi et al., 2012; Li et al., 2011: Lintsen et al., 2009; Saravelos et al., 2016; Turner et al., 2013; Zaig et al., 2012) with only Karlidere et al.'s (2008) and Taguchi et al.'s (2015) collecting data on the day before and on the day of embryo transfer, respectively.

Risk of bias within studies: The quality of the studies was high or median (An et al., 2011; Gourounti et al., 2011; Gurhan et al., 2009; Hashemi et al., 2012; Karlidere et al., 2008: Li et al., 2011: Lintsen et al., 2009; Saravelos et al., 2016; Taguchi et al., 2015; Turner et al., 2013; Zaig et al., 2012) with none scoring a 'low' score (4 or less stars). All studies used pregnancy diagnosis (An et al., 2011; Gourounti et al., 2011; Gurhan et al., 2009; Hashemi et al., 2012; Karlidere et al., 2008: Li et al., 2011: Lintsen et al., 2009; Saravelos et al., 2016; Taguchi et al., 2015; Turner et al., 2013; Zaig et al., 2012), with pregnancy ultrasound scan used to verify pregnancy in most cases (An et al., 2011; Gourounti et al., 2011; Karlidere et al., 2008; Li et al., 2011; Lintsen et al., 2009; Saravelos et al., 2016; Turner et al., 2013; Zaig et al., 2012). Most studies were published after 2010 (An et al., 2011; Gourounti et al., 2011; Hashemi et al., 2012; Li et al., 2011; Saravelos et al., 2016; Taguchi et al., 2015; Turner et al., 2013; Zaig et al., 2012) and used the STAI to measure state anxiety ((An et al., 2011; Gourounti et al., 2011; Gurhan et al., 2009; Hashemi et al., 2012; Karlidere et al., 2008: Lintsen et al., 2009; Saravelos et al., 2016; Turner et al., 2013; Zaig et al., 2012)), with half of the depression studies using BDI ((An et al., 2011; Gurhan et al., 2009; Karlidere et al., 2008: Saravelos et al., 2016). The most common ART procedure was IVF (Gurhan et al., 2009; Li et al., 2011; Taguchi

et al., 2015; Turner et al., 2013; Zaig et al., 2012) or IVF/ICSI (An et al., 2011; Gourounti et al., 2011; Hashemi et al., 2012; Karlidere et al., 2008; Lintsen et al., 2009; Saravelos et al., 2016).

Only five studies reported baseline data (An et al 2011; Gurhan et al 2009: Li et al 2011; Linsten et al., 2009; and Turner et al., 2013). All baseline data are reported in Table 1. Only two studies (An et al. 2011; Linsten et al., 2009) reported statistical comparisons between patient's anxiety baseline (time 1) scores and anxiety during ART treatment (time 2) by pregnancy outcomes and both studies found no significant anxiety score gains from baseline to during ART treatment. Table 1 also includes the main summary of each study's results. Each study's calculated standardised means differences between pregnant and not pregnant women for depression and state anxiety scores and changes in depression and state anxiety scores from baseline to during treatment are presented in forest plots (figures 2-5).

Synthesis of Results

Depression During ART

Results of individual studies and synthesis of results: Depression was measured in eight studies (An et al 2011; Gourounti et al 2011: Gurhan et al 2009: Karlidere et al 2008: Li et al 2011; Saravelos et al 2016; Taguchi et al 2015; Zaig et al 2012). Results revealed that women who achieved pregnancy reported lower depression mean scores than women who did not achieve a pregnancy -0.302 (95% CI: -0.551 - -0.054, z = -2.387, p = 0.017) with significantly high levels of heterogeneity (I^2= 77.142%, p = 0.001). See Figure 2 for forest plots.

Risk of bias

The effects of depression remained consistent in the sensitivity analyses, which considered studies which only used first time ARTpatients (An et al., 2011; Gurhan et al., 2009; Karlidere et al., 2008; Li et al., 2011; Lintsen et al., 2009), reported pregnancies diagnosed by ultrasound s, single cycle results only (An et al 2011; Gourounti et al 2011; Gurhan et al 2009 Karlidere et al 2008; Li et al 2011; Taguchi et al 2015; Zaig et al 2012), data collected before or during oocyte retrieval (An et al 2011; Gourounti et al 2011; Gurhan et al 2009; Li et al 2011; Saravelos et al 2016; Zaig et al 2012) and high quality studies (An et al., 2011; Gourounti et al., 2011; Karlidere et al., 2008; Li et al., 2011; Saravelos et al., 2016; Zaig et al., 2012) (See Table 2). However, the effect became smaller when the analysis was conducted in fewer studies reporting IVF outcomes (not ICSI) (Gurhan et al 2009; Li et al 2011; Taguchi et al 2015; Zaig et al 2012) and recent studies (An et al 2011; Gourounti et al 2011; Li et al 2011; Saravelos et al 2016; Taguchi et al 2015; Zaig et al 2012).

Publication bias

Data analyses generally indicated low levels of publication bias risk for depression. The trim and fill data analyses revealed only 1 additional study would be needed to ensure the funnel plot was generally symmetrical and Egger's regression intercept was not significant (t(6) =1.77).

State Anxiety During ART

Results of individual studies and synthesis of results: State anxiety was measured in ten studies (An et al 2011; Gourounti et al 2011; Gurhan et al 2009; Hashemi et al 2012; Karlidere et al 2008; Li et al 2011; Lintsen et al 2009; Saravelos et al 2016; Turner et al 2013; Zaig et al 2012). Results revealed that women who achieved a pregnancy reported lower state anxiety

mean scores than women who did not achieve a pregnancy -0.335 (95% CI: -0.582 - -0.087: z=-2.649: p=0.008) with evidence of high levels of significant heterogeneity (I^2 =81.339%, p = 0.001). See figure 3 for forest plot of results from individual studies and synthesis of results. Mixed effects meta-regression was performed to test whether mean maternal age moderated the effect of state anxiety on outcomes. This revealed no effect of age (slope = -0.06, 95% CI [-0.20, 0.08]. We were not able to perform moderator analyses on any other variables due to the small numbers of studies (less than 10 studies per moderator variable).

Risk of bias analysis

The effect for state anxiety was robust under different methodological conditions, such as studies which diagnosed pregnancy by ultrasound (not pregnancy test) (An et al., 2011; Gourounti et al., 2011; Karlidere et al., 2008; Li et al., 2011; Lintsen et al., 2009; Saravelos et al., 2016; Turner et al., 2013; Zaig et al., 2012), single cycle outcomes (An et al 2011; Gourounti et al 2011; Gurhan et al 2009; Hashemi et al 2012; Karlidere et al 2008; Li et al 2011; Lintsen et al 2009; Turner et al 2013; Zaig et al 2012), data collected before or during oocyte retrieval (An et al 2011; Gourounti et al 2011; Gurhan et al 2009; Hashemi et al 2012; Li et al 2011; Lintsen et al 2009; Saravelos et al 2016; Turner et al 2013; Zaig et al 2012) and high quality studies (An et al., 2011; Gourounti et al., 2011; Hashemi et al., 2012; Karlidere et al., 2008; Li et al., 2011; Lintsen et al., 2009; Saravelos et al., 2016; Zaig et al., 2012). However, the effects became smaller for first time ART users (An et al., 2011; Gurhan et al., 2009; Karlidere et al., 2008; Li et al., 2011; Lintsen et al., 2009), IVF only (Gurhan et al 2009; Li et al 2011; Turner et al 2013; Zaig et al 2012), and recent studies only An et al 2011; Gourounti et al 2011; Hashemi et al 2012; Li et al 2011; Saravelos et al 2016; Turner et al 2013; Zaig et al 2012). See Table 1.

Publication bias

Some evidence of publication bias was present for state anxiety. The trim and fill data analyses for state anxiety revealed 2 additional studies were needed to make the funnel plot symmetrical. However, state anxiety's Egger's regression intercept was not significant (t(8) =2.235).

Changes in depression scores from baseline to during ART treatment

Results of individual studies and synthesis of results: Depression was reported at baseline and during ART treatment in three studies (An et al 2011; Gurhan et al 2009: Li et al 2011). Results revealed that changes in reported depression scores from baseline (time 1) to during ART treatment (time 2) were not associated with ART outcomes -0.106 (95% CI: -0.296 - 0.085, z = -1.088) with low and non-significant heterogeneity (I^2= 0.00%). See Figure 4 for forest plots.

Risk of bias analysis

The non-significant effects of depression remained consistent in the sensitivity analyses, (see table 2).

Publication bias

There was limited evidence of publication bias, possibly due to the small number of included studies. The trim and fill data analyses for changes in depression revealed no additional studies were needed to make the funnel plot symmetrical. Egger's regression intercept was also not significant (t(1) =0.38).

Changes in state anxiety scores from baseline to during ART treatment

Results of individual studies and synthesis of results: State anxiety was reported at baseline (time 1) and during ART treatment (time 2) in five studies (An et al 2011; Gurhan et al 2009; Li et al 2011; Lintsen et al 2009; Turner et al 2013). Results revealed that changes in reported state anxiety mean scores from baseline to during ART treatment were not associated with ART outcomes -0.056(95% CI: -0.195 - 0.082, z = -0.794) with low, non-significant heterogeneity (I^2< 0.00%). See Figure 5 for forest plots.

Risk of bias analysis

The small effect for state anxiety was generally consistent under different methodological conditions (see table 2).

Publication bias

Some evidence of publication bias was present for changes in state anxiety. The trim and fill data analyses for state anxiety revealed 2 additional studies were needed to make the funnel plot symmetrical. However, the state anxiety's Egger's regression intercept was not significant (t(3) =3.28).

Discussion

Summary of evidence

This is the first meta-analysis that investigated the relationship between anxiety and depression experienced during ART and ART outcomes. The analysis of the available research has shown that depression and state anxiety scores reported during ART treatment stages

(most studies reported before or during the oocyte retrieval stage) are significantly associated with ART treatment outcomes. Although the significant effect sizes for depression (-0.302) and state anxiety (-0.335) were small, they were generally consistent under different methodological conditions and there was little evidence of publication bias. However, there was no evidence that changes in depression or state anxiety scores from baseline (time 1) to during ART treatment (time 2) were associated with ART treatment outcomes. Heterogeneity was high for depression and state anxiety during ART suggesting that the effects varied between studies. However, our findings were consistent under different methodological assumptions, although the number of studies included in the latter analyses were small.

Previous meta-analyses have found small associations between baseline depression, state and trait anxiety and stress and reduced pregnancy chances with ART (Purewal et al., in press; Mattheisen et al., 2011). However, we found baseline depression (d=-0.177) and state anxiety (d=-0.096) demonstrated a weaker effect on ART outcomes (Purewal et al., 2017a) than the effect size found reported in this current paper. It therefore appears that depression and state anxiety taken at certain stages of the ART treatment cycle (i.e., before or during-mainly the oocyte retrieval stage) are relevant factors in predicting ART outcomes, but there is no evidence that changes in levels of depression or state anxiety from baseline to during treatment is associated with ART outcomes. However, the numbers of studies in the baseline vs during ART treatment analyses were relatively small and more extensive investigations are needed for definitive answers.

These findings are interesting and help to frame future investigations. High levels of heterogeneity were obtained for the effect of depression and state anxiety scores during ART, but we were unable to fully investigate the source of heterogeneity because there were insufficient numbers of studies to test for moderator effects (such as BMI, number of oocytes retrieved, poor responders)c) (Deeks et al., 2009). However, we were able to investigate the moderating effect of age on state anxiety and ART outcomes and found no impact, but the sample size was small (K=10). These results are interesting and highlight there are many potential explanations for the associations between depression and state anxiety during ART procedures and ART outcomes that could lead to variability across studies. Increases in anxiety or depression scores may be associated with other variables that are linked to negative outcomes. Future research needs to examine whether women with poor prognosis (e.g., older women, women with high BMI, smoking, previous unsuccessful IVF experiences, knowledge of negative test results or of poor responses, medical comorbidities) experience a greater association between state anxiety and depression and ART outcomes than women with better prognosis (see Van Loendersloot et al. 2010). Some of these relationships are complex and inter-related. For example, depression and anxiety are often comorbid with obesity and binge eating (Luppino et al., 2010; Nicholls et al., 2016). Obesity has also been found to negatively impact ART outcomes (Purewal et al. in press; Rittenberg et al., 2011; Metwally et al., 2007).

Alternatively, there is some suggestion, largely from animal studies, that anxiety and depression may be linked to changes in immune system function associated with miscarriage (for review, see Qu et al., 2017). It is possible that these factors vary across study populations or have different effects across different IVF treatment protocols, contributing to

heterogeneity. Moreover, the association between depression and anxiety scores during ART treatment and outcomes could arise because women who respond poorly to the stimulation cycle may experience higher levels of anxiety or depression to the knowledge of poor test results. For example, the number of oocytes retrieved predicts IVF outcomes (Stolwijk et al., 1996; Smeenk et al., 2000) a higher number of oocytes is associated with lower state anxiety and depression scores (Gourounti et al. (2011). Boivin & Takeman (1995) also reported that greater stress during ART treatment was significantly correlated with lower numbers of oocyte retrieved and embryos transferred. More research is necessary to tease out these relationships and understand the underlying mechanisms (Purewal et al. in press). Future research should adopt a more holistic approach that investigates how psychological variables interconnect with physiological factors. However, given the ethical issues that would need to be confronted to explore some of these factors (e.g. manipulating whether patients should be blind to their baseline results and to how they are responding to the treatment) it may remain difficult to fully map the factors underpinning these associations.

Whether anxiety and depression during ART are markers for other factors linked to poorer outcomes, or contributors to poorer outcomes, a conservative response to our findings might be to provide tailored psychosocial care for patients during different ART stages to help them manage the psychological and physical toil of undergoing certain aspects of treatment (i.e., receiving news of poor response). This is likely to have a positive influence, as a number of studies have examined the effect of psychological interventions on ART outcomes with generally positive results. For example, Hämmerli et al.'s (2009) meta-analysis found psychological interventions were effective at improving ART pregnancy rates but did not reduce depression or anxiety. Chow et al. (2016) recently found evidence that psychosocial

interventions improved psychological and pregnancy outcomes in their critical review. Frederiksen et al. (2015) performed a meta-analysis on 39 studies and found significant effects of psychosocial interventions on ART clinical pregnancy and that reductions in anxiety were associated with improvement in pregnancy rates. These findings indicate that attempts to manage and reduce psychological distress during ART can be successful in improving pregnancy outcomes. Recent research has further highlighted the importance of psychosocial adjustment in women who went through treatment successfully (Toscano & Montgomery, 2009) and unsuccessfully (Gameiro & Finnigan, 2017). The psychological treatment and support needs of infertile patients who fail to fulfil their goal of parenthood has been previously described (Boivin et al, 2005). A recent systematic review and meta-analysis of the longer term mental health of infertile patients who failed to become parents emphasizes the need for appropriately tailored psychosocial support for those individuals who eventually relinquish their parenthood goals (Gameiro and Finnigan, 2017).

The small number of studies included in this review is a limitation, particularly in the data analyses examining changes in levels of state anxiety and depression and its association with ART outcomes. So, we cannot say with authority yet, whether changes in levels of anxiety or depression are associated with ART outcomes. Further, only a small number of studies included first time ART patients (n=5), with all the others including a mix of first time ART patients and patients who have had previous unsuccessful treatment. This is a shortcoming because rates of depression and anxiety are known to increase after ART treatment failure (Milazzo et al, 2016).

In conclusion, depression and state anxiety during ART treatment are associated with poor ART outcomes, but there is no evidence that changes in the levels of anxiety and depression from baseline to during ART treatment are associated with ART outcomes. However, the numbers of studies were small and more detailed empirical research is necessary to make a more definitive review. These findings help to frame future research questions and investigations and could help target psychological support during different stages of treatment. However, more detailed empirical research is necessary, which measures women's psychological functioning during the course of treatment and its association with ART outcomes.

Funding

This research was funded by the British Academy small grant award (SG100026).

Declaration of interests

The authors have no interests to declare.

Figure legends

Figure 1- Figure 1: Flow Diagram of studies included in the meta-analysis.

Figure 2 – Forest plot of depression scores during ART shown as standardised mean differences between pregnant and not pregnant women. (d= -0.302; 95% CI: -0.551 - -0.054, z = -2.387, p = 0.017; I^2= 77.142%, p = 0.001).

Figure 3- Forest plot of state anxiety scores during ART shown as standardised mean differences between pregnant and not pregnant women. (d= -0.335; 95% CI: -0.582 - -0.087: z=-2.649, p=0.008; I^2 =81.339%, p = 0.001).

Figure 4 – Forest plot of changes in depression scores from baseline to during ART shown as standardised mean differences between pregnant and not pregnant women. (d=-0.106; 95% CI: -0.296 - 0.085, z = -1.088; I^2= 0.00%).

Figure 5 – Forest plot of changes in state anxiety scores from baseline to during ART shown as standardised mean differences between pregnant and not pregnant women. (d=-0.056; 95% CI: -0.195 - 0.082, z = -0.794; I^2= 0.00%).

Key message

Depression and state anxiety during ART treatment are associated with poor ART outcomes, but there is no evidence that changes in the levels of anxiety and depression from baseline to during ART treatment affect ART outcomes. Depression and state anxiety during ART may have a stronger effect on ART outcomes than baseline depression/anxiety.

References

An, Y., Wang, Z., Ji, H., Zhang, Y. and Wu, K., 2011. Pituitary-adrenal and sympathetic nervous system responses to psychiatric disorders in women undergoing in vitro fertilization treatment. Fertil. Seril, 6, 404-408.

An, Y., Sun, Z., Li, L., Zhang, Y. and Ji, H., 2013. Relationship between psychological stress and reproductive outcome in women undergoing in vitro fertilization treatment: psychological and neurohormonal assessment. J. Assist. Reprod. Genet.,30, 35-41

Boivin, J. and Takefman, J.E., 1995. Stress level across stages of in vitro fertilization in subsequently pregnant and nonpregnant women. Fertil. Steril., 64, 802-810.

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

Boivin, J., E. Griffiths, C.A. Venetis, 2011. Emotional distress in infertile women and failure of assisted reproductive technologies: meta-analysis of prospective psychosocial studies. BMJ, 342, d223.

Borenstein, M., Hedges, L., Higgins, J., Rothstein, H. 2005. Comprehensive meta-analysis version 2. Englewood, NJ: Biostat, 104.

Chow, K. M., Cheung, M. C., I. K. Cheung, 2016. Psychosocial interventions for infertile couples: a critical review. J. Clin. Nurs., 15, 2101-2113.

Deeks, J., Bossuyt, P., Gatsonis, C., 2009. Cochrane handbook for systematic reviews of diagnostic test accuracy version 1.0. 0. The Cochrane Collaboration.

Duval, S., Tweedie, R., 2000. Trim and fill: a simple funnel‐plot‐based method of testing and adjusting for publication bias in meta‐analysis. Biometrics, 56, 455-63.

Eugster, A., Vingerhoets, A.J., 1999. Psychological aspects of in vitro fertilization: a review. Soc. Sci. Med, 48, 575–89.

Farquhar, C., Rishworth, J.R., Brown, J., Nelen, W.L., Marjoribanks, J., 2015. Assisted reproductive technology: an overview of Cochrane Reviews. The Cochrane Library. 2015 Jan 1.

Frederiksen, Y., Farver-Vestergaard, I., Skovgård, N.G., Ingerslev, H.J., Zachariae, R., 2015. Efficacy of psychosocial interventions for psychological and pregnancy outcomes in infertile women and men: a systematic review and meta-analysis. B.M.J open, 5, .e006592.

Gameiro, S., Boivin, J., Peronace, L. and Verhaak, C.M., 2012. Why do patients discontinue fertility treatment? A systematic review of reasons and predictors of discontinuation in fertility treatment. Hum. Reprod. update, 18, 652-669.

Gameiro, S. and Finnigan, A., 2017. Long-term adjustment to unmet parenthood goals following ART: a systematic review and meta-analysis. Hum Reprod Update, 23, 322-337.

Gourounti, K., Anagnostopoulos, F., Vaslamatzis, G., 2011. The relation of psychological stress to pregnancy outcome among women undergoing in-vitro fertilization and intracytoplasmic sperm injection. Women. Health, 51, 321-339.

Gressier, F., Letranchant, A., Cazas, O., Sutter-Dallay, A.L., Falissard, B., Hardy, P., 2015. Post-partum depressive symptoms and medically assisted conception: a systematic review and meta-analysis. Hum Reprod, 30, 2575-86.

Gürhan, N., Atici, D., Akyüz, A. and Kisa, S., 2009. Association of depression and anxiety with oocyte and sperm numbers and pregnancy outcomes during in vitro fertilization treatment. Psychol. Rep., 104, 796-806.

Hämmerli, K., Znoj, H., and J. Barth., 2009. The efficacy of psychological interventions for infertile patients: a meta-analysis examining mental health and pregnancy rate. Hum. Reprod. Update., 15, 279-295.

Hashemi, S., Simbar, M., Ramezani-Tehrani, F., Shams, J., Majd, H.A., 2012. Anxiety and success of in vitro fertilization. Eur. J. Obstet. Gynecol. Reprod. Biol, 164,60-64.

Li, X.H., Ma, Y.G., Geng, L.H., Qin, L., Hu, H. and Li, S.W., 2011. Baseline psychological stress and ovarian norepinephrine levels negatively affect the outcome of in vitro fertilisation. Gynecol. Endocrinol, 27, 139-43.

Lintsen, A.M.E., Verhaak, C.M., Eijkemans, M.J.C., Smeenk, J.M.J. and Braat, D.D.M., 2009. Anxiety and depression have no influence on the cancellation and pregnancy rates of a first IVF or ICSI treatment. Hum. Reprod, 24, 1092-8.

Van Loendersloot, L.L., Van Wely, M., Limpens, J., Bossuyt, P.M.M., Repping, S., Van Der Veen, F., 2010. Predictive factors in in vitro fertilization (IVF): a systematic review and meta-analysis. Hum reprod update, 16, 577-589.

Luppino, F.S., de Wit, L.M., Bouvy, P.F., Stijnen, T, Cuijpers, P, Penninx, B.W., Zitman, F.G., 2010. Overweight, obesity, and depression: a systematic review and meta-analysis of longitudinal studies. Arch. Gen. Psychiatry., 1;220-9.

Karlidere, T., Bozkurt, A., Ozmenler, K.N., Ozsahin, A., Kucuk, T., Yetkin, S., 2008. The influence of emotional distress on the outcome of in-vitro fertilization (IVF) and/or intracytoplasmic sperm injection (ICSI) treatment among infertile Turkish women. Isr. J. Psychiatry. Relat. Sci, 45, 55.

Kushnir, V.A., Barad, D.H., Albertini, D.F., Darmon, S.K. and Gleicher, N., 2017. Systematic review of worldwide trends in assisted reproductive technology 2004–2013. Reprod. Biol. Endocrinol. 15, 6.

Maheshwari, A., Bhattacharya, S., Johnson, N.P., 2008. Predicting fertility. Hum. Fertil. (Camb), 11, 109-17.

Matthiesen, S.M.S., Frederiksen, Y., Ingerslev, H.J. and Zachariae, R., 2011. Stress, distress and outcome of assisted reproductive technology (ART): a meta-analysis. Hum. Reprod, 26, 2763-76.

Metwally, M., T.C. Li, Ledger, W.L., 2997. The impact of obesity on female reproductive function. Obes. Rev., 8,515-23.

Milazzo, A., Mnatzaganian, G., Elshaug, A. G., Hemphill, S. A., Hiller, J. E., & Astute Health Study Group, 2016. Depression and Anxiety Outcomes Associated with Failed Assisted Reproductive Technologies: A Systematic Review and Meta-Analysis. PloS. one, *11,* e0165805.

Moher, D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group, 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ, b2535

Nicholls, W., Devonport, T.J., Blake, M, 2016.The association between emotions and eating behaviour in an obese population with binge eating disorder. Obes. Rev., 17,30-42.

Purewal, S., Chapman, S., van den Akker, O.A.B, 2017a. A systematic review and meta-analysis of psychological predictors of successful assisted reproductive technologies. BMC Res Notes,10,711.

Purewal, S., Chapman, S., van den Akker, O.A.B, 2017b. A systematic review and meta-analysis of lifestyle and body mass index predictors of successful assisted reproductive technologies. J Psychosom Obstet Gynaecol, 1-17.

Qu, F., Wu, Y., Zhu, Y.H., Barry, J., Ding, T., Baio, G., Muscat, R., Todd, B.K., Wang, F.F. and Hardiman, P.J., 2017. Sci Rep,7.

Rittenberg, V., Seshadri, S., Sunkara, S. K., Sobaleva, S., Oteng-Ntim, E., El-Toukhy, T., 2011. Effect of body mass index on IVF treatment outcome: an updated systematic review and meta-analysis. Reprod. Biomed. Online., 23, 421-39.

Saravelos, S.H., Wong, A.W., Kong, G.W., Huang, J., Klitzman, R., Li, T.C., 2016. Pain during embryo transfer is independently associated with clinical pregnancy in fresh/frozen assisted reproductive technology cycles. J. Obstet. Gynaecol. Res, 42, 684-693.

Smeenk, J.M.J., Stolwijk, A.M., Kremer, J.A.M. Braat, D.D.M., 2000. External validation of the Templeton model for predicting success after IVF. Hum. Reprod, 15, 1065-1068.

Sterne, J. A., & Egger, M., 2001. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. J. Clin. Epidemiol., 54, 1046-55.

Stolwijk, A.M., Zielhuis, G.A., Hamilton, C.J.C.M., Straatman, H.M.P.M., Hollanders, J.M.G., Goverde, H.J., Van Dop, P.A., Verbeek, A.L.M., 1996. Pregnancy: Prognostic models for the probability of achieving an ongoing pregnancy after in-vitro fertilization and the importance of testing their predictive value. Hum. Reprod., 11, 2298-2303.

Stroup, D.F., Berlin, J.A., Morton, S.C., Olkin, I., Williamson, G.D., Rennie, D., Moher, D., Becker, B.J., Sipe, T.A. and Thacker, S.B., 2000. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Jama., 283, 2008-2012.

Taguchi, S., T. Hayashi, Y. Tada, K. Kitaya, M. Funabiki, Y. Iwaki, M. Karita, Nakamura, Y, 2015. Do combined psychological stress examinations predict pregnancy outcome in an assisted reproductive technology program? Clin. Exp. Obstet. Gynecol,42,309-10.

Toscano, S.E., Montgomery, R.M., 2009. The lived experience of women pregnant (including preconception) post in vitro fertilization through the lens of virtual communities. Health Care Women Int, 30, 1014-1036.

Turner, K., Reynolds-May, M.F., Zitek, E.M., Tisdale, R.L., Carlisle, A.B. and Westphal, L.M., 2013. Stress and anxiety scores in first and repeat IVF cycles: a pilot study. PLoS. One, 8, e63743.

Wells, G. A., Shea, B., O'Connell, D., Peterson, J., Welch, V., Losos, M., Tugwell, P. 2009. The Newcastle-Ottawa Scale (NOS) for assessing the quality if nonrandomized studies in meta-analyses. 2009. Epub Available from: URL: http://www. ohri. ca/programs/clinical_epidemiology/oxford. htm [cited 2009 Oct 19].

Zaig, I., F. Azem, S. Schreiber, Y. Gottlieb-Litvin, Meiboom, H., Bloch, M, 2012. Women's psychological profile and psychiatric diagnoses and the outcome of in vitro fertilization: Is there an association? Arch. Womens. Ment. Health,15, 353-9.

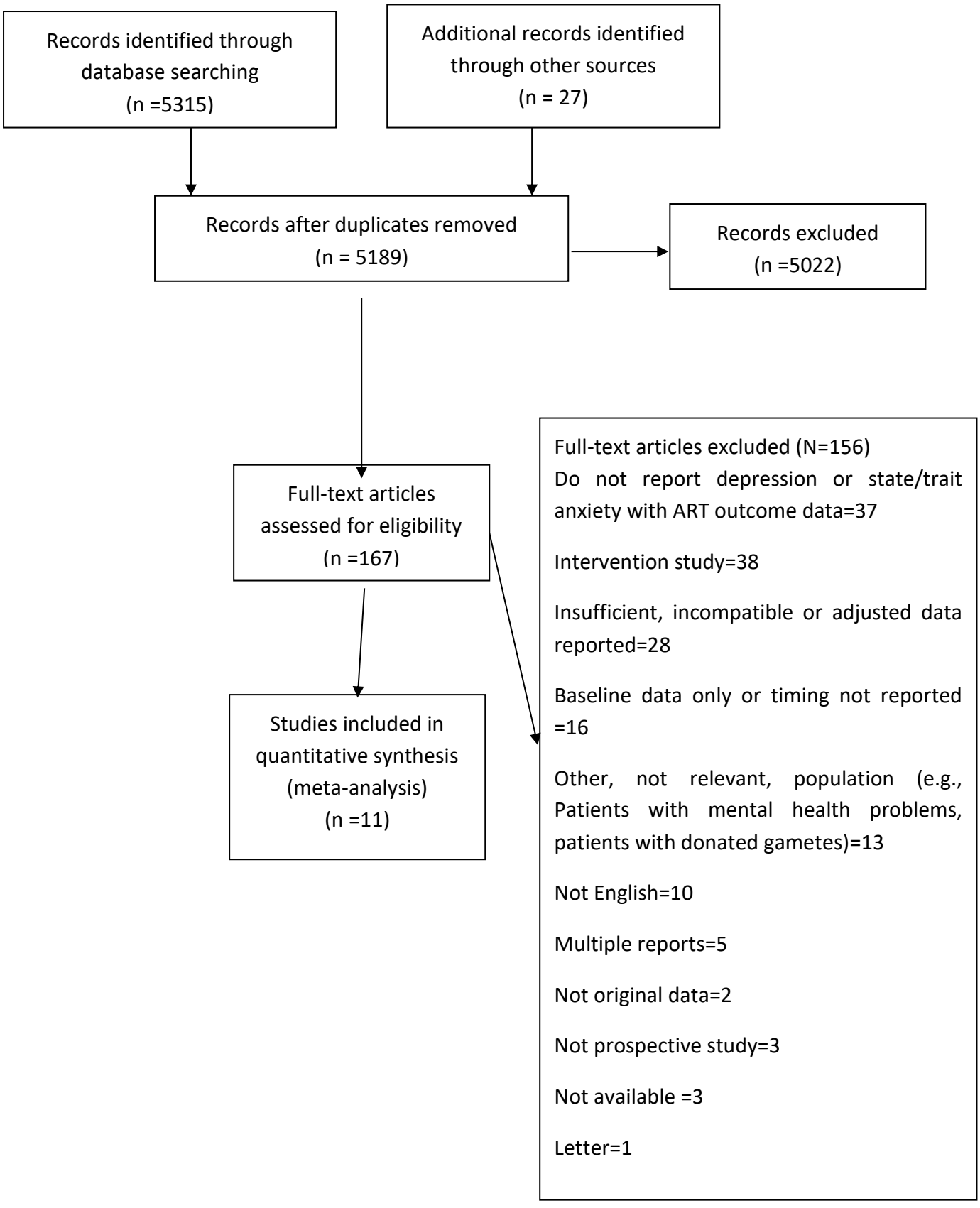


Figure 1: PRISMA 2009 Flow Diagram of studies included in the meta-analysis

Table1: Study characteristics

Authors and country	Psychological variable &Measurement	Time of assessment and period of enrolment	Method of **pregnancy** diagnosis	Study Design and sample size	Treatment	Main findings	Newcastle-Ottawa Quality Score
1. **An et al.2011. China.**	State Anxiety-*STAI* Depression –*BDI*	Baseline and day of oocyte retrieval Period of enrolment – 2009-2010	ultrasound scan	Prospective, cohort study N=264. All first time ART users.	IVF,ICSI. Data from single treatment cycle.	**<u>At baseline</u>** **Depression (NS)** Pregnant (n=92, 1.5±1.3) Not pregnant (n=172, 1.6±1.5). **State anxiety (NS)** Pregnant (n=92, 36.1±8.8) Not pregnant (n=172, 37.6±10.0). **<u>At day of oocyte retrieval</u>** **Depression (NS)** Pregnant (n=92, 1.6±1.6) Not pregnant (n=172, 1.9±18). **State anxiety (NS)**	Selection **** Comparability ** Outcome ***

						Pregnant (n=92, 38.7±6.7) Not pregnant (n=172, 39.7±7.4).	
2. **Gourounti et al 2011. Greece**	State Anxiety- *STAI* Depression- *Centre for Epidemiological Studies Depression Scale (CES-D)*	Before oocyte retrieval Period of enrolment November 2008 and July 2009	ultrasound scan	Prospective, cohort study N=160. Mix of first and multiple time ART users.	IVF,ICSI Data from single treatment cycle.	**<u>Before oocyte retrieval</u>** **Depression (p=<.001)** Pregnant (n=41, 7.9±6.9) Not pregnant (n=119, 15.0±9.6). **State anxiety (p=<.001)** Pregnant (n=41, 33.7±7.3) Not pregnant (n=119, 43.5±9.7).	Selection**** Comparability ** Outcome ***
3. **Gurhan et al 2009. Turkey.**	Depression –*BDI* State Anxiety - *STAI*	Baseline and day of oocyte retrieval Period of enrolment	HCG test	Prospective, cohort study N=80	IVF First cycle of treatment.	**<u>At baseline</u>** **Depression (p=<.05)** Positive test (n=39, 10.1±5.2) Negative test **(n=41, 12.7±6.2).**	Selection **** Comparability Outcome **

		- September 2004 and July 2005		All first time ART users.	Data from single treatment cycle.	**State anxiety (NS)** Positive test (n=39, 45.0±4.6) Negative test (n=41, 47.6±7.8). **<u>At day of oocyte retrieval</u>** **Depression (p=.01)** Positive test (n=39, 15.2±5.7) Negative test (n=41, 19±7.3). **State anxiety (p=.01)** Positive test (n=39, 51.7±5.1) Negative test (n=41, 55.0±5.5).	
4. **Hashemi et al 2012. Iran.**	State-anxiety - *STAI*	Before oocyte retrieval Period of enrolment	HCG test	Prospective, cohort study N=180. Mix of first and	IVF, ICSI, ZIFT Data from single	**<u>Before oocyte retrieval</u>** **State anxiety (NS)** Positive test (n=19, 46.32±10.8)	Selection *** Comparability ** Outcome **

		not reported.		multiple time ART users.	treatment cycle.	Negative test (n=161, 47.45±10.6).	
5. **Karlidere et al 2008. Turkey.**	Depression –*BDI* State Anxiety - *STAI*	Day before embryo transfer Period of enrolment June 2001 to July 2003	Ultrasound scan	Prospective, cohort study N=104 All first time ART users.	IVF, ICSI Data from single treatment cycle.	**<u>Day before embryo transfer</u>** **Depression (p=.001)** Pregnant (n=49, 6.21±4.34) Not pregnant (n=55, 10.55±5.49). **State anxiety (p=0.001)** Pregnant (n=49, 33.21±7.91) Not pregnant (n=55, 40.14±8.37).	Selection **** Comparability** Outcome ***
6. **Li et al 2011. China**	State Anxiety – *Zung self rating anxiety scale (SAS)*	Day of oocyte retrieval.	Ultrasound scan	Prospective, cohort study N=107	IVF Data from single	**<u>At baseline</u>** **Depression (NS)** Pregnant (n=50, 52.66±12.34) Not pregnant (n=57, 54.06±11.34).	Selection **** Comparability ** Outcome ***

	Depression – *Zung self rating depression scale (SDS)*	Period of enrolment 2007-2008.		All first time ART users.	treatment cycle.	**State anxiety (NS)** Pregnant (n=50, 39.66±8.04) **<u>At day of oocyte retrieval</u>** **Depression (NS)** Pregnant (n=50, 55.27±9.18) Not pregnant (n=57, 56.39±10.93). **State anxiety (NS)** Pregnant (n=50, 42.66±7.12) Not pregnant (n=57, 41.96±9.23).	
7. **Lintsen et al 2009. The Netherlands.**	State Anxiety - *STAI*	Day before oocyte retrieval. Period of enrolment 2002-2004.	Ultrasound scan	Prospective, cohort study N=690 (at baseline), of which n=413had completed questionnaire on day	IVF,ICSI Data from single treatment cycle.	**<u>Baseline</u>** **Depression (NS)** Pregnant (n=196, 1.2±1.8) Not pregnant (n=494, 1.4±2.4). **State anxiety (NS)** Pregnant (n=196, 176±4.7)	Selection **** Comparability ** Outcome **

				before oocyte retrieval. All first time ART users.		Not pregnant (n=494, 17.7±5.0). **<u>Day before oocyte retrieval</u>** **State anxiety (NS)** Pregnant (n=122, 18.4±5.8) Not pregnant (n=291, 18.5±5.8).	
8. **Saravelos et al 2016. Hong Kong.**	Depression –*BDI* State Anxiety - *STAI*	Day of oocyte retrieval. Period of enrolment 2011-2014.	Ultrasound Scan.	Prospective, cohort study N=360 First time ART users or multiple users not specified.	IVF, ICSI Data from multiples treatment cycle.	**<u>Day of oocyte retrieval</u>** **Depression (NS)** Pregnant (n=175, 7.8±8.2) Not pregnant (n=185, 7.5±7.4). **State anxiety (NS)** Pregnant (n=175, 55.1±10) Not pregnant (n=185, 54.8±8.6).	Selection **** Comparability ** Outcome ***
9. **Taguchi et al 2015. Japan.**	*Zung self rating depression scale (SDS)*	Day of embryo transfer.	HCG test	Prospective, cohort study N=113	IVF Data from	**<u>Day of embryo transfer</u>** **Depression (NS)**	Selection *** Comparability **

		Period of enrolment April 2012 to May 2012.		First time ART users or multiple users not specified.	single treatment cycle.	Pregnant (n=36, 37.2±6.3) Not pregnant (n=77, 36.7±6.8).	Outcome *
10. **Turner et al 2013. US.**	State Anxiety - *STAI*	Day before oocyte retrieval. Period of enrolment June 2009-September 2009.	ultrasound scan	Prospective, cohort study. N=36 included in baseline sample and n=35 included in day before oocyte retrieval sample. First or second time ART users.	IVF Data from single treatment cycle.	**<u>Baseline</u>** **State anxiety (NS)** Pregnant (n=15, 37.53±12.33) Not pregnant (n=21, 43.57±14.44). **<u>Day before oocyte retrieval</u>** **State anxiety (P=0.05)** Pregnant (n=15, 34.93±11.18) Not pregnant (n=20, 44.35±13.63).	Selection *** Comparability Outcome **

11. **Zaig et al 2012. Israel.**	Depression - *CESD Center for Epidemiologic Studies Depression scale* State Anxiety - *STAI*	At ovulation induction Period of enrolment January 2006 to December 2007.	Ultrasound scan	Prospective, cohort study N=108 First or second time ART users.	IVF Data from single treatment cycle.	**<u>At ovulation induction</u>** **Depression (NS)** Pregnant (n=45, 34.06±9.4) Not pregnant (n=63, 34.93±9.47. **State anxiety (NS)** Pregnant (n=45, 42.42±11.4) Not pregnant (n=63, 44.07±11.79).	Selection **** Comparability Outcome ***

Note: BDI – Beck's Depression Inventory; ICSI = intracytoplasmic sperm injection; IVF = in vitro fertilisation; NS= non-significant differences between women who were pregnant and women who were not pregnant; p= value; STAI – State-Trait Anxiety Inventory-State scale. The sample size refers to data that is extracted from the papers and used in the meta-analysi

Table 2: Sensitivity analyses

	d [95% CI OR]	Heterogeneity (I^2)
DEPRESSION DURING ART TREATMENT		
Pregnancy diagnosed with ultrascan only (k=6)	-0.323 [-0.617,-0.029], z=-2.152, p=0.03	80.784%, p=0.001
First ART (k=5)	-0.357 [-0.637, 0.077], z=-2.498, p=0.013	66.010%, p=0.019
Single cycle only (k=7)	-0.363 [-0.628, -0.097], z=-2.678, p=0.007	72.610%, p=0.001
Only IVF (k=4)	-0.174 [-0.428, 0.080], z=-1.342, NS	37.932%, NS
Data collected before or during oocyte retrieval period (k=6)	-0.268 [-0.519, -0.018] z=-2.100, p=0.036	72.556%, p=0.003
High quality (k=6)	-0.323 [-0.617, -0.029] z=-2.152, p=0.031	80.784%, p=0.001
Recent only (k=6)	-0.175 [-0.408, 0.058], z=-1.470, NS	69.277%, p=0.006
STATE ANXIETY DURING ART TREATMENT		
Pregnancy diagnosed with ultrascan only (k=8)	-0.330 [-0.614, -0.046], z=-2.279: p=0.023	84.317%, p=0.001
First ART (k=5)	-0.280 [-0.588, 0.028], z=-1.075, NS	475.503%, p=0.001
Single cycle (k=9)	-0.388 [-0.666, -0.110], z=-2.737, p=0.006	80.793%, p=0.001
IVF only (k=4)	-0.332 [-0.731, 0.068], z=-1.625, NS	66.888%, p=0.029
Data collected before or during oocyte retrieval period (K=9)	-0.273 [-0.515, -0.032], z=-2.215, p=0.027	78.766%, p=0.001

High quality (k= 8)	-0.261 [-0.524, -0.003], z=-1.941, p=0.052	82.857%, p=0.001
Recent only [a] (k=7)	-0.280 [-0.587, 0.026], z=-1.793, NS	81.130%, p=0.001
CHANGES IN DEPRESSION FROM BASELINE TO DURING ART TREATMENT		
Pregnancy diagnosed with ultrascan only (k=2)	-0. 088 [-0.299, 0.123], z=-0.816, NS	<0.001%, NS
First ART (k=2)	-0. 088 [-0.299, 0.123], z=-0.816, NS	<0.001%, NS
Single cycle only (k=2)	-0. 088 [-0.299, 0.123], z=-0.816, NS	<0.001%, NS
Only IVF (k=2)	-0. 093 [-0.380, 0.195], z=-0.633, NS	<0.001%, NS
Data collected before or during oocyte retrieval period (k=3)	-0.106 [-0.296, 0.085], z = -1.088, NS	<0.00%, NS
High quality (k=2)	-0.088 [-0.299, 0.123], z = -1.816, NS	<0.001%, NS
Recent only [a] (k=2)	-0. 088 [-0.299, 0.123], z=-0.816, NS	<0.001%, NS
CHANGES IN STATE ANXIETY FROM BASELINE TO DURING ART TREATMENT		
Pregnancy diagnosed with ultrascan only (k= 4)	-0.048 [-0.194, 0.098], z=-0.941: NS	<0.001%, NS
First ART (k=3)	-0.031 [-0.180, 0.119], z=-0.401, NS	<0.001%, NS
Single cycle (k=4)	-0.048 [-0.194, 0.098], z=-0.641, NS	<0.001%, NS
IVF only (k=3)	-0.130 [-0.396, 0.136], z=-0.957, NS	<0.001%, NS

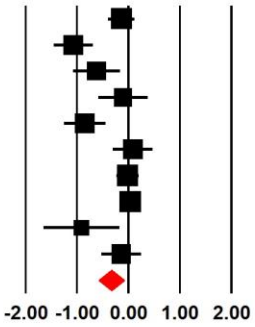
Data collected before or during oocyte retrieval period (K=5)	-0.056 [-0.195, 0.082], z = -0.794, NS	<0.00%, NS
High quality (k=3)	-0.031 [-0.180, 0.119], z=-0.401, NS	<0.001%, NS
Recent only [a] (k=3)	-0.091 [-0.293, 0.111], z=-0.887, NS	<0.001%, NS

[a]Studies published from 2010 onwards; NS = p value was not significant.

Study name	Std diff in means	Lower limit	Upper limit	Z-Value	p-Value
An et al 2011	-0.140	-0.393	0.114	-1.079	0.280
Gourounti et al 2011	-1.071	-1.445	-0.697	-5.614	0.000
Gurhan et al 2009	-0.622	-1.070	-0.173	-2.714	0.007
Hashemi et al 2012	-0.106	-0.582	0.369	-0.439	0.661
Karlidere et al 2008	-0.850	-1.252	-0.448	-4.143	0.000
Li et al 2011	0.084	-0.296	0.464	0.434	0.664
Lintsen et al 2009	-0.017	-0.229	0.194	-0.160	0.873
Saravelos et al 2016	0.032	-0.174	0.239	0.306	0.760
Turner et al 2013	-0.910	-1.639	-0.181	-2.446	0.014
Zaig et al 2012	-0.142	-0.525	0.241	-0.726	0.468
	-0.335	-0.582	-0.087	-2.649	0.008

Statistics for each study

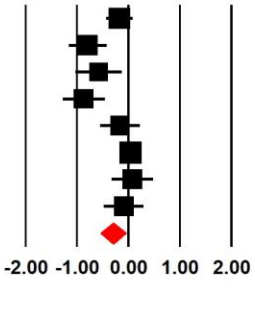
Std diff in means and 95% CI



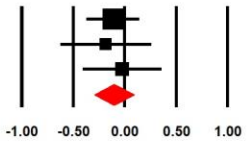
Study name	Std diff in means	Lower limit	Upper limit	Z-Value	p-Value
An et al 2011	-0.173	-0.427	0.080	-1.338	0.181
Gourounti et al 2011	-0.789	-1.155	-0.424	-4.236	0.000
Gurhan et al 2009	-0.578	-1.026	-0.131	-2.534	0.011
Karlidere et al 2008	-0.871	-1.274	-0.468	-4.239	0.000
Li et al 2011	-0.164	-0.544	0.217	-0.843	0.399
Saravelos et al 2016	0.038	-0.168	0.245	0.365	0.715
Taguchi et al 2015	0.075	-0.321	0.471	0.372	0.710
Zaig et al 2012	-0.092	-0.475	0.291	-0.472	0.637
	-0.302	-0.551	-0.054	-2.387	0.017

Statistics for each study

Std diff in means and 95% CI



Study name	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
An et al 2011	-0.115	0.129	0.017	-0.369	0.138	-0.893	0.372
Gurhan et al 2009	-0.183	0.224	0.050	-0.622	0.257	-0.815	0.415
Li et al 2011	-0.026	0.194	0.038	-0.405	0.354	-0.132	0.895
	-0.106	0.097	0.009	-0.296	0.085	-1.088	0.276



Study name	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
An et al 2011	-0.070	0.129	0.017	-0.323	0.183	-0.540	0.589
Gurhan et al 2009	-0.132	0.224	0.050	-0.571	0.307	-0.589	0.556
Li et al 2011	-0.041	0.194	0.038	-0.421	0.339	-0.211	0.833
Lintsen et al 2009	-0.000	0.108	0.012	-0.211	0.211	-0.000	1.000
Turner et al 2013	-0.429	0.358	0.128	-1.131	0.273	-1.197	0.231
	-0.056	0.071	0.005	-0.195	0.082	-0.794	0.427